Speeding up random walk mixing
by starting from a uniform vertex[*]

Alberto Espuny Díaz[†]   Patrick Morris[‡]   Guillem Perarnau[§]
Oriol Serra[¶]

Abstract

The theory of rapid mixing random walks plays a fundamental role in the study of modern randomised algorithms. Usually, the mixing time is measured with respect to the worst initial position. It is well known that the presence of bottlenecks in a graph hampers mixing and, in particular, starting inside a small bottleneck significantly slows down the diffusion of the walk in the first steps of the process. The average mixing time is defined to be the mixing time starting at a uniformly random vertex and hence is not sensitive to the slow diffusion caused by these bottlenecks.

In this paper we provide a general framework to show logarithmic average mixing time for random walks on graphs with small bottlenecks. The framework is especially effective on certain families of random graphs with heterogeneous properties. We demonstrate its applicability on two random models for which the mixing time was known to be of order $(\log n)^2$, speeding up the mixing to order $\log n$. First, in the context of smoothed analysis on connected graphs, we show logarithmic average mixing time for randomly perturbed graphs of bounded degeneracy. A particular instance is the Newman-Watts small-world model. Second, we show logarithmic average mixing time for supercritically percolated expander graphs. When the host graph is complete, this application gives an alternative proof that the average mixing time of the giant component in the supercritical Erdős-Rényi graph is logarithmic.

[†]Technische Universität Ilmenau, 98684 Ilmenau, Germany. E-mail: alberto.espuny-diaz@tu-ilmenau.de
[‡]Universitat Politècnica de Catalunya (UPC), Barcelona, Spain. E-mail: pmorrismaths@gmail.com
[§]Universitat Politècnica de Catalunya (UPC), Barcelona, Spain. E-mail: guillem.perarnau@upc.edu
[¶]Universitat Politècnica de Catalunya (UPC), Barcelona, Spain. E-mail: oriol.serra@upc.edu

Speeding up random walk mixing by starting from a uniform vertex

1 Introduction

Random walks on graphs are one of the fundamental tools for sampling (see, e.g., [38]). Applications are numerous in areas such as computer science, discrete mathematics and statistical physics. Prominent examples include the polynomial-time algorithm to estimate the volume of a convex body [19], computing the matrix permanent [28] or the use of Glauber dynamics to sample from Gibbs distributions, in particular from proper colourings [42].

Most usually, the size of the sampling space is exponential in the input size, and fully exploring this space is computationally intractable. The Markov chain Monte Carlo (MCMC) method consists of running a random walk in an appropriately chosen graph, whose vertex set is the sample space, until its distribution is arbitrarily close to equilibrium, regardless of the initial state. At that time we say the walk has mixed, and the time until it does is called the (worst-case) *mixing time*. To obtain efficient sampling algorithms it suffices to prove that the mixing time is poly-logarithmic in the input size.

The connection between rapid mixing and expanders is well-established. In the context of random walks, expansion is measured by means of a graph parameter called *conductance*; see Section 2.2 for the precise definition. Jerrum and Sinclair [28] gave an upper bound on the mixing time depending on the conductance and the logarithm of the minimum stationary value. This bound is central in the theory of Markov chains.

Random environments are particularly interesting sampling spaces and, in the last 20 years, researchers have developed the theory of random walks on random graphs. As expected, the good expansion properties of random graphs ensure rapid mixing. By the Jerrum-Sinclair bound, graphs with conductance bounded away from zero mix in logarithmically many steps and usually exhibit *cut-off*, that is, the distribution converges rapidly to the stationary distribution in a small window of time. Good examples are random graph models with control on the degrees, such as random regular graphs [34], random graphs with given degree sequences [6, 4], their directed analogues [9, 12], or graphs perturbed by random perfect matchings [27].

Nonetheless, the presence of small obstructions slows down the mixing. A canonical example is the giant component of a sparse Erdős-Rényi graph $G(n, c/n)$ with $c > 1$. This component contains relatively small *bottlenecks*, that is, connected sets that only have few edges connecting them to the rest of the graph. In such cases, tools like the Jerrum-Sinclair bound fail to pin down the correct order of the mixing time. Fountoulakis and Reed [23] introduced a strengthening of the bound that is sensitive to small bottlenecks and used it to show that the mixing time of the largest component in $G(n, c/n)$ is asymptotically almost surely (a.a.s. for short) $O(\log^2 n)$ [24]. Indeed, this is the correct order as the component contains paths of degree 2 vertices (also referred to as *bare paths*) whose length is of order $\log n$. Starting at the centre of such paths, a random walk takes $\Omega(\log^2 n)$ steps in expectation to escape from it. We remark that the mixing time in the supercritical random graph $G(n, c/n)$ was also bounded independently by Benjamini, Kozma and Wormald [5], using a different approach investigating the *anatomy* of the giant component.

However, these local bottlenecks are a negligible part of the giant component and the rest of the component has good expansion properties. This suggests that, if the random walk started outside the bottlenecks, the mixing time would decrease. This was implicit in the work of Benjamini, Kozma and Wormald [5] and their description of the giant component, and such a speeding up of mixing time was also conjectured explicitly by Fountoulakis and Reed [24]. Berestycki, Lubetzky, Peres and Sly [6] confirmed their prediction, showing that there exists $a = a(c)$ such that the mixing time starting at a uniformly random vertex is asymptotically $a \log n$ with high probability (they in fact

proved much more, establishing the value of $a(c)$ precisely as well as cut-off for the random walk). This result reinforces the idea that, in certain heterogeneous scenarios, averaging over the starting position yields more efficient sampling algorithms.

The goal of this paper is to provide a general framework to show logarithmic average-case mixing time for random walks on graphs with small bottlenecks.

1.1 Average mixing times

Given an n-vertex graph G, the *lazy random walk* over G is a Markov chain with state space $V(G)$ which can be defined as follows. If at any given time we are in a vertex $u \in V(G)$, the lazy random walk stays in u with probability $1/2$, and with probability $1/2$ it moves to a uniformly random neighbour of u in G. If G is a connected graph, it is well known that the lazy random walk over G is ergodic and its distribution converges to the (unique) stationary distribution π_G (see, e.g., [33] for a comprehensive review of random walks and mixing times).

The total variation distance $d_{\mathrm{TV}}(\mu, \nu)$ between two probability distributions ν and μ on the vertex set $V(G)$ of a graph G is defined as

$$d_{\mathrm{TV}}(\mu, \nu) := \max_{A \subseteq V(G)} |\mu(A) - \nu(A)| = \frac{1}{2} \sum_{v \in V(G)} |\mu(v) - \nu(v)|. \tag{1.1}$$

Let P_G be the transition matrix of the lazy random walk over G. For $\varepsilon > 0$, the *ε-mixing time* $t_{\mathrm{mix}}(G, \varepsilon)$ of this lazy random walk is defined as

$$t_{\mathrm{mix}}(G, \varepsilon) := \min \left\{ t \in \mathbb{N}_0 : \max_{u \in V(G)} d_{\mathrm{TV}}(\mu_0^u P_G^t, \pi_G) \leq \varepsilon \right\},$$

where μ_0^u is the distribution supported entirely on $u \in V(G)$.

If instead of considering the worst-case initial vertex we consider a uniformly random vertex $v \in V(G)$, then the quantity $d_{\mathrm{TV}}(\mu_0^v P_G^t, \pi_G)$ is a random variable. We define the *average ε-mixing time* $\bar{t}_{\mathrm{mix}}(G, \varepsilon)$ of the lazy random walk, to be the time at which the expectation of this random variable falls below the ε. That is,

$$\bar{t}_{\mathrm{mix}}(G, \varepsilon) := \min \left\{ t \in \mathbb{N}_0 : \frac{1}{n} \sum_{u \in V(G)} d_{\mathrm{TV}}(\mu_0^u P_G^t, \pi_G) \leq \varepsilon \right\}.$$

Remark 1.1. In this work, we will focus on the quantity $\bar{t}_{\mathrm{mix}}(G, \varepsilon)$, which we believe is a natural candidate for tracking mixing times starting from a uniform vertex. Nonetheless, other related quantities have been used to measure the mixing time from a uniform starting point.

Indeed, for a vertex $u \in V(G)$, define

$$t_{\mathrm{mix}}^{(u)}(G, \varepsilon) := \min \left\{ t \in \mathbb{N}_0 : d_{\mathrm{TV}}(\mu_0^u P_G^t, \pi_G) \leq \varepsilon \right\}$$

and consider the random variable $t_{\mathrm{mix}}^{(U_n)} = t_{\mathrm{mix}}^{(U_n)}(G, \varepsilon)$, where U_n is a vertex chosen uniformly at random from $V(G)$. This notion was the one studied by Berestycki, Lubetzky, Peres and Sly [6]. It is natural to compare $\bar{t}_{\mathrm{mix}}$ to $\mathbb{E}(t_{\mathrm{mix}}^{(U_n)})$: in the first case, we average the total variation distance over starting vertices and take the smallest time t when this average is smaller than ε; in the second one, we average the mixing times over the starting vertices (see Figure 1). In general as functions, neither of these notions is stronger than the other, in that one can design examples of trajectories for total variation distances $d_{\mathrm{TV}}(\mu_0^u P_G^t, \pi_G)$ for different vertices u, showing that $\bar{t}_{\mathrm{mix}}$ cannot be bounded by a function of $\mathbb{E}(t_{\mathrm{mix}}^{(U_n)})$ and vice versa. However, bounding either $\mathbb{E}(t_{\mathrm{mix}}^{(U_n)})$ or $\bar{t}_{\mathrm{mix}}$ implies

that $t_{\mathrm{mix}}^{(U_n)}$ is small with high probability. In the first case this is a direct application of Markov's inequality. In the second one, define $d_u(t) := d_{\mathrm{TV}}(\mu_0^u P_G^t, \pi_G)$, for a vertex u, then $\bar{t}_{\mathrm{mix}}(G, \varepsilon)$ is the time t at which the expected value of $d_u(t)$ (averaged over starting points) is less than ε. By Markov's inequality, $d_{U_n}(\bar{t}_{\mathrm{mix}}(G, \varepsilon^2)) \leq \varepsilon$ with probability at least $1 - \varepsilon$.
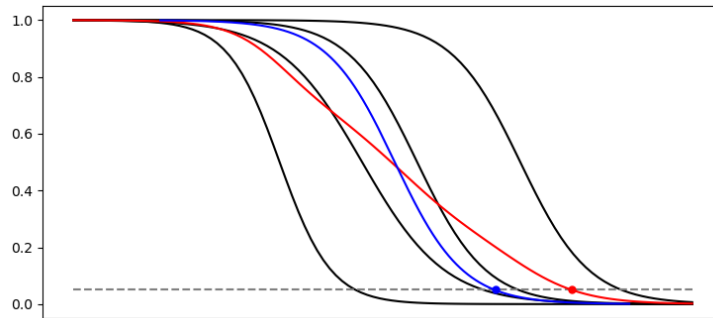

Figure 1: Schematic plot of the total variation distance starting at different vertices and the two average mixing times for $\varepsilon = 0.05$. In red, the function $\frac{1}{n} \sum_{u \in V(G)} d_{\mathrm{TV}}(\mu_0^u P_G^t, \pi_G)$ and the dot representing $\bar{t}_{\mathrm{mix}}(G, \varepsilon)$. In blue, the average of mixing times at different thresholds and the dot representing $\mathbb{E}(t_{\mathrm{mix}}^{(U_n)}(G, \varepsilon))$.

A related but very different notion is the time it takes to mix starting at μ_V, the uniform distribution over V:

$$t_{\mathrm{mix}}^{(V)}(G, \varepsilon) := \min \left\{ t \in \mathbb{N}_0 : d_{\mathrm{TV}}(\mu_V P_G^t, \pi_G) \leq \varepsilon \right\}.$$

A similar notion has been studied for directed graphs, where the initial distribution is the in-degree one; see, e.g., [9, Theorem 3]. In general, this latter notion of average mixing time is much smaller than the previous notions and we expect this to also be the case in the settings studied here, although we do not explore this direction.

Remark 1.2. In the literature, the *mixing time* of the random walk is often defined as $t_{\mathrm{mix}}(G) := t_{\mathrm{mix}}(G, 1/4)$, since the distance to the stationary distribution is contractive after this time. However, this might not be the case for $\bar{t}_{\mathrm{mix}}$. Consider for instance the lollipop graph $L_{n,k}$: a clique on k vertices and a path on $n - k$ vertices joined by an edge incident to one of the endpoints of the path. If k and $n - k$ are both very large, then, after one step, the total variation distance is roughly 0 if we start at the clique (almost all the mass of $\pi_{L_{n,k}}$ is supported on the clique), and roughly 1 if we start at the path. Taking $k = \lceil \alpha n \rceil$, then

$$\frac{1}{n} \sum_{u \in V(G)} d_{\mathrm{TV}}(\mu_0^u P_{L_{n,k}}, \pi_{L_{n,k}}) \sim 1 - \alpha. \tag{1.2}$$

If $\alpha > 3/4$, then $\bar{t}_{\mathrm{mix}}(L_{n,k}, 1/4) = 1$. However, the time required to further decrease the distance to the stationary distribution is of order $\Omega(n^2)$, as this is the time required for the walk starting at a typical vertex in the path to hit the clique.

1.2 Our results

Our results will apply to graphs satisfying certain natural structural conditions, which we formalise in the following definition.

Definition 1.3. *Let G be an n-vertex graph. For $\alpha > 0$, we say that a set $S \subseteq V(G)$ is α-thin in G if*

$$|\partial_G(S)| := e_G(S, V(G) \setminus S) < \alpha|S|.$$

For $D > 0$, we say that a set $S \subseteq V(G)$ is D-loaded in G if

$$e_G(S) > D|S|.$$

We say that G is an (α, D)-spreader graph if it satisfies the following three properties:

(S1) *For all $(\log n)^{1/5} \leq k \leq (1 - 1/D^2)n$, the number of G-connected α-thin sets $S \subseteq V(G)$ with $|S| = k$ is less than $n\mathrm{e}^{-\sqrt{k}}$.*

(S2) *For all $(\log n)^{1/5} \leq k \leq (1 - 1/D^2)n$, the number of G-connected α^{-1}-loaded sets $S \subseteq V(G)$ with $|S| = k$ is less than $n\mathrm{e}^{-\sqrt{k}}$.*

(S3) *No set $S \subseteq V(G)$ with $|S| \geq \alpha n$ is D-loaded in G.*

Note that, for $k > \log^2 n$, one has that $n\mathrm{e}^{-\sqrt{k}} < 1$, and thus the conditions on an n-vertex graph G being an (α, D)-spreader graph guarantee that there are no G-connected vertex subsets of size between $\log^2 n$ and $(1 - 1/D^2)n$ that have too few edges leaving the set (S1) or too many edges contained inside the set (S2). These pseudo-random conditions on expansion and edge distribution arise naturally in the context of random graph models. Indeed, the density of a random graph within any vertex set and across any vertex partition is expected to be the same as the density of the whole graph, and concentration inequalities in conjunction with union bounds can be used to derive the non-existence of such bad connected vertex sets with high probability. Moreover, the conditions of Definition 1.3 bound the number of bad vertex sets of size between $(\log n)^{1/5}$ and $\log^2 n$, with exponential decay as one can expect from concentration inequalities on binomial random variables. In the context of the current work, the conditions on spreader graphs will guarantee that all bottlenecks are small and they are scarce in the graph.

To digest the notion of spreader graphs, one can think of $\alpha > 0$ as an arbitrarily small constant and D as arbitrarily large. The parameter $\alpha > 0$ controls conditions (S1) and (S2) in the sense that, as α shrinks, these conditions become easier to satisfy and thus the definition of spreader graphs captures more graphs. Similarly, the parameter D controls (S3) and imposes in particular that the spreader graphs are sparse with bounded average degree. It should be noted that, due to D appearing in (S1) and (S2) and α appearing in (S3), our definition is not actually monotone in these parameters. This is a technical subtlety that is needed in our proof to guarantee a trade-off between the conditions. However, in all applications, the restraints given by D in (S1) and (S2) and α in (S3) are never critical, as we have very good control over the edge distribution in all linear sets.

We also remark that the constant $1/5$ could be replaced by any constant $\zeta < 1/4$. Indeed, for sets smaller than $(\log n)^\zeta$, we impose no restriction. The point is that, as we will focus on connected spreader graphs G, even if a small set is an extreme bottleneck, the random walk will not get stuck there for too long before exploring the set enough to escape. In our proof, these bottlenecks contribute a factor $(\log n)^{4\zeta}$ (due to a connected set of size s having *conductance* at least $1/s^2$ and Theorem 2.1 giving a quadratic dependence on conductance, see Section 2.2 for details), hence a choice of $\zeta < 1/4$ guarantees that this contribution is negligible. It may be possible to replace this constraint of $1/4$ by $1/2$ but not beyond this.

Finally, we remark that the constants α and D could be replaced with functions that depend on n and the definition of spreader graphs could be adjusted so that our main

theorem would still give bounds on average mixing times. However, as our focus is on sparse graphs with constant average degree, we do not pursue this direction here.

Remark 1.4. The definition of (α, D)-spreader graphs bears resemblance with that of α-*AN graphs* (or α-*decorated expanders*) introduced in [5]. An α-AN graph G is defined in terms of the existence of an expander subgraph B whose complement is formed by a small number of small components, similar to what can be deduced from (S1)-(S3), and additionally requiring that not too many components of $G - B$ are connected to each $v \in V(B)$. The backbone of the main result in [5] is to show that random walks on α-AN graphs mix in $O(\log^2 n)$ steps.

Our main theorem provides a tool to prove logarithmic average mixing time for (α, D)-spreader graphs.

Theorem 1.5. *For all $\varepsilon > 0$, $D \geq 4$ and $0 < \alpha < 1/D^2$, there exists a $C > 0$ such that the following holds for all n sufficiently large. Suppose G is an n-vertex connected (α, D)-spreader graph. Then,*

$$\bar{t}_{\mathrm{mix}}(G, \varepsilon) \leq C \log n.$$

We believe that in many cases, as in our two applications below, this theorem can be used to quickly derive optimal bounds for average mixing times in settings where worst-case mixing times are established via conductance bounds.

The proof of Theorem 1.5 bears some similarities with the proof in [6]. Both use the idea of contracting badly connected sets and coupling the random walks in the original and the contracted graphs. However, our proof is conceptually simpler as it does not use the anatomy of the giant component [16], a powerful description of the largest component in the supercritical regime. Instead, we rely on the Fountoulakis-Reed bound for mixing [23] and recent progress on hitting time lemmas [35].

1.3 Application 1: Smoothed analysis on connected graphs

The idea of studying the effect of random perturbations on a given structure arose naturally in several distinct settings. In theoretical computer science, Spielman and Teng [40] (see also [41]) introduced the notion of *smoothed analysis* of algorithms. By randomly perturbing an input to an algorithm, they could interpolate between a worst-time case analysis and an average case analysis, leading to a better understanding of the practical performance of algorithms on real life instances. This has been hugely influential, leading to the study of smoothed analysis in a host of different settings, including numerical analysis [39, 43], satisfiability [22, 14], data clustering [3], multilinear algebra [7] and machine learning [29]. Almost simultaneously, in graph theory, Bohman, Frieze and Martin [8] introduced the model of randomly perturbed graphs which, as with smoothed analysis, allows one to understand the interplay between an extremal and probabilistic viewpoint. The majority of work on the subject has focused on dense graphs [10, 11, 26].

In the context of random walk mixing, it can be seen that small random perturbations cannot speed up the mixing time on dense graphs significantly. Indeed, the canonical examples leading to torpid mixing (e.g., two cliques connected by a long path) are robust with respect to that property. Smoothed analysis of sparse graphs was introduced by Krivelevich, Reichman and Samotij [31]. Here one starts with a connected graph of bounded degree (in fact, bounded degeneracy often suffices) and applies a small random perturbation by adding a copy of the binomial random graph $R \sim G(n, \delta/n)$ for small $\delta > 0$. Although this perturbation is very slight, they showed that it greatly improves the expansion properties of the graph. A graph G is said to be Δ-degenerate if there is some ordering of the vertices of G such that each vertex has at most Δ neighbours in G that precede it in the ordering. To be precise, Krivelevich, Reichman and Samotij

proved that, for any $\Delta \in \mathbb{N}$ and $\delta > 0$, if G is an n-vertex Δ-degenerate connected graph and $R \sim G(n, \delta/n)$, then $G' := G \cup R$ a.a.s. satisfies that $t_{\mathrm{mix}}(G') = O(\log^2 n)$. By considering, for example, a path on n vertices, which has mixing time $\Omega(n^2)$, we see a vast improvement after a slight random perturbation. We also note that the result is tight on such examples, as the randomly perturbed path a.a.s. contains bare paths of length $\Omega(\log n)$.

Our first application of Theorem 1.5 shows that we can improve the mixing time yet further in this model by starting from a uniformly chosen vertex, as in this case we avoid the small bottlenecks that remain if the initial graph had poor expansion.

Theorem 1.6. *For any $\varepsilon, \delta > 0$ and $\Delta \in \mathbb{N}$, there exists a $C > 0$ such that the following holds. Let G be an n-vertex Δ-degenerate connected graph, choose $R \sim G(n, \delta/n)$ and let $G' := G \cup R$. Then, a.a.s.*

$$\bar{t}_{\mathrm{mix}}(G', \varepsilon) \leq C \log n.$$

Remark 1.7. Theorem 1.6 is tight, up to the constant factor C, for *all* graphs with maximum degree Δ. Indeed, this follows from the fact that $N^k(v) = O((2\bar{d})^k \log n)$ for *all* vertices $v \in V(G')$, where $N^k(v)$ denotes the number of vertices that are at distance at most k from v in G' and $\bar{d} := \Delta + \delta$ is an upper bound on the average degree in G'. Such an upper bound can be shown easily by induction, see for example [13], and setting $k = c \log n$ for $c > 0$ sufficiently small shows that at least half of the vertices cannot be reached from v in k steps and hence $\bar{t}_{\mathrm{mix}}(G', \varepsilon) \geq k$. Nonetheless, the converse of the inequality in Theorem 1.6 is not true for all Δ-degenerate graphs. Consider for instance a star: it is 1-degenerate, but the mixing time of the randomly perturbed star is $O(1)$ as we mix in the step after visiting the centre of the star for the first time.

Some time before the systematic study of random perturbations in the combinatorial and theoretical computer science communities discussed above, the notion appeared in physics literature with the study of so-called *small-world networks*. Here we will concentrate on a model introduced by Newman and Watts [37, 36] where, for some fixed $k \in \mathbb{N}$, $\delta > 0$ and $n \in \mathbb{N}$ large, one starts with n-vertices of the graph ordered as $v_1, \ldots, v_n$, adds all edges $v_i v_j$ for which $i + 1 \leq j \leq i + k$ (with addition modulo n), and then adds all remaining edges independently with probability $p = \delta/n$. We denote the resulting random graph as $H_{n,k,\delta}$. It is easy to see that this graph fits into the framework of Krivelevich, Reichman and Samotij [31], and so their result implies that, for any $k \in \mathbb{N}$ and $\delta > 0$, the Newman-Watts small world network $H_{n,k,\delta}$ a.a.s. satisfies $t_{\mathrm{mix}}(H_{n,k,\delta}) = O(\log^2 n)$. In fact, this was established before their work by Addario-Berry and Lei [1], improving on a previous bound of $O(\log^3 n)$ due to Durrett [18]. Here, as a direct consequence of Theorem 1.6, we conclude that the average mixing time on the Newman-Watts small world network is of order $O(\log n)$.

Corollary 1.8. *For all $k \in \mathbb{N}$ and $\varepsilon, \delta > 0$, there exists a $C > 0$ such that the following holds. The Newman-Watts small world network $H_{n,k,\delta}$ a.a.s. satisfies*

$$\bar{t}_{\mathrm{mix}}(H_{n,k,\delta}, \varepsilon) \leq C \log n.$$

1.4 Application 2: Giant components in random subgraphs of expanders

For $p \in [0, 1]$ and a graph G, we define G_p to be the graph with the same vertex set where each edge of G is retained in G_p independently with probability p. The graph G is called the host graph, and the random subgraph G_p, the p-percolated one. Percolation on graphs is a well-established topic in probability theory. Most classically, if the host graph is the complete graph on n vertices K_n, then its p-percolated subgraph is the Erdős-Rényi graph $G(n, p)$. For any graph G, let $L_1(G)$ denote a largest connected component in G and let $\ell_1(G)$ denote its order. In their seminal paper [21], Erdős and Rényi proved a

phase transition for $\ell_1(G(n,p))$. Namely, writing $p = c/n$ for some constant c, if $c < 1$ then a.a.s. $\ell_1(G(n,p)) = O(\log n)$, while if $c > 1$ then a.a.s. $\ell_1(G(n,p)) = \Omega(n)$ and the largest component, which is the unique component of linear size, is known as the *giant component*.

A central question in random graph theory is whether other host graphs exhibit the same phenomenon [2]. One quickly observes that, in order for G_p to have a sharp threshold for the component structure, the host graph G should satisfy some additional properties. A natural property to consider is the pseudo-random notion of *expansion*. There is a strong connection between expansion and the graph spectrum. Given the eigenvalues of the adjacency matrix of a d-regular graph G, say $d = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$, we let $\lambda(G) := \max\{|\lambda_2|, |\lambda_n|\}$ be the second largest eigenvalue. We then define an (n, d, λ)-*graph* to be a d-regular graph G on n vertices with $\lambda(G) = \lambda$. When λ is small compared to d, an (n, d, λ)-graph is said to be an *expander* and it enjoys many of the same properties as a random graph with the same density. We refer the reader to the excellent survey of Krivelevich and Sudakov [32] on the subject.

In terms of percolation, Frieze, Krivelevich and Martin [25] proved that, if G is an (n, d, λ)-graph with $\lambda = o(d)$, then $\ell_1(G_p)$ undergoes a phase transition at $p = 1/d$. They obtained the following description of the supercritical regime: for $\delta > 0$ and $p = (1+\delta)/d$, and provided that $\lambda \leq \delta^4 d$, a.a.s. $\ell_1(G_p) \sim y(\delta)n$ for some $y(\delta) \in (0, 1)$. Moreover, as in $G(n,p)$, the largest component $L_1(G_p)$ is a.a.s. the unique component of linear size. Very recently, Diskin and Krivelevich [17] studied the mixing time of percolated (n, d, λ)-graphs. More precisely, they showed that, in the supercritical regime, there exists $C = C(\delta)$ such that a.a.s. $t_{\mathrm{mix}}(L_1(G_p)) \leq C \log^2 n$. This is indeed optimal for some graphs, in particular for Erdős-Rényi random graphs [24, 5], as discussed above.

Our next application of our main result shows that for percolated pseudo-random graphs the average mixing time is logarithmic.

Theorem 1.9. *For all $\delta > 0$ sufficiently small and all $\varepsilon > 0$, there exists a $C > 0$ such that, if $p = (1+\delta)/d$ and G is an (n, d, λ)-graph with $\lambda \leq \delta^4 d$, then a.a.s.*

$$\bar{t}_{\mathrm{mix}}(L_1(G_p), \varepsilon) \leq C \log n.$$

Similarly as in Remark 1.7, it can be proven that Theorem 1.9 is tight up to multiplicative constant for *all* (n, d, λ)-graphs.

As a consequence, for $G = K_n$ we obtain the following.

Corollary 1.10. *For all $\delta > 0$ sufficiently small and all $\varepsilon > 0$, there exists a $C > 0$ such that, for $p = (1+\delta)/n$, a.a.s.*

$$\bar{t}_{\mathrm{mix}}(L_1(G(n,p)), \varepsilon) \leq C \log n.$$

By Remark 1.1, $t_{\mathrm{mix}}^{(U_n)} = O(\log n)$ a.a.s., where U_n is chosen uniformly at random from $V(G)$. This result aligns with [6], although theirs is much stronger, showing cut-off for $t_{\mathrm{mix}}^{(U_n)}$ as previously mentioned.

1.5 Organisation

The rest of this paper is organised as follows. In Section 2 we introduce all the necessary notation, definitions and tools for our proofs. We use these to prove Theorem 1.5 in Section 3. This section is structured in subsections where we build different tools to be used in our main proof; in particular, we discuss the main ideas of the proof in Section 3.1. Sections 4 and 5 are devoted to proving Theorems 1.6 and 1.9, respectively. Finally, we discuss some open problems in Section 6.

2 Preliminaries

2.1 Basic notation

Let $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ denote the set of non-negative integers. If n is a positive integer, we set $[n] := \{1, \ldots, n\}$. Throughout, we will consider both simple graphs and multigraphs. The word *graph* will refer to simple graphs, that is, each pair of vertices forms at most one edge. Our multigraphs, which will be allowed to have parallel edges but no loops, will be clearly identified as such. All our graphs are labelled, so whenever we discuss an n-vertex (multi)graph G, we implicitly assume that $V(G) = [n]$. Given a (multi)graph $G = (V, E)$ and disjoint sets $A, B \subseteq V(G)$, we write $E_G(A)$ for the (multi)set of edges of G contained in A, and $E_G(A, B)$ for the (multi)set of edges with one endpoint in A and the other in B. We set $e_G(A) := |E_G(A)|$ and $e_G(A, B) := |E_G(A, B)|$. If $A = \{a\}$, we write $E_G(a, B) := E_G(\{a\}, B)$, and similarly in all related notation. For simplicity, we write $e(G) := e_G(V(G))$. We write $G[A] := (A, E_G(A))$. We say that A is *G-connected* if $G[A]$ is connected. For each vertex $v \in V(G)$, we write $\deg_G(v) := e_G(\{v\}, V(G) \setminus \{v\})$ for its degree. We denote $\partial_G(A) := E_G(A, V(G) \setminus A)$ and $\deg_G(A) := \sum_{v \in A} \deg_G(v) = 2e_G(A) + |\partial_G(A)|$.

In many of our statements we will consider an n-vertex graph satisfying a set of conditions or a conclusion, which are often asymptotic in nature. This is in fact an abuse of notation. To be precise, one must consider a sequence $(G_k)_{k \geq 1}$ of graphs on an increasing number of vertices so that the graphs in the sequence satisfy the conditions. This abuse of notation greatly simplifies the statements, so we will assume it throughout. (This also includes any asymptotic statements about random graphs.) For any sequence of graphs $(G_k)_{k \geq 1}$ with $|V(G_k)| \to \infty$, we say that a graph property $\mathcal{P}$ holds asymptotically almost surely (a.a.s.) if $\lim_{k \to \infty} \mathbb{P}[G_k \in \mathcal{P}] = 1$.

2.2 Random walks

Given an arbitrary connected multigraph G, the lazy random walk over G is a Markov chain on state space $V(G)$ defined by the transition matrix $P_G = (P_G(i, j))_{i,j \in V(G)}$ given by

$$P_G(i, j) = \begin{cases} 1/2 & \text{if } i = j, \\ e_G(i, j)/(2 \deg_G(i)) & \text{if } i \neq j. \end{cases} \tag{2.1}$$

That is, the lazy random walk is a sequence of random variables $(X_t)_{t \geq 0}$ with probability distributions $(\mu_t)_{t \geq 0}$, respectively, over $V(G)$, where μ_0 is the starting distribution and, for each $t \geq 1$, the distribution of μ_t is obtained from the distribution of μ_{t-1} as $\mu_t = \mu_{t-1} P_G = \mu_0 P_G^t$. The sequence of distributions thus depends only on G and the starting distribution. In the special case when there is a vertex $x \in V(G)$ such that $\mu_0(x) = 1$, we will write $(\mu_t^x)_{t \geq 0}$ to denote the resulting sequence of distributions.

If G is connected, the lazy random walk over G converges to a stationary distribution π_G (that is, a distribution satisfying $\pi_G = \pi_G P_G$), independently of the starting distribution μ_0. It is well known (see, e.g., [33]) that this stationary distribution satisfies

$$\pi_G(u) = \frac{\deg_G(u)}{2e(G)} \tag{2.2}$$

for all $u \in V(G)$. Given a set $S \subseteq V(G)$, we define $\pi_G(S) := \sum_{v \in S} \pi_G(v)$. It follows from (2.2) that

$$\pi_G(S) = \frac{\deg_G(S)}{2e(G)} = \frac{2e_G(S) + |\partial_G(S)|}{2e(G)}. \tag{2.3}$$

We define $\pi_{\min}(G) := \min_{v \in V(G)} \pi_G(v)$ and $\pi_{\max}(G) := \max_{v \in V(G)} \pi_G(v)$.

Recall the definition of mixing times in the introduction. The mixing time of a random walk on a connected (multi)graph G is deeply tied with the concept of *conductance*. Given a set $S \subseteq V(G)$, we define

$$Q_G(S) := \sum_{i \in S} \sum_{j \in V(G) \setminus S} \pi_G(i) P_G(i,j) = \frac{|\partial_G(S)|}{4e(G)}, \tag{2.4}$$

where the equality follows from (2.1) and (2.2). Observe that $Q_G(S) = Q_G(V(G) \setminus S)$. Finally, we define the *conductance* $\Phi_G(S)$ of S as

$$\Phi_G(S) := \frac{Q_G(S)}{\pi_G(S) \pi_G(V(G) \setminus S)}. \tag{2.5}$$

From the definitions in (2.3), (2.4) and (2.5) and the fact that $\deg_G(A) \leq 2e(G)$ for any set $A \subseteq V(G)$, it follows that

$$\Phi_G(S) = \frac{e(G)|\partial_G(S)|}{\deg_G(S) \deg_G(V(G) \setminus S)} \geq \frac{|\partial_G(S)|}{2 \deg_G(S)}. \tag{2.6}$$

Our approach to estimate the mixing time of the lazy random walk over a multigraph G is based on ideas of Fountoulakis and Reed [23, 24]. Roughly speaking, their main contribution is the fact that the mixing time of an abstract irreducible, reversible, aperiodic Markov chain (which we may represent using a weighted graph H on its state space) can be bounded from above using the conductances of different H-connected sets of states of various sizes. The fact that we may restrict ourselves to H-connected sets is crucial to obtain tighter bounds than would be obtained through other classical means. For simplicity, here we only state a version of the result of Fountoulakis and Reed [23] which is applicable to our setting. For any $p \in (\pi_{\min}(G), 1)$, we let $\Phi_G(p)$ be the minimum conductance $\Phi_G(S)$ over all G-connected sets $S \subseteq V(G)$ such that $p/2 \leq \pi_G(S) \leq p$ (if no such set S exists, we set $\Phi_G(p) = 1$).

Theorem 2.1 (Fountoulakis and Reed [23]). *Let G be a connected multigraph. There exists an absolute constant C_0 such that*

$$t_{\mathrm{mix}}(G) \leq C_0 \sum_{j=1}^{\lceil \log_2 \pi_{\min}(G)^{-1} \rceil} \Phi_G^{-2}(2^{-j}).$$

Another parameter of interest is the *hitting time* to a vertex (or set of vertices) in the random walk on a multigraph G. Given any $v \in V(G)$ and the lazy random walk $(X_t)_{t \geq 0}$ with starting distribution μ_0, we define the *hitting time* to v as

$$\tau_G(\mu_0, v) := \inf\{t \in \mathbb{N}_0 : X_t = v, X_0 \sim \mu_0\}.$$

In more generality, given any set $S \subseteq V(G)$, we define the *hitting time* to S as

$$\tau_G(\mu_0, S) := \inf\{t \in \mathbb{N}_0 : X_t \in S, X_0 \sim \mu_0\}.$$

Given any vertex $u \in V(G)$, let P_u be the matrix obtained from the transition matrix P_G by removing the row and column corresponding to u. If P_u is primitive (i.e., all entries of P_u^m are positive for some $m \geq 1$), by Perron-Frobenius, the largest eigenvalue of P_u, denoted by λ_u, is real, of multiplicity 1 and satisfies $\lambda_u < 1$.

We will make use of the first visit time lemma of Cooper and Frieze [15]. Here we state a more recent version with weaker hypotheses due to Manzo, Quattropani and Scoppola [35].

Theorem 2.2 (First Visit Time Lemma, Manzo, Quattropani and Scoppola [35])**.** *Let G be an n-vertex connected multigraph. Suppose that there exist a real number $c > 2$ and a diverging sequence $T = T(n)$ such that the following conditions hold:*

(HP1) *Fast mixing:* $\max_{x,y \in V(G)} |\mu_T^x(y) - \pi_G(y)| = o(n^{-c})$.

(HP2) *Small $\pi_{\max}$:* $T \cdot \pi_{\max}(G) = o(1)$.

(HP3) *Large $\pi_{\min}$:* $\pi_{\min}(G) = \omega(n^{-2})$.

Then, for all $u \in V(G)$, we have

$$\sup_{t \geq 0} \left| \frac{\mathbb{P}[\tau_G(\pi_G, u) > t]}{\lambda_u^t} - 1 \right| \xrightarrow[n \to \infty]{} 0 \tag{2.7}$$

and

$$\left| \frac{1 - \lambda_u}{\pi_G(u)/R_T(u)} - 1 \right| \xrightarrow[n \to \infty]{} 0, \tag{2.8}$$

where

$$R_T(u) := \sum_{t=0}^{T} \mu_t^u(u) > 1$$

is the expected number of indices $t \in [T] \cup \{0\}$ for which the lazy random walk $(X_t)_{t \geq 0}$ on G starting at $X_0 = u$ satisfies $X_t = u$.

From the intuitive point of view, the theorem says that the hitting time to u is roughly distributed as a geometric random variable with success probability $\pi_G(u)/R_T(u)$. If one wants to hit u by independently sampling vertices according to π_G, then it would be a geometric random variable with success probability $\pi_G(u)$. The factor $R_T(u)$ is the price to pay for taking into account the geometry of the graph: the more likely it is to return from u to u, the less connected u is to the rest of the graph, and the smaller the probability to hit it at a given (large) time is.

Remark 2.3. In the proof of Theorem 2.2, one can check that, if we only want (2.7) to hold for a given $u \in V(G)$, then (HP2) can be replaced by

(HP2′) *Small $\pi_G(u)$:* $T \cdot \pi_G(u) = o(1)$.

Remark 2.4. The following holds as a corollary of Theorem 2.2. For any fixed $D > 0$ and n-vertex connected multigraph G satisfying (HP1), (HP2) (or (HP2′)) and (HP3) with the additional property that $e(G) \leq Dn$, if $u \in V(G)$ and $t_0 = t_0(n)$ is such that $\lambda_u^{t_0} = 1 - o(1)$, then

$$\frac{1}{n} \sum_{v \in V(G)} \mathbb{P}[\tau_G(\mu_0^v, u) \leq t_0] = o(1). \tag{2.9}$$

Indeed, for $\varepsilon > 0$, Theorem 2.2 implies that $\mathbb{P}[\tau_G(\pi_G, u) \leq t_0] \leq \varepsilon^2$. Then, $B_\varepsilon := \{v \in V(G) : \mathbb{P}[\tau_G(\mu_0^v, u) \leq t_0] \geq \varepsilon\}$ satisfies $\pi_G(B) \leq \varepsilon$. Moreover, as G is connected, we have that $\pi_G(B) \geq |B|/(2e(G)) \geq |B|/(2Dn)$ and so $|B| \leq 2\varepsilon Dn$. Therefore,

$$\sum_{v \in V(G)} \mathbb{P}[\tau_G(\mu_0^v, u) \leq t_0] = \sum_{v \in V(G) \setminus B} \mathbb{P}[\tau_G(\mu_0^v, u) \leq t_0] + \sum_{v \in B} \mathbb{P}[\tau_G(\mu_0^v, u) \leq t_0]$$

$$\leq (2D + 1)\varepsilon n,$$

which can be made arbitrarily small, by taking ε small with respect to D. This establishes (2.9).

3 A general approach to average mixing times

3.1 Proof overview

As discussed in the introduction, the main tool we will use to bound the mixing times of random walks is the result of Fountoulakis and Reed (Theorem 2.1) which relates the (worst-case) mixing time of a random walk in a graph G to the conductance (see (2.5)) of the G-connected vertex subsets S of G. We think of vertex subsets S whose conductance is *poor* (those for which $\Phi_G(S) = o(1)$) as bottlenecks: they have more edges internally in S than leaving S and so the random walk is likely to get held up in S. The spreader graphs (see Definition 1.3) we are interested in studying here have only few small bottlenecks. Indeed, any vertex subset which can lead to small bottlenecks must either be thin or loaded and our upper bounds on the number of these sets in a spreader graph readily imply that any vertex set with poor conductance is of at most polylogarithmic size (size $(\log^2 n)$ to be precise, see Remark 3.3). Now, if a set with poor conductance is very small (size at most $(\log n)^{1/5}$), it will not slow down mixing significantly as our random walk will not get stuck for very long in these sets before leaving them. Therefore, it is the intermediate size sets which pose a problem, and we will first show in Section 3.2 (see Lemma 3.4) that the set U of bad vertices contained in *some* intermediate set which has poor conductance contains a negligible proportion of the overall vertex set of our spreader graph.

Intuitively, we can then see how starting at an average vertex in G speeds up the mixing time. Indeed, we are very unlikely to start at a bad vertex in U and, moreover, we are in fact very unlikely to visit a vertex in U in the first $O(\log n)$ time steps, by which time we aim to show that the distribution of the random walk is already well-mixed. In order to formalise this intuition, we adjust our spreader graph G by shrinking the intermediate sets with poor conductance and thus removing troublesome small bottlenecks. The resulting (multi-)graph we will call G^*. Using that the number of bad vertices $|U|$ is negligible, or rather that the number of edges incident to U, $\deg_G(U)$, is negligible (Lemma 3.4), we show in Section 3.3 that switching from G to G^* does not have a big effect on the edge distribution and that, in particular, the stationary distributions on G and G^* are comparable. We then show in Section 3.4 that, after contracting intermediate sets with poor conductance, we can apply Theorem 2.1 of Fountoulakis and Reed to conclude that the *worst-case* mixing time in G^* is logarithmic. Here we will need that G^* is defined carefully to preserve connectivity between sets after contractions (see (3.7)). Finally, we will prove Theorem 1.5 by coupling the random walk from an average vertex on G with the random walk in G^*. As the random walk in G^* from *any* starting point mixes rapidly, we can conclude that the random walk in G also mixes rapidly as long as the two random walks stay coupled for long enough. For this, our final ingredient is to show that the random walk in G is unlikely to hit our bad vertices U, which we do in Section 3.5 by appealing to the First Visit Time Lemma (Theorem 2.2) of Manzo, Quattropani and Scoppola.

3.2 Badly connected sets

We will make use of the following simple definition.

Definition 3.1. *For $\gamma > 0$ and a connected multigraph G, we say a set $S \subseteq V(G)$ is γ-bad in G if*

$$\frac{|\partial_G(S)|}{\deg_G(S)} < \gamma,$$

and that it is γ-good otherwise.

The following lemma gives us a basic property of bad sets in a connected multigraph

and quickly ties them with our notion of (α, D)-spreader graphs. Note that the notions of thin and loaded sets extend naturally to multigraphs.

Lemma 3.2. *Let G be a multigraph. For $0 < \alpha \le 1$, if a set $S \subseteq V(G)$ is $(\alpha^2/4)$-bad in G, then*

(1) $|\partial_G(S)| \le 2e_G(S)$, *and*

(2) *either S is α-thin in G or it is α^{-1}-loaded in G (or both).*

Proof. The assertion (1) follows easily since, if $|\partial_G(S)| > 2e_G(S)$, then

$$\frac{|\partial_G(S)|}{\deg_G(S)} = \frac{|\partial_G(S)|}{2e_G(S) + |\partial_G(S)|} \ge \frac{|\partial_G(S)|}{2|\partial_G(S)|} = \frac{1}{2} \ge \frac{\alpha^2}{4},$$

a contradiction.

For the second assertion, suppose that S is neither α-thin nor α^{-1}-loaded in G. Then, we have that

$$\frac{|\partial_G(S)|}{\deg_G(S)} = \frac{|\partial_G(S)|}{2e_G(S) + |\partial_G(S)|} \ge \frac{|\partial_G(S)|}{4e_G(S)} \ge \frac{\alpha|S|}{4\alpha^{-1}|S|} = \frac{\alpha^2}{4},$$

a contradiction. Here we used that $|\partial_G(S)| \le 2e_G(S)$ from (1) in the first inequality and the definitions of α-thin and α^{-1}-loaded in the second. $\qquad\square$

We will also make use of the following simple observation.

Remark 3.3. Let $\alpha, D > 0$, and let G be an n-vertex graph satisfying (S1) and (S2). Then, there are *no* G-connected α-thin or α^{-1}-loaded sets S of size $(\log n)^2 \le |S| \le (1 - 1/D^2)n$.

Given any $\alpha > 0$ and $D > 0$, we define $\beta := 1/D^2$ and $\gamma := \alpha^2/4$. Given any n-vertex graph G, let

$$\mathcal{S}(\alpha, D) := \left\{ S \subseteq V(G) : S \text{ is } G\text{-connected and } \gamma\text{-bad}, (\log n)^{1/5} \le |S| \le (1 - \beta)n \right\}. \quad (3.1)$$

Further, we define $U(\alpha, D) \subseteq V(G)$ to be the set of vertices that lie in sets in $\mathcal{S}(\alpha, D)$, that is,

$$U(\alpha, D) := \bigcup_{S \in \mathcal{S}(\alpha, D)} S. \quad (3.2)$$

We now give bounds on $|U(\alpha, D)|$ and $\deg_G(U(\alpha, D))$.

Lemma 3.4. *Let $\alpha \in (0, 1]$ and $D > 0$. Let G be an n-vertex connected graph which satisfies (S1) and (S2). If n is sufficiently large, then*

$$|U(\alpha, D)| \le n \exp(-(\log n)^{1/11}) \quad (3.3)$$

and

$$\deg_G(U(\alpha, D)) \le n \exp(-(\log n)^{1/11}). \quad (3.4)$$

In particular,

$$\pi_G(U(\alpha, D)) \le \exp(-(\log n)^{1/11}). \quad (3.5)$$

Proof. Let $\mathcal{S} = \mathcal{S}(\alpha, D)$ and $U = U(\alpha, D)$. By Lemma 3.2(2), all sets $S \in \mathcal{S}$ must be α-thin or α^{-1}-loaded. By (S1) and (S2), for each $(\log n)^{1/5} \le k \le (1 - \beta)n$, the number of α-thin or α^{-1}-loaded sets of size k in G is less than $2n \exp(-\sqrt{k})$. Moreover, as stated in Remark 3.3, there are no α-thin or α^{-1}-loaded sets of size $k \ge (\log n)^2$. Thus,

$$|\mathcal{S}| < 2n \sum_{k=(\log n)^{1/5}}^{(\log n)^2} e^{-\sqrt{k}} \le 2n(\log n)^2 e^{-(\log n)^{1/10}}. \quad (3.6)$$

Therefore, by the definition of U, the bound on the size of the largest $S \in \mathcal{S}$, and assuming n is sufficiently large, we conclude that

$$|U| \leq \sum_{S \in \mathcal{S}} |S| \leq |\mathcal{S}|(\log n)^2 < 2n(\log n)^4 \mathrm{e}^{-(\log n)^{1/10}} \leq n\mathrm{e}^{-(\log n)^{1/11}}.$$

We now turn our attention to (3.4). By the definition of U, we have that

$$\deg_G(U) \leq \sum_{S \in \mathcal{S}} \deg_G(S),$$

and using Lemma 3.2(1) this simplifies to

$$\deg_G(U) \leq \sum_{S \in \mathcal{S}} 4e_G(S).$$

Now, for each $S \in \mathcal{S}$, let $b(S)$ be some G-connected set such that $(\log n)^2 \leq |b(S)| \leq (\log n)^3$ and $S \subseteq b(S)$. Note that this is possible because G is connected and every $S \in \mathcal{S}$ has size less than $(\log n)^2$. By (S2) and the bounds on $|b(S)|$, for each $S \in \mathcal{S}$ we have that

$$e_G(b(S)) \leq \alpha^{-1}|b(S)| \leq \alpha^{-1}(\log n)^3.$$

Therefore, using (3.6) and for n sufficiently large,

$$\deg_G(U) \leq \sum_{S \in \mathcal{S}} 4e_G(S) \leq \sum_{S \in \mathcal{S}} 4e_G(b(S)) \leq 4\alpha^{-1}(\log n)^3|\mathcal{S}| \leq n\mathrm{e}^{-(\log n)^{1/11}}.$$

In particular, since G is connected (so $e(G) \geq n/2$), it follows from (2.3) that

$$\pi_G(U) = \frac{\deg_G(U)}{2e(G)} \leq \mathrm{e}^{-(\log n)^{1/11}}. \qquad \square$$

3.3 Stationary distributions

Let $\alpha > 0$ and $D > 0$ be given, let G be some n-vertex graph, and consider the set $U = U(\alpha, D)$. We now define $G^* = G^*(\alpha, D)$ to be the multigraph obtained by contracting all the connected components of $G[U]$ to single vertices. To be more precise, let $U = U_1 \cup \ldots \cup U_t$ be a partition of U into sets each of which induces a connected component in $G[U]$ and let $U^* \coloneqq \{u_1, \ldots, u_t\}$ be a set of t new vertices. Then, let $V(G^*) = (V(G) \setminus U) \cup U^*$ and, for each $x \in V(G)$, let

$$f(x) = \begin{cases} x & \text{if } x \notin U, \\ u_i & \text{if } x \in U_i \subseteq U. \end{cases} \tag{3.7}$$

In particular, $f(U) = U^*$. Finally, we define the multiset

$$E(G^*) \coloneqq \{f(x)f(y) : xy \in E(G), f(x) \neq f(y)\}.$$

Observe that G^* is connected if and only if G is connected, and that U^* must be an independent set in G^*.

Given some connected graph G, we want to compare the behaviour of the lazy random walk on G and its contracted form G^*. In particular, we wish to compare their stationary distributions. In order to do this, we need to make them comparable by having them on the same state space. Let us describe this in full generality. Let $H_1 = (V_1, E_1)$ and $H_2 = (V_2, E_2)$ be two connected multigraphs (possibly with $V_1 \cap V_2 \neq \varnothing$). Then, we define an auxiliary multigraph $H \coloneqq H_1 \cup H_2 = (V_1 \cup V_2, E_1 \cup E_2)$ as the union of H_1 and H_2.

Given the stationary distributions π_{H_1} and π_{H_2}, we define two distributions σ_1 and σ_2 on H, where, for each $v \in V(H)$,

$$
\begin{aligned}
\sigma_1(v) &= \begin{cases} \pi_{H_1}(v) & \text{if } v \in V_1, \\ 0 & \text{otherwise}; \end{cases} \\
\sigma_2(v) &= \begin{cases} \pi_{H_2}(v) & \text{if } v \in V_2, \\ 0 & \text{otherwise}. \end{cases}
\end{aligned}
\tag{3.8}
$$

For the sake of notation, for any vertex $v \in V(H) \setminus V_1$ we set $\deg_{H_1}(v) = 0$ and, similarly, for any $v \in V(H) \setminus V_2$ we set $\deg_{H_2}(v) = 0$. With this setup, we abuse notation slightly and write $d_{\mathrm{TV}}(\pi_{H_1}, \pi_{H_2})$ for $d_{\mathrm{TV}}(\sigma_1, \sigma_2)$.

Lemma 3.5. *Let $D \geq 1$ and $0 < \alpha < 1/D^2$. Let G be an n-vertex connected graph which satisfies* (S1) *and* (S2)*. If n is sufficiently large, we have that*

$$
d_{\mathrm{TV}}(\pi_G, \pi_{G^*}) \leq \exp(-(\log n)^{1/12}).
$$

Proof. Let $U = U(\alpha, D)$, and let $U^* = f(U) \subseteq V(G^*)$. We also fix $\tilde{G} := G \cup G^*$, and recall the distributions defined in (3.8). Observe that

$$
\sum_{v \in V(\tilde{G})} |\deg_G(v) - \deg_{G^*}(v)| = \sum_{v \in U} \deg_G(v) + \sum_{v \in U^*} \deg_{G^*}(v) \leq 2 \deg_G(U).
\tag{3.9}
$$

The equality uses the fact that, for all $v \in V(G) \setminus U$, we have $\deg_G(v) = \deg_{G^*}(v)$. In the inequality, we simply note that $\sum_{v \in U} \deg_G(v) = \deg_G(U)$ by definition and that, by the definition of G^*, this gives an upper bound for the second sum. Moreover, from (1.1) and the triangle inequality we have

$$
\begin{aligned}
d_{\mathrm{TV}}(\pi_G, \pi_{G^*}) &= \frac{1}{2} \sum_{v \in V(\tilde{G})} \left| \frac{\deg_G(v)}{2e(G)} - \frac{\deg_{G^*}(v)}{2e(G^*)} \right| \\
&\leq \frac{1}{2} \sum_{v \in V(\tilde{G})} \left(\left| \frac{\deg_G(v) - \deg_{G^*}(v)}{2e(G)} \right| + \left| \frac{\deg_{G^*}(v)}{2e(G)} - \frac{\deg_{G^*}(v)}{2e(G^*)} \right| \right).
\end{aligned}
\tag{3.10}
$$

The second term in the sum can be evaluated as

$$
\begin{aligned}
\sum_{v \in V(\tilde{G})} \left| \frac{\deg_{G^*}(v)}{2e(G)} - \frac{\deg_{G^*}(v)}{2e(G^*)} \right| &= \sum_{v \in V(\tilde{G})} \frac{e(G) - e(G^*)}{2e(G)e(G^*)} \deg_{G^*}(v) \\
&= \frac{e(G) - e(G^*)}{e(G)} \\
&= \frac{1}{2e(G)} \sum_{v \in V(\tilde{G})} (\deg_G(v) - \deg_{G^*}(v)) \\
&\leq \frac{1}{2e(G)} \sum_{v \in V(\tilde{G})} |\deg_G(v) - \deg_{G^*}(v)|.
\end{aligned}
$$

Introducing this in (3.10) together with (3.9) and using Lemma 3.4 and that $e(G) \geq n/2$, we conclude that

$$
d_{\mathrm{TV}}(\pi_G, \pi_{G^*}) \leq \sum_{v \in V(\tilde{G})} \left| \frac{\deg_G(v) - \deg_{G^*}(v)}{2e(G)} \right| \leq \frac{\deg_G(U)}{e(G)} \leq \exp(-(\log n)^{1/12}). \qquad \square
$$

3.4 Mixing time after contractions

The following result shows the mixing properties of contracted spreader graphs.

Proposition 3.6. *For all $D \geq 4$, $0 < \alpha < 1/D^2$ and $\varepsilon > 0$, there exists a $C > 0$ such that the following holds for all n sufficiently large. Suppose G is an n-vertex connected (α, D)-spreader graph. Then,*

$$t_{\mathrm{mix}}(G^*(\alpha, D), \varepsilon) \leq C \log n.$$

In order to prove Proposition 3.6, we will rely on the following lemma.

Lemma 3.7. *For all $D \geq 4$ and $0 < \alpha < \beta = 1/D^2$, the following holds for all n sufficiently large. Suppose G is an n-vertex connected (α, D)-spreader graph. Then, taking $G^* = G^*(\alpha, D)$, for all G^*-connected $S^* \subseteq V(G^*)$ such that $(\log n)^{1/2}/n \leq \pi_{G^*}(S^*) \leq 1/2$ we have that $\Phi_{G^*}(S^*) \geq \alpha^2/8$.*

Proof. Let $U := U(\alpha, D)$ and $\gamma := \alpha^2/4$. Recall our definition of $f \colon V(G) \to V(G^*)$ from (3.7). The proof will make use of the following claim.

Claim 3.8. *Any G^*-connected set $S^* \subseteq V(G^*)$ with $|S^*| \geq 2$ and such that $(\log n)^{1/5} \leq |f^{-1}(S^*)| \leq (1-\beta)n$ is γ-good in G^*.*

Proof. Take any such $S^* \subseteq V(G^*)$ and let $S := f^{-1}(S^*) = \{v \in V(G) : f(v) \in S^*\}$. Note that S must be G-connected. We claim that the bounds on $|S|$ imply that it must be γ-good in G. Indeed, if S was γ-bad in G, it would be a G-connected subset of U (recall (3.1) and (3.2)) and so would be mapped by f to a single vertex. As f is surjective and $|S^*| \geq 2$, this is clearly not possible. It also follows from the definition of G^* that $e_G(S) \geq e_{G^*}(S^*)$ and $|\partial_G(S)| = |\partial_{G^*}(S^*)|$. Therefore,

$$\frac{|\partial_{G^*}(S^*)|}{\deg_{G^*}(S^*)} = \frac{|\partial_{G^*}(S^*)|}{2e_{G^*}(S^*) + |\partial_{G^*}(S^*)|} = \frac{|\partial_G(S)|}{2e_{G^*}(S^*) + |\partial_G(S)|} \geq \frac{|\partial_G(S)|}{2e_G(S) + |\partial_G(S)|} \geq \gamma,$$

and so S^* is γ-good in G^*. ◄

We will also need the fact that none of the sets from the statement are too large.

Claim 3.9. *Let $S^* \subseteq V(G^*)$ be a G^*-connected set such that $\pi_{G^*}(S^*) \leq 1/2$. Then, $|S^*| \leq (1 - 2\beta)n$.*

Proof. Assume for a contradiction that there is such a set with $|S^*| > (1 - 2\beta)n$. Letting $\bar{S}^* := V(G^*) \setminus S^*$, we have that $\pi_{G^*}(S^*) + \pi_{G^*}(\bar{S}^*) = 1$ and $|\partial_{G^*}(S^*)| = |\partial_{G^*}(\bar{S}^*)|$, implying that

$$e_{G^*}(\bar{S}^*) \geq e_{G^*}(S^*) \geq |S^*| - 1 \geq (1 - 3\beta)n > 3n/4,$$

where we used here that S^* is G^*-connected and the fact that $\beta = 1/D^2 \leq 1/16$. Let $T \subseteq V(G)$ be some set of size $3\beta n$ such that $f^{-1}(\bar{S}^*) \subseteq T$, noting that this is possible since, by Lemma 3.4, we have $|f^{-1}(\bar{S}^*)| \leq |\bar{S}^*| + |U| \leq 3\beta n$. Now, by (S3), we have that $e_{G^*}(\bar{S}^*) \leq e_G(T) \leq D|T| \leq 3D\beta n \leq 3n/4$, a contradiction, where we again appealed to the facts that $\beta = 1/D^2$ and $D \geq 4$. ◄

Now suppose that there exists some G^*-connected set $S^* \subseteq V(G^*)$ with $(\log n)^{1/2}/n \leq \pi_{G^*}(S^*) \leq 1/2$ and $\Phi_{G^*}(S^*) < \alpha^2/8$. By the bound on the conductance, it follows from (2.6) that S^* is γ-bad in G^* and so, by Lemma 3.2(1), we have $|\partial_{G^*}(S^*)| \leq 2e_{G^*}(S^*)$. This implies that

$$4e_{G^*}(S^*) \geq \deg_{G^*}(S^*) = \pi_{G^*}(S^*) \cdot 2e(G^*)$$

$$\geq \frac{(\log n)^{1/2}}{n} 2e(G^*)$$

$$\geq \frac{(\log n)^{1/2}(2e(G) - \deg_G(U))}{n}$$

$$\geq \frac{(\log n)^{1/2}}{2}, \tag{3.11}$$

where in the last inequality we used the fact that $\deg_G(U) = o(n)$ from Lemma 3.4 and the fact that $e(G) \geq n/2$ as G is connected.

Observe that, since $\gamma < 1$ and G^* is connected, no set $S^* \subseteq V(G^*)$ with $|S^*| = 1$ can be γ-bad, so we must have $|S^*| \geq 2$. Then, Claim 3.8 implies that $S := f^{-1}(S^*)$ has size $|S| < (\log n)^{1/5}$ or $|S| > (1-\beta)n$. If $|S| > (1-\beta)n$, then $|S^*| \geq |S| - |U| > (1-2\beta)n$ (again by Lemma 3.4), and we know this cannot happen by Claim 3.9, so we must have $|S| < (\log n)^{1/5}$. As, trivially, any vertex set $S' \subseteq V(G)$ has $e_G(S') \leq |S'|^2$, we have that $e_{G^*}(S^*) \leq e_G(S) \leq (\log n)^{2/5}$. This contradicts (3.11). $\qquad\square$

With this, we can prove Proposition 3.6.

Proof of Proposition 3.6. Let $G^* := G^*(\alpha, D)$. Observe that (S3) implies $e(G^*) \leq e(G) \leq Dn$. By Lemma 3.7, for any G^*-connected set $S^* \subseteq V(G^*)$ with $(\log n)^{1/2}/n \leq \pi_{G^*}(S^*) \leq 1/2$ we have $\Phi_{G^*}(S^*) \geq \alpha^2/8$. Consider now any G^*-connected set $S^* \subseteq V(G^*)$ with $\pi_{G^*}(S^*) \leq (\log n)^{1/2}/n$ (in particular, $S^* \neq V(G^*)$). The fact that G^* is connected together with (2.4) and (2.5) ensures that

$$\Phi_{G^*}(S^*) \geq \frac{|\partial_{G^*}(S^*)|}{4e(G^*) \cdot \pi_{G^*}(S^*)} \geq \frac{1}{4Dn\pi_{G^*}(S^*)}.$$

Let J denote the set of indices $j \geq 1$ such that $2^{-j} \leq (\log n)^{1/2}/n$, and note that $\min(J) \leq \log n$. It then follows that

$$\sum_{j=1}^{\lceil \log_2 \pi_{\min}(G^*)^{-1} \rceil} \Phi_{G^*}^{-2}(2^{-j}) \leq \log n \cdot \frac{64}{\alpha^4} + \sum_{j \in J} 2^{-2j}(4Dn)^2$$

$$\leq \log n \cdot \frac{64}{\alpha^4} + 2\max_{j \in J}\{2^{-2j}\} \cdot 16D^2 n^2 \leq \left(\frac{64}{\alpha^4} + 32D^2\right)\log n,$$

where in the last inequality we use the definition of J. By Theorem 2.1 we have

$$t_{\mathrm{mix}}(G^*) \leq C_0\left(\frac{64}{\alpha^4} + 32D^2\right)\log n,$$

where C_0 is some absolute constant. Since the total variation distance decreases exponentially fast after the mixing time (see, e.g., [33, section 4.5]), we get

$$t_{\mathrm{mix}}(G^*, \varepsilon) \leq C_0\left(\frac{64}{\alpha^4} + 32D^2\right)\lceil \log_2(1/\varepsilon) \rceil \log n,$$

and the proposition holds by taking C appropriately. $\qquad\square$

3.5 Hitting time of bad vertices

Let $D \geq 4$ and $0 < \alpha < 1/D^2$ and consider a connected (α, D)-spreader graph G. We now wish to study the hitting time to the set of bad vertices $U(\alpha, D)$ in G and show that a.a.s. it is not too small.

Lemma 3.10. *Let $D \geq 4$ and $0 < \alpha < 1/D^2$. Let G be an n-vertex connected (α, D)-spreader graph, and let $U = U(\alpha, D)$. Then,*

$$\frac{1}{n}\sum_{v \in V(G) \setminus U} \mathbb{P}[\tau_G(\mu_0^v, U) \leq (\log n)^2] = o(1).$$

Proof. In this proof we will use a new auxiliary multigraph $\hat{G}$. Let us introduce it here. Consider the multigraph $G^* = G^*(\alpha, D)$, and let $U^* \coloneqq f(U) \subseteq V(G^*)$. Recall that the definition of U^* implies that it is an independent set in G^*. Now, $\hat{G}$ is obtained from G^* by contracting U^* to a single new vertex u^*. The fact that U^* is an independent set in G^* guarantees that $e(G^*) = e(\hat{G})$, and since vertices outside U^* do not see their degree changed by this operation, it follows that $\pi_{\hat{G}}(v) = \pi_{G^*}(v)$ for all $v \in V(G) \setminus U$ and $\pi_{\hat{G}}(u^*) = \pi_{G^*}(U^*)$.

Let $(X_t)_{t \geq 0}$ and $(\hat{X}_t)_{t \geq 0}$ denote lazy random walks on G and $\hat{G}$ starting on some vertex $v \in V(G) \setminus U$. Let us denote $\tau_U^v \coloneqq \tau_G(\mu_0^v, U)$ and $\hat{\tau}_{u^*}^v \coloneqq \tau_{\hat{G}}(\mu_0^v, u^*)$. Define the natural coupling $(X_t, \hat{X}_t)_{t \geq 0}$ as follows: for any $t \geq 1$, while $X_t \notin U$, let $\hat{X}_t = X_t$; if there is a $t \geq 1$ such that $X_t \in U$, then for the smallest such t we let $\hat{X}_t = u^*$; otherwise (that is, for all $t > \tau_U^v$), we let X_t and $\hat{X}_t$ evolve independently. Observe that this is indeed a valid coupling since for all $v \in V(G) \setminus U$ we have $e_G(v, U) = e_{\hat{G}}(v, u^*)$. With this natural coupling, conditional on $\tau_U^v > t$, we have $X_t = \hat{X}_t$ and so $\tau_U^v = \hat{\tau}_{u^*}^v$; in particular, for any $T_0 \geq 0$ we have

$$\mathbb{P}[\tau_U^v > T_0] = \mathbb{P}[\hat{\tau}_{u^*}^v > T_0]. \tag{3.12}$$

We will study the hitting times $\hat{\tau}_{u^*}^v$ using Theorem 2.2 on $\hat{G}$ with $T = T(n) = (\log n)^6$ and $c = 3$. We start with the following claim (which we make no efforts to optimise).

Claim 3.11. *We have that* $t_{\mathrm{mix}}(\hat{G}) \leq (\log n)^4$.

Proof. First we prove that, for any $\hat{G}$-connected set $\hat{S} \subseteq V(\hat{G})$ such that $\pi_{\hat{G}}(\hat{S}) \leq 1/2$, we have that $\Phi_{\hat{G}}(\hat{S}) \geq 1/(2 \log n)$. Indeed, fix such an $\hat{S}$ and define $S^* \subseteq V(G^*)$ as

$$S^* = \begin{cases} \hat{S} & \text{if } u^* \notin \hat{S}, \\ (\hat{S} \setminus \{u^*\}) \cup U^* & \text{if } u^* \in \hat{S}. \end{cases}$$

Further, let $S^* = S_1 \cup \ldots \cup S_r$ for some $r \in \mathbb{N}$ be a decomposition of S^* into G^*-connected components (note that $r = 1$ if $u^* \notin \hat{S}$). Now, as $\deg_{\hat{G}}(\hat{S}) = \deg_{G^*}(S^*)$ and $e(\hat{G}) = e(G^*)$, we have that $\pi_{G^*}(S^*) \leq 1/2$ and hence $\pi(S_i) \leq 1/2$ for all $i \in [r]$. Moreover, as U^* is an independent set in G^*, we have that $|\partial_{\hat{G}}(\hat{S})| = \sum_{i=1}^r |\partial_{G^*}(S_i)|$ and $\deg_{\hat{G}}(\hat{S}) = \sum_{i=1}^r \deg_{G^*}(S_i)$.

Returning to analyse the conductance of $\hat{S}$, from (2.6), we have that

$$\Phi_{\hat{G}}(\hat{S}) \geq \frac{|\partial_{\hat{G}}(\hat{S})|}{2 \deg_{\hat{G}}(\hat{S})} = \frac{\sum_{i=1}^r |\partial_{G^*}(S_i)|}{2 \sum_{i=1}^r \deg_{G^*}(S_i)} \geq \frac{1}{2} \min_{i \in [r]} \frac{|\partial_{G^*}(S_i)|}{\deg_{G^*}(S_i)} = \frac{|\partial_{G^*}(S_{i_0})|}{2 \deg_{G^*}(S_{i_0})}, \tag{3.13}$$

letting $i_0 \in [r]$ be a minimising index. If $\deg_{G^*}(S_{i_0}) \leq \log n$, then we are done due to the fact that $|\partial_{G^*}(S_{i_0})| \geq 1$ as G^* is connected. If $\deg_{G^*}(S_{i_0}) > \log n$, then $\pi_{G^*}(S_{i_0}) = \deg_{G^*}(S_{i_0})/(2e(G^*)) \geq (\log n)^{1/2}/n$, using that $e(G^*) \leq e(G) \leq Dn$ from (S3). Therefore, using (2.6) and (3.13), we have that

$$\Phi_{\hat{G}}(\hat{S}) \geq \frac{|\partial_{G^*}(S_{i_0})|}{2 \deg_{G^*}(S_{i_0})} = \frac{\Phi_{G^*}(S_{i_0}) \deg_{G^*}(V(G^*) \setminus S_{i_0})}{2e(G^*)} \geq \frac{\Phi_{G^*}(S_{i_0})}{2} \geq \frac{\alpha^2}{16},$$

where we used Lemma 3.7 in last inequality and the fact that $\pi_{G^*}(S_{i_0}) = 1 - \pi_{G^*}(V(G^*) \setminus S_{i_0}) \leq 1/2$ in the penultimate inequality.

So we have established that $\Phi_{\hat{G}}(\hat{S}) \geq 1/(2 \log n)$ for all $\hat{G}$-connected sets $\hat{S}$ with $\pi_{\hat{G}}(\hat{S}) \leq 1/2$. Now notice that $\pi_{\min}(\hat{G}) \geq 1/(2e(\hat{G})) \geq 1/(2Dn)$ due to (S3), and hence, when applying Theorem 2.1, there are logarithmically many terms in the sum. This establishes the desired upper bound on $t_{mix}(\hat{G})$. ◄

Using (1.1) and Claim 3.11 and as the total variation distance decreases exponentially fast after the mixing time (see, e.g., [33, section 4.5]), we have

$$\max_{x,y \in V(\hat{G})} |\mu_0^x P_{\hat{G}}^T(y) - \pi_{\hat{G}}(y)| \le 2 \max_{x \in V(\hat{G})} d_{\mathrm{TV}}(\mu_0^x P_{\hat{G}}^T, \pi_{\hat{G}}) = o(n^{-3})$$

and (HP1) is satisfied.

Let us now prove that $\pi_{\hat{G}}(u^*)$ is small. By Remark 2.3, to prove our statement it suffices to have (HP2$'$) for $u = u^*$. By Lemma 3.4 (3.5), $\pi_G(U) = o((\log n)^{-6})$. Moreover, as mentioned earlier, $\pi_{\hat{G}}(u^*) = \pi_{G^*}(U^*)$. By Lemma 3.5, we have that

$$\pi_{\hat{G}}(u^*) = \pi_G(U) + O(d_{\mathrm{TV}}(\pi_G, \pi_{G^*})) = o((\log n)^{-6}). \tag{3.14}$$

It follows that $T \cdot \pi_{\hat{G}}(u^*) = o(1)$ and (HP2$'$) holds.

Finally, since G is a connected (α, D)-spreader graph with D fixed and $\alpha < 1$, $\hat{G}$ is a connected multigraph with $e(\hat{G}) \le e(G) \le Dn$, by (S3). It follows by (2.2) that $\pi_{\min}(\hat{G}) \ge 1/2e(\hat{G}) = \omega(n^{-2})$ and (HP3) is satisfied.

Now, recalling the relevant definitions from Theorem 2.2 and letting $T_0 := \lceil (\log n)^2 \rceil$, as n goes to infinity we have

$$\lambda_{u^*}^{T_0} = \left(1 - (1 + o(1))\frac{\pi_{\hat{G}}(u^*)}{R_T(u^*)}\right)^{T_0} \ge 1 - (1 + o(1))\frac{T_0 \pi_{\hat{G}}(u^*)}{R_T(u^*)} = 1 - o(1),$$

where we used $(1 - x)^{T_0} \ge 1 - T_0 x$, (3.14) and $R_T(u^*) \ge 1$. Therefore, appealing to Remark 2.4, we have that

$$\frac{1}{n} \sum_{v \in V(G) \setminus U} \mathbb{P}[\hat{\tau}_{u^*}^v \le T_0] \le \frac{1}{|V(\hat{G})|} \sum_{v \in V(G) \setminus U} \mathbb{P}[\hat{\tau}_{u^*}^v \le T_0]$$

$$\le \frac{1}{|V(\hat{G})|} \sum_{v \in V(\hat{G})} \mathbb{P}[\hat{\tau}_{u^*}^v \le T_0] = o(1).$$

Together with (3.12), this completes the proof of the lemma. $\qquad\square$

3.6 Proof of the main theorem

We are finally ready to prove Theorem 1.5.

Proof of Theorem 1.5. Let $U = U(\alpha, D)$ and $G^* = G^*(\alpha, D)$. Let $(X_t)_{t \ge 0}$ and $(X_t^*)_{t \ge 0}$ denote lazy random walks on G and G^*, respectively. For $x \in V(G)$, let $\tau_U^x := \tau_G(\mu_0^x, U)$. Similarly as in the proof of Lemma 3.10, we consider a natural coupling $(X_t, X_t^*)_{t \ge 0}$ of the random walks so that for $t < \tau_U$ we let $X_t^* = X_t$ and otherwise we let the walks evolve independently.

First, observe that, by Lemma 3.5 and the triangle inequality (and adopting the abuse of notation introduced in Section 3.3), for any $x \in V(G)$ we have

$$d_{\mathrm{TV}}(\mu_0^x P_G^t, \pi_G) \le d_{\mathrm{TV}}(\mu_0^x P_G^t, \pi_{G^*}) + d_{\mathrm{TV}}(\pi_G, \pi_{G^*}) = d_{\mathrm{TV}}(\mu_0^x P_G^t, \pi_{G^*}) + o(1). \tag{3.15}$$

For every $x \in V(G) \setminus U$ and $y \in V(G)$, we can write

$$\begin{aligned}
\mu_0^x P_G^t(y) &= \mathbb{P}[X_t = y \mid X_0 = x] \\
&= \mathbb{P}[X_t = y, \tau_U > t \mid X_0 = x] + \mathbb{P}[X_t = y, \tau_U \le t \mid X_0 = x] \\
&= \mathbb{P}[X_t^* = y, \tau_U > t \mid X_0 = x] + \mathbb{P}[X_t = y, \tau_U \le t \mid X_0 = x] \\
&\le \mu_0^x P_{G^*}^t(y) + \mathbb{P}[X_t = y, \tau_U \le t \mid X_0 = x],
\end{aligned}$$

where we let $\mu_0^x P_{G^*}^t(y) = 0$ if $y \in U$. Let $A := \{y \in V(G) : \mu_0^x P_G^t(y) \geq \pi_{G^*}(y)\}$. It follows that

$$
\begin{aligned}
d_{\text{TV}}(\mu_0^x P_G^t, \pi_{G^*}) &= \sum_{y \in A} (\mu_0^x P_G^t(y) - \pi_{G^*}(y)) \\
&\leq \sum_{y \in A} (\mu_0^x P_{G^*}^t(y) - \pi_{G^*}(y)) + \sum_{y \in A} \mathbb{P}[X_t = y, \tau_U \leq t \mid X_0 = x] \\
&\leq d_{\text{TV}}(\mu_0^x P_{G^*}^t, \pi_{G^*}) + \mathbb{P}[\tau_U^x \leq t].
\end{aligned}
\tag{3.16}
$$

Let $C > 0$ be the constant given by Proposition 3.6 with $\varepsilon/2$ playing the role of ε, and let $T_0 := \lceil C \log n \rceil$. By Proposition 3.6, we have that $d_{\text{TV}}(\mu_0^x P_{G^*}^{T_0}, \pi_{G^*}) \leq \varepsilon/2$. Combining these bounds with (3.15) and (3.16), we obtain for all $x \in V(G) \setminus U$ that

$$
d_{\text{TV}}(\mu_0^x P_G^{T_0}, \pi_G) \leq \varepsilon/2 + \mathbb{P}[\tau_U^x \leq T_0] + o(1).
$$

By Lemma 3.4 (3.3) and Lemma 3.10, we conclude that

$$
\frac{1}{n} \sum_{x \in V(G)} d_{\text{TV}}(\mu_0^x P_G^{T_0}, \pi_G) \leq \frac{|U|}{n} + \frac{n - |U|}{n} \left(\frac{\varepsilon}{2} + o(1) \right) + \frac{1}{n} \sum_{x \in V(G) \setminus U} \mathbb{P}[\tau_U^x \leq T_0] \leq \varepsilon
$$

and $\bar{t}_{\text{mix}}(G, \varepsilon) \leq T_0$, concluding the proof. $\qquad \square$

4 Smoothed analysis on connected graphs

We next want to show applications of Theorem 1.5. We use this section to prove Theorem 1.6.

Proof of Theorem 1.6. Let $D := 2(\Delta + 1 + \delta)$ and $0 < \alpha < \delta/D^2$ be a sufficiently small constant. If G' is a connected (α, D)-spreader graph, the statement follows from Theorem 1.5. Since G' is connected by assumption, it suffices to show that a.a.s. G' is an (α, D)-spreader graph.

We need to verify that (S1), (S2) and (S3) hold a.a.s. in $G' = G \cup R$. The fact that (S2) and (S3) hold a.a.s. follows directly from the fact that G is Δ-degenerate and that $e(R[S]) < \max\{2, 2\delta\}|S|$ for all $S \subseteq V(R)$ (see, for example, [31, Lemma 8]). The fact that property (S1) holds a.a.s. in G' follows from the proof of [31, Theorem 3]. Indeed, for each $(\log n)^{1/5} \leq k \leq (1 - 1/D^2)n$, let X_k denote the number of G'-connected α-thin sets $S \subseteq V(G)$ with $|S| = k$. Then (see [31, eq. (4)] and the claim immediately after), one can check that the expected number of such sets satisfies

$$
\mathbb{E}[X_k] \leq \sum_{m=1}^{\alpha k} \sum_{b=m}^{\alpha k} n^m \exp\left(C\alpha \log(1/\alpha)k\right) \left(\frac{\delta}{n}\right)^{m-1} \mathbb{P}\left[\text{Bin}\left(k(n-k), \frac{\delta}{n}\right) < \alpha k \right],
$$

where C is some absolute constant. Following again the proof in [31] and choosing α sufficiently small and n sufficiently large, one concludes that

$$
\mathbb{E}[X_k] \leq k^2 n \exp\left(\left(C\alpha \log(1/\alpha) - \frac{\delta}{8D^2} \right) k \right) \leq n \exp\left(-\frac{\delta}{20D^2} k \right).
$$

Property (S1) then follows by Markov's inequality and a union bound over all $(\log n)^{1/5} \leq k \leq (1 - 1/D^2)n$. $\qquad \square$

5 Random subgraphs of expanders

In order to prove Theorem 1.9, we will rely on several known properties of the giant component of a random subgraph of an (n, d, λ)-graph. Recall from Section 2.1 that when we refer to asymptotic statements holding in *an* (n, d, λ)-graph, implicitly what is meant is that the statement holds for any sequence $(G_n)_{n \geq 1}$ of (n, d, λ)-graphs that satisfy the stated condition.

Lemma 5.1. *Let $\delta > 0$ be a sufficiently small constant and let G be an (n, d, λ)-graph with $\lambda \leq \delta^4 d$. Let $p = (1 + \delta)/d$ and let L_1 be a largest component in G_p. Then, a.a.s. the following properties hold:*

(a) *L_1 has $(1 + o(1))(2\delta + g(\delta))n$ vertices, where $g(\delta) = o(\delta)$ as δ tends to 0.*

(b) *There exists some absolute constant $c > 0$ such that, for any G_p-connected $S \subseteq V(L_1)$ with $16 \log n / \delta^2 \leq |S| \leq \delta^2 n / 50$, we have that*

$$|\partial_{G_p}(S)| \geq \frac{c\delta^2 |S|}{\log(1/\delta)}.$$

(c) *There exists some absolute constant $c' > 0$ such that, for any $S \subseteq V(L_1)$ with $\delta^2 n / 50 \leq |S| \leq 12\delta n / 11$, we have that*

$$|\partial_{G_p}(S)| \geq \frac{c'\delta^2 |S|}{\log(1/\delta)}.$$

Proof. Statement (a) follows from [25, Theorem 1]; see also the discussion following [17, Theorem 1.1]. Statement (b) is a consequence of [17, Theorem 1 (1)], while (c) is given in [17, Theorem 2]. □

With this, we can prove Theorem 1.9.

Proof of Theorem 1.9. Let $L_1 \coloneqq L_1(G_p)$. Fix $D \coloneqq 12$, $c_0 \coloneqq \min\{c, c'\}$ (where c and c' are the absolute constants from Lemma 5.1(b) and (c)) and $\alpha \coloneqq c_0 \delta^2 / (D^2 \log(1/\delta))$. By Theorem 1.5, it suffices to show that a.a.s. L_1 is an (α, D)-spreader graph. That is, we need to show that (for sufficiently small $\delta > 0$) a.a.s. L_1 satisfies properties (S1)–(S3).

We are first going to show that (S1) and (S2) hold a.a.s. in G_p, rather than L_1, for sets of size at least $(\log n)^{1/6}$. Similarly as happened in the proof of Theorem 1.6, in this case properties (S1) (for $|S| \leq \delta^2 n / 50$) and (S2) can be obtained by following the proofs of [17, Theorem 1 (1)] and [17, Lemma 2.4], respectively. Let us give here a brief sketch. Let us first consider (S1) (for $|S| \leq \delta^2 n / 50$). For each $(\log n)^{1/6} \leq k \leq \delta^2 n / 50$, let X_k denote the number of G_p-connected sets $S \subseteq V(G)$ with $|S| = k$ which are α-thin. Then, following [17, Theorem 1 (1)], we have $\mathbb{E}[X_k] \leq 3n \exp(-\delta^2 k / 8)$. By Markov's inequality and a union bound over all $(\log n)^{1/6} \leq k \leq \delta^2 n / 50$, we conclude that a.a.s.

(d) for all $(\log n)^{1/6} \leq k \leq \delta^2 n / 50$, the number of G_p-connected α-thin sets $S \subseteq V(G)$ with $|S| = k$ is less than $ne^{-\sqrt{k}}$.

Similarly, for each $k \geq (\log n)^{1/6}$, one can check from the proof of [17, Lemma 2.4] that, if we let Y_k denote the number of G_p-connected sets $S \subseteq V(G)$ with $|S| = k$ such that $e_{G_p}(S) \geq 10k$, then $\mathbb{E}[Y_k] \leq n \exp(-2k)$. Again by Markov's inequality and a union bound over all $k \geq (\log n)^{1/6}$, we conclude that a.a.s.

(e) for all $(\log n)^{1/6} \leq k \leq n$, the number of G_p-connected α^{-1}-loaded sets $S \subseteq V(G)$ with $|S| = k$ is less than $ne^{-\sqrt{k}}$.

We next work towards property (S3). We are going to show that no set $S \subseteq V(G)$ with $|S| \geq \delta\alpha n$ is D-loaded in G_p. Fix some $\delta\alpha n \leq k \leq n$ and let $\eta := k/n$. If $\delta > 0$ is sufficiently small, we have that $\delta^4 < \alpha\delta \leq \eta$. Now, by the expander mixing lemma (see, for example, [17, Lemma 2.1]), for any set $S \subseteq V(G)$ with $|S| = k = \eta n$ we have that

$$e_G(S) \leq \frac{dk^2}{n} + \lambda k \leq (\eta + \delta^4)dk \leq 2\eta dk,$$

using that $\lambda \leq \delta^4 d$. Hence, the probability that S contains Dk edges in G_p is at most

$$\binom{2\eta dk}{Dk} p^{Dk} \leq \left(\frac{2\eta de}{D}\right)^{Dk} \left(\frac{1+\delta}{d}\right)^{Dk} \leq \left(\frac{6\eta}{D}\right)^{Dk}.$$

Taking a union bound over all possible sets of size k, we have that the probability of there existing a D-loaded set of size k in G_p is at most

$$\binom{n}{k} \left(\frac{6\eta}{D}\right)^{Dk} \leq \left(\frac{ne}{k}\right)^k \left(\frac{6\eta}{D}\right)^{Dk} = \left(\frac{e}{\eta} \left(\frac{6\eta}{D}\right)^D\right)^k = \left(e\eta^{D-1} \left(\frac{1}{2}\right)^D\right)^k \leq \left(\frac{1}{2}\right)^{\alpha\delta n}.$$

By a union bound over all $\alpha\delta n \leq k \leq n$ we conclude that a.a.s.

  (f) no set $S \subseteq V(G)$ with $|S| \geq \delta\alpha n$ is D-loaded in G_p.

Condition on the event that Lemma 5.1(a) and (c) as well as (d), (e) and (f) hold in G_p, which a.a.s. occurs. Let $n' := |V(L_1)| = (1 + o(1))(2\delta + g(\delta))n$ by (a). It follows that $(\log n')^{1/5} \geq (\log n)^{1/6}$ and so, by (d), (S1) holds in L_1 for sets of size at most $\delta^2 n/50$; similarly, (S2) holds by (e), and (S3) holds by (f). Thus, it only remains to establish (S1) for sets $S \subseteq V(L_1)$ such that $\delta^2 n/50 \leq |S| \leq (1 - 1/D^2)n'$. For $\delta^2 n/50 \leq |S| \leq 12\delta n/11$, this is immediate from (c), and we can also use (c) for larger sets S. Indeed, suppose $12\delta n/11 \leq |S| \leq (1 - 1/D^2)n'$ and let $\bar{S} := V(L_1) \setminus S$. It follows from (a), by taking a sufficiently small δ, that

$$|\bar{S}| = |L_1| - |S| \leq n' - \frac{12}{11}\delta n < \frac{n'}{2} \leq \frac{12}{11}\delta n,$$

where in the last two inequalities we use that $\frac{\delta}{2\delta + g(\delta)}$ tends to $1/2$ as δ tends to 0 due to the fact that $g(\delta) = o(\delta)$. We also have that $|\bar{S}| \geq n'/D^2 \geq |S|/D^2 \geq \delta^2 n/50$, using also here that δ is sufficiently small. Hence, we can apply (c) to $\bar{S}$ and we obtain that

$$|\partial_{L_1}(S)| = |\partial_{L_1}(\bar{S})| \geq \frac{c'\delta^2|\bar{S}|}{\log(1/\delta)} \geq \frac{c'\delta^2|S|}{D^2 \log(1/\delta)} \geq \alpha|S|. \qquad \square$$

6 Open problems

Theorem 1.5 is only effective on graphs where the mixing is slowed down by few small bottlenecks. This is the case in the two applications presented. Nevertheless, there are other cases where both small and large bottlenecks exist. It would be interesting to study average-case mixing times in such scenarios and determine which improvement with respect to the worst-case can be attained. One such example is the small-world model of Kleinberg [30], whose mixing time has been studied in [20].

In recent years, the theory of random walks in random directed graphs has attracted a considerable amount of attention. As in the case of random regular graphs, under mild conditions on the bidegree sequence, the mixing time is logarithmic [9, 12]. From the point of view of smoothed analysis, a natural question is whether randomly perturbing a deterministic strongly connected digraph can yield logarithmic mixing time.

Conductance-based bounds such as Jerrum-Sinclair and Fountoulakis-Reed are not valid in the non-reversible setting, which requires new ideas. Finally, we mention an analogous result to the mixing time in randomly perturbed connected graphs by Krivelevich, Reichman and Samotij [31], for graphs perturbed by a random perfect matching, has been obtained by Hermon, Sly and Sousi [27]. Considering such a model in the directed setting would also be interesting.

References

[1] L. Addario-Berry and T. Lei. The mixing time of the Newman-Watts small-world model. *Adv. Appl. Probab.*, 47(1):37–56, 2015, doi: 10.1239/aap/1427814580. MR3327314

[2] M. Ajtai, J. Komlós, and E. Szemerédi. Largest random component of a k-cube. *Combinatorica*, 2:1–7, 1982, doi: 10.1007/BF02579276. MR0671140

[3] D. Arthur, B. Manthey, and H. Röglin. Smoothed analysis of the k-means method. *J. ACM*, 58(5):No. 19, 2011, doi: 10.1145/2027216.2027217. MR2853654

[4] A. Ben-Hamou and J. Salez. Cutoff for nonbacktracking random walks on sparse random graphs. *Ann. Probab.*, 45(3):1752–1770, 2017, doi: 10.1214url/16-AOP1100. MR3650414

[5] I. Benjamini, G. Kozma, and N. Wormald. The mixing time of the giant component of a random graph. *Random Struct. Algorithms*, 45(3):383–407, 2014, doi: 10.1002/rsa.20539. MR3252922

[6] N. Berestycki, E. Lubetzky, Y. Peres, and A. Sly. Random walks on the random graph. *Ann. Probab.*, 46(1):456–490, 2018, doi: 10.1214/17-AOP1189. MR3758735

[7] A. Bhaskara, M. Charikar, A. Moitra, and A. Vijayaraghavan. Smoothed analysis of tensor decompositions. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '14, pages 594–603, New York, NY, USA, 2014. Association for Computing Machinery, doi: 10.1145/2591796.2591881.

[8] T. Bohman, A. Frieze, and R. Martin. How many random edges make a dense graph Hamiltonian? *Random Struct. Algorithms*, 22(1):33–42, 2003, doi: 10.1002/rsa.10070. MR1943857

[9] C. Bordenave, P. Caputo, and J. Salez. Random walk on sparse random digraphs. *Probab. Theory Relat. Fields*, 170(3):933–960, 2018, doi: 10.1007/s00440-017-0796-7. MR3773804

[10] J. Böttcher, J. Han, Y. Kohayakawa, R. Montgomery, O. Parczyk, and Y. Person. Universality for bounded degree spanning trees in randomly perturbed graphs. *Random Struct. Algorithms*, 55(4):854–864, 2019, doi: 10.1002/rsa.20850. MR4025392

[11] J. Böttcher, R. Montgomery, O. Parczyk, and Y. Person. Embedding spanning bounded degree graphs in randomly perturbed graphs. *Mathematika*, 66(2):422–447, 2020, doi: 10.1112/mtk.12005. MR4130332

[12] X. S. Cai, P. Caputo, G. Perarnau, and M. Quattropani. Rankings in directed configuration models with heavy tailed in-degrees. *arXiv e-prints*, 2021, arXiv:2104.08389. MR4677741

[13] F. Chung and L. Lu. The diameter of sparse random graphs. *Adv. Appl. Math.*, 26(4):257–279, 2001, doi: 10.1006/aama.2001.0720. MR1826308

[14] A. Coja-Oghlan, U. Feige, A. Frieze, M. Krivelevich, and D. Vilenchik. On smoothed k-CNF formulas and the Walksat algorithm. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '09, pages 451–460, New York, NY, USA, 2009. Association for Computing Machinery, Society for Industrial and Applied Mathematics, doi: 10.1137/1.9781611973068.50. MR2809346

[15] C. Cooper and A. Frieze. The cover time of random regular graphs. *SIAM J. Discrete Math.*, 18(4):728–740, 2005, doi: 10.1137/S0895480103428478. MR2157821

[16] J. Ding, E. Lubetzky, and Y. Peres. Anatomy of the giant component: the strictly supercritical regime. *Eur. J. Comb.*, 35:155–168, 2014, doi: 10.1016/j.ejc.2013.06.004. MR3090494

[17] S. Diskin and M. Krivelevich. Expansion in supercritical random subgraphs of expanders and its consequences. *arXiv e-prints*, 2022, arXiv:2205.04852. MR4515692

[18] R. Durrett. *Random graph dynamics*, volume 20 of *Camb. Ser. Stat. Probab. Math.* Cambridge University Press, Cambridge, 2006, doi: 10.1017/CBO9780511546594. MR2271734

[19] M. Dyer, A. Frieze, and R. Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *J. Assoc. Comput. Mach.*, 38(1):1–17, 1991, doi: 10.1145/102782.102783. MR1095916

[20] M. E. Dyer, A. Galanis, L. A. Goldberg, M. Jerrum, and E. Vigoda. Random walks on small world networks. *ACM Trans. Algorithms*, 16(3):article 37, 2020, doi: 10.1145/3382208. MR4120525

[21] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci., Ser. A*, 5(1):17–60, 1960. MR0125031

[22] U. Feige. Refuting smoothed 3CNF formulas. In *48th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '07, pages 407–417, Providence, RI, USA, 2007. Institute of Electrical and Electronics Engineers, doi: 10.1109/FOCS.2007.16.

[23] N. Fountoulakis and B. A. Reed. Faster mixing and small bottlenecks. *Probab. Theory Related Fields*, 137(3-4):475–486, 2007, doi: 10.1007/s00440-006-0003-8. MR2278465

[24] N. Fountoulakis and B. A. Reed. The evolution of the mixing rate of a simple random walk on the giant component of a random graph. *Random Struct. Algorithms*, 33(1):68–86, 2008, doi: 10.1002/rsa.20210. MR2428978

[25] A. Frieze, M. Krivelevich, and R. Martin. The emergence of a giant component in random subgraphs of pseudo-random graphs. *Random Struct. Algorithms*, 24(1):42–50, 2004, doi: 10.1002/rsa.10100. MR2020308

[26] J. Han, P. Morris, and A. Treglown. Tilings in randomly perturbed graphs: bridging the gap between Hajnal-Szemerédi and Johansson-Kahn-Vu. *Random Struct. Algorithms*, 58(3):480–516, 2021, doi: 10.1002/rsa.20981. MR4234994

[27] J. Hermon, A. Sly, and P. Sousi. Universality of cutoff for graphs with an added random matching. *Ann. Probab.*, 50(1):203–240, 2022, doi: 10.1214/21-AOP1532. MR4385126

[28] M. Jerrum and A. Sinclair. Conductance and the rapid mixing property for markov chains: the approximation of permanent resolved. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, STOC '88, pages 235–244, New York, NY, USA, 1988. Association for Computing Machinery, doi: 10.1145/62212.62234. MR1025467

[29] A. T. Kalai, A. Samorodnitsky, and S.-H. Teng. Learning and smoothed analysis. In *50th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '09, pages 395–404, Atlanta, GA, USA, 2009. Institute of Electrical and Electronics Engineers, doi: 10.1109/FOCS.2009.60. MR2648420

[30] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, STOC '00, pages 163–170, New York, NY, USA, 2000. Association for Computing Machinery, doi: 10.1145/335305.335325. MR2114529

[31] M. Krivelevich, D. Reichman, and W. Samotij. Smoothed analysis on connected graphs. *SIAM J. Discrete Math.*, 29(3):1654–1669, 2015, doi: 10.1137/151002496. MR3395121

[32] M. Krivelevich and B. Sudakov. Pseudo-random graphs. In *More sets, graphs and numbers*, pages 199–262. Springer, 2006, doi: 10.1007/978-3-540-32439-3_10. MR2223394

[33] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009, doi: 10.1090/mbk/058. With a chapter by James G. Propp and David B. Wilson. MR2466937

[34] E. Lubetzky and A. Sly. Cutoff phenomena for random walks on random regular graphs. *Duke Math. J.*, 153(3):475–510, 2010, doi: 10.1215/00127094-2010-029. MR2667423

[35] F. Manzo, M. Quattropani, and E. Scoppola. A probabilistic proof of Cooper and Frieze's "First Visit Time Lemma". *arXiv e-prints*, 2021, arXiv:2101.10748. MR4332217

[36] M. E. J. Newman and D. J. Watts. Renormalization group analysis of the small-world network model. *Phys. Lett., A*, 263(4-6):341–346, 1999, doi: 10.1016/S0375-9601(99)00757-4. MR1732095

[37] M. E. J. Newman and D. J. Watts. Scaling and percolation in the small-world network model. *Phys. Rev. E*, 60:7332–7342, 1999, doi: 10.1103/PhysRevE.60.7332.

[38] D. Randall. Rapidly mixing markov chains with applications in computer science and physics. *Comput. Sci. Eng.*, 8(2):30–41, 2006, doi: 10.1109/MCSE.2006.30.

[39] A. Sankar, D. A. Spielman, and S.-H. Teng. Smoothed analysis of the condition numbers and growth factors of matrices. *SIAM J. Matrix Anal. Appl.*, 28(2):446–476, 2006, doi: 10.1137/S0895479803436202. MR2255338

[40] D. A. Spielman and S.-H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *J. ACM*, 51(3):385–463, 2004, doi: 10.1145/990308.990310. MR2145860

[41] D. A. Spielman and S.-H. Teng. Smoothed analysis: an attempt to explain the behavior of algorithms in practice. *Commun. ACM*, 52(10):76–84, 2009, doi: 10.1145/1562764.1562785.

[42] E. Vigoda. Improved bounds for sampling colorings. *J. Math. Phys.*, 41(3):1555–1569, 2000, doi: 10.1063/1.533196. MR1757969

[43] V. H. Vu and T. Tao. The condition number of a randomly perturbed matrix. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '07, pages 248–255, New York, NY, USA, 2007. Association for Computing Machinery, doi: 10.1145/1250790.1250828. MR2402448

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4], ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System: https://vtex.lt/services/ejms-peer-review/
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: https://imstat.org/shop/donation/