Electronic Colloquium on Computational Complexity, Report No. 46 (2019)

Random walks and forbidden minors II: A poly($d\varepsilon^{-1}$)-query tester for minor-closed properties of bounded degree graphs

Akash Kumar* C. Seshadhri† Andrew Stolman‡

Abstract

Let G be a graph with n vertices and maximum degree d. Fix some minor-closed property $\mathcal{P}$ (such as planarity). We say that G is ε-far from $\mathcal{P}$ if one has to remove εdn edges to make it have $\mathcal{P}$. The problem of property testing $\mathcal{P}$ was introduced in the seminal work of Benjamini-Schramm-Shapira (STOC 2008) that gave a tester with query complexity triply exponential in ε^{-1}. Levi-Ron (TALG 2015) have given the best tester to date, with a quasipolynomial (in ε^{-1}) query complexity. It is an open problem to get property testers whose query complexity is poly($d\varepsilon^{-1}$), even for planarity.

In this paper, we resolve this open question. For any minor-closed property, we give a tester with query complexity $d \cdot \text{poly}(\varepsilon^{-1})$. The previous line of work on (independent of n, two-sided) testers is primarily combinatorial. Our work, on the other hand, employs techniques from spectral graph theory. This paper is a continuation of recent work of the authors (FOCS 2018) analyzing random walk algorithms that find forbidden minors.

*Department of Computer Science, Purdue University. akumar@purdue.edu (Supported by NSF CCF-1618981.)

†Department of Computer Science, University of California, Santa Cruz. sesh@ucsc.edu (Supported by NSF TRIPODS grant CCF-1740850 and NSF CCF-1813165)

‡Department of Computer Science, University of California, Santa Cruz. astolman@ucsc.edu (Supported by NSF TRIPODS grant CCF-1740850)

ISSN 1433-8092

1 Introduction

The classic result of Hopcroft-Tarjan gives a linear time algorithm for deciding planarity [HT74]. As the old theorems of Kuratowski and Wagner show, planarity is characterized by the non-existence of K_5 and $K_{3,3}$ minors [Kur30, Wag37]. The monumental graph minor theorem of Robertson-Seymour proves that any property of graphs closed under minors can be expressed by the non-existence of a finite list of minors [RS95a, RS95b, RS04]. Moreover, given a fixed graph, H, the property of being H-minor-free can be decided in quadratic time [KKR12]. Thus, any minor-closed property of graphs can be decided in quadratic time.

What if an algorithm is not allowed to read the whole graph? This question was first addressed in the seminal result of Benjamini-Schramm-Shapira (BSS) in the language of property testing [BSS08]. Consider the model of random access to a graph adjacency list, as introduced by Goldreich-Ron [GR02]. Let $G = (V, E)$ be a graph where $V = [n]$ and the maximum degree is d. We have random access to the list through *neighbor queries*. There is an oracle that, given $v \in V$ and $i \in [d]$, returns the ith neighbor of v (if no neighbor exists, it returns $\perp$).

For a property $\mathcal{P}$ of graphs with degree bound d, the distance of G to $\mathcal{P}$ is the minimum number of edge additions/removals required to make G have $\mathcal{P}$, divided by dn. We say that G is ε-far from $\mathcal{P}$ if the distance to $\mathcal{P}$ is more than ε. A property tester for $\mathcal{P}$ is a randomized procedure that takes as input (query access to) G and a proximity parameter, $\varepsilon > 0$. If $G \in \mathcal{P}$, the tester must accept with probability at least $2/3$. If G is ε-far from $\mathcal{P}$, the tester must reject with probability at least $2/3$. A tester is one-sided if it accepts $G \in \mathcal{P}$ with probability 1.

Let $\mathcal{P}$ be some minor-closed property such as planarity. BSS proved the remarkable result that any such $\mathcal{P}$ is testable in time independent of n. Their query complexity was triply exponential in (d/ε). Hassidim-Kelner-Nguyen-Onak improved this complexity to singly exponential, introducing the novel concept of partition oracles [HKNO09]. Levi-Ron gave a more efficient analysis, proving the existence of testers with query complexity quasi-polynomial in (d/ε) [LR15]. For the special cases of outerplanarity and bounded treewidth, $\mathrm{poly}(d/\varepsilon)$ query testers are known [YI15, EHNO11].

It has been a significant open problem to get $\mathrm{poly}(d/\varepsilon)$ query testers for all minor-closed properties. In Open Problem 9.26 of Goldreich's recent book on property testing, he states the "begging question of whether [the query complexity bound of testing minor-closed properties] can be improved to a polynomial [in $1/\varepsilon$]" [Gol17]. Even for classic case of planarity, this was unknown.

In this paper, we resolve this open problem.

Theorem 1.1. *Let $\mathcal{P}$ be any minor-closed property of graphs with degree bound d. There exists a (two-sided) tester for $\mathcal{P}$ that runs in $d^2 \cdot \mathrm{poly}(\varepsilon^{-1})$ time.*

Thus, properties such as planarity, series-parallel graphs, embeddability in bounded genus surfaces, linkless embeddable, and bounded treewidth are all testable in time $d^2 \cdot \mathrm{poly}(\varepsilon^{-1})$.

By the graph minor theorem of Robertson-Seymour [RS04], Theorem 1.1 is a corollary of our main result for testing H-minor-freeness. As alluded to earlier, for any minor-closed property $\mathcal{P}$, there exists a finite list of graphs $\{H_1, H_2, \ldots, H_b\}$ satisfying the following condition. A graph G is in $\mathcal{P}$ iff for all $i \leq b$, G does not contain an H_i-minor. Let $\mathcal{P}_{H_i}$ be the property of being H_i-minor-free. The characterization implies that if G is ε-far from $\mathcal{P}$, there exists $i \leq b$ such that G is $\Omega(\varepsilon)$-far from $\mathcal{P}_{H_i}$. Thus, property testers for H_i-minor freeness imply property testers for $\mathcal{P}$ (with constant blowup in the proximity parameter).

Our main quantitative theorem follows.

Theorem 1.2. *There is an absolute constant c such that the following holds. Fix a graph H with r vertices. The property of being H-minor-free is testable in $d(r/\varepsilon)^c$ queries and $d^2(r/\varepsilon)^{2c}$ time.*

We stress that c is independent on r. Currently, our value of c is likely more than 100, and we have not tried to optimize the exponent of ε. We believe that significant improvement is possible, even by just tightening the current analysis. It would be of significant interest to get a better bound, even for the case of planarity.

1.1 Related work

Property testing on bounded-degree graphs is a large topic, and we point the reader to Chapter 9 of Goldreich's book [Gol17]. Graph minor theory is immensely deep, and Chapter 12 of Diestel's book is an excellent reference [Die10]. We will focus on the work regarding property testing of H-minor-freeness.

As mentioned earlier, this line of work started with Benjamini-Schramm-Shapira [BSS08]. Their tester basically approximates the frequency of all subgraphs with radius $2^{1/\varepsilon}$, which leads to the large dependence in d/ε. Central to their result (and subsequent) work is the notion of *hyperfiniteness*. A hyperfinite class of graphs has the property that the removal of a small constant fraction of edges leaves connected components of constant size. Hassidim-Kelner-Nguyen-Onak design *partition oracles* for hyperfinite graphs to get improved testers [HKNO09, LR15]. These oracles are local procedures that output the connected component that a vertex lies in, without explicit knowledge of any global partition. This is extremely challenging as one has to maintain consistency among different queries. The final construction is an intricate recursive procedure that makes $\exp(d/\varepsilon)$ queries. Levi-Ron gave a significantly simpler and more efficient analysis leading to their query complexity of $(d\varepsilon^{-1})^{\log \varepsilon^{-1}}$. Newman-Sohler show how partition oracles lead to testers for any property of hyperfinite graphs [NS13].

Given the challenge of $\text{poly}(d\varepsilon^{-1})$ testers for planarity, there has been focus on other minor-closed properties. Yoshida-Ito give such a tester for outerplanarity [YI15], which was subsumed by a $\text{poly}(d\varepsilon^{-1})$ tester by Edelman et al for bounded treewidth graphs [EHNO11]. Nonetheless, $\text{poly}(d\varepsilon^{-1})$ testers for planarity remained open.

Unlike general (two-sided) testers, one-sided testers for H-minor-freeness must have a dependence on n. BSS conjectured that the complexity of testing H-minor-freeness (and specifically planarity) is $\Theta(\sqrt{n})$. Czumaj et al [CGR+14] showed such a lower bound for any H containing a cycle, and gave an $\widetilde{O}(\sqrt{n})$ tester when H is a cycle. Fichtenberger-Levi-Vasudev-Wötzel give an $\widetilde{O}(n^{2/3})$ tester for H-minor-freeness when H is $K_{2,k}$, the $(k\times 2)$-grid or the k-circus graph [FLVW17]. Recently, Kumar-Seshadhri-Stolman (henceforth KSS) nearly resolved the BSS conjecture with an $n^{1/2+o(1)}$-query one-sided tester for H-minor-freeness [KSS18a]. The underlying approach uses the proof strategy of the bipartiteness tester of Goldreich-Ron [GR99].

The body of work on two-sided (independent of n) testers is primarily combinatorial. The proof of Theorem 1.2 is a significant deviation from this line of work, and is inspired by the spectral graph theoretic methods in KSS. As we explain in the next section, we do not require the full machinery of KSS, but we do follow the connections between random walk behavior and graph minors. The tester of Theorem 1.2 is simpler than those of Hassidim et al and Levi-Ron, who use recursive algorithms to construct partition oracles [HKNO09, LR15].

1.2 Main ideas

Let us revisit the argument of KSS, that gives an $n^{1/2+o(1)}$-query one-sided tester for H-minor-freeness. We will take great liberties with parameters, to explain the essence. The proof of Theorem 1.2 is inspired by the approach in KSS, but the proof details deviate significantly. We discover that the full machinery is not required. But the main idea is to exploit connections between random walk behavior and graph minor-freeness.

First, we fix a random walk length $\ell = n^{\delta} \gg 1/\varepsilon$, for small constant $\delta > 0$. One of the building blocks is a random walk procedure that finds H-minors by performing $\sqrt{n} \cdot \mathrm{poly}(\ell)$ random walks of length ℓ. For our purposes, it is not relevant what the algorithm is, and we simply refer to this as the "random walk procedure". One of the significant concepts in KSS is the notion of a *returning random walk*. For any subset of vertices $S \subset V$, an S-returning random walk of length ℓ is a random walk that starts from S and ends at S. For any vertex $s \in S$, we use $\mathbf{q}_{[S],s,\ell}$ to denote the $|S|$-dimensional vector of probabilities of an S-returning walk of length ℓ starting from s.

KSS proves the following two key lemmas. We use c to denote some constant that depends only on H.

1. Suppose there is a subset $S \subseteq V$, $|S| \geq n/\ell$, with the following property. For at least half the vertices $s \in S$, $\|\mathbf{q}_{[S],s,\ell}\| \leq \ell^{-c}$. Then, whp, the $\sqrt{n} \cdot \mathrm{poly}(\ell)$-time random walk procedure finds an H-minor.

2. Suppose there is a subset $S \subseteq V$, $|S| \geq n/\ell$, with the following property. For at least half the vertices $s \in S$, $\|\mathbf{q}_{[S],s,\ell}\| > \ell^{-c}$. Then, for every such vertex s, there is a cut of conductance at most $1/\ell$ contained in S, where all vertices (in the cut) are reached with probability at least $1/\mathrm{poly}(\ell)$ by ℓ-length S-returning walks from s.

To get a one-sided tester, we run the $\sqrt{n} \cdot \mathrm{poly}(\ell)$ random walk procedure. If it does not find an H-minor, then the antecedent of the second part above is true for all S such that $|S| \geq n/\ell$. The consequent basically talks of local partitioning *within* S, even though random walks are performed in the whole graph G. The statement is proven using arguments from local partitioning theorems of Spielman-Teng [ST12]. By iterating the argument, we can prove the existence of a set of εdn edges, whose removal breaks G into connected components of size at most $\mathrm{poly}(\ell)$. Moreover, a superset of any piece can be "discovered" by performing $\mathrm{poly}(\ell)$ random walks (of length ℓ) from some starting vertex. Roughly speaking, each piece has a distinct starting vertex. Thus, if G was ε-far from being H-minor-free, an ε-fraction (by size) of the pieces will contain H-minors. A procedure that picks $\mathrm{poly}(\ell)$ random vertices (to hit the starting vertex of these pieces) and runs $\mathrm{poly}(\ell)$ random walks of length ℓ will, whp, cover a subgraph that contains an H-minor. We refer to this as the "local search procedure", which runs in $\mathrm{poly}(\ell)$ time.

This sums up the KSS approach. Observe that in the first case above, by the probabilistic method, we are guaranteed the existence of a minor. Let us abstract out the argument as follows. Let $\boldsymbol{Q}$ be the statement/condition: there exists a subset $S \subseteq V$, $|S| \geq n/\ell$ such that for at least half the vertices $s \in S$, $\|\mathbf{q}_{[S],s,\ell}\| \leq \ell^{-c}$. KSS basically proves the following lemmas, which we refer to subsequently as Lemma 1 and Lemma 2.

1. $\boldsymbol{Q} \Rightarrow G$ contains an H-minor.

2. $\neg\boldsymbol{Q} \Rightarrow$ If G is ε-far from being H-minor-free, the local search procedure finds an H-minor whp.

We now have an approach to get a $\mathrm{poly}(\varepsilon^{-1})$ tester. Suppose we could set the random walk length ℓ to be $\mathrm{poly}(\varepsilon^{-1})$. And suppose we could test the condition $\boldsymbol{Q}$ in time $\mathrm{poly}(\varepsilon^{-1})$. We could then run local search on top of this, and get a bonafide tester.

A simple adaptation of proofs of both Lemma 1 and Lemma 2 run into some fundamental difficulties. The proof of Lemma 1 crucially requires ℓ to be n^δ (or at least $\Omega(\log n)$). The existence of the minor is shown through the success of the $\sqrt{n} \cdot \mathrm{poly}(\ell)$ random walk procedure. Constant length random walks cannot find an H-minor, even if G was $\Omega(1)$-far from being H-minor-free (G could be a 3-regular expander).

From hyperfiniteness to $\ell = \mathrm{poly}(\varepsilon^{-1})$. We employ a different (and simpler) approach to reduce the walk length. A classic result of Alon-Seymour-Thomas asserts that any H-minor-free bounded-degree graph G satisfies the following "hyperfinite" decomposition: for any $\alpha \in (0, 1)$, we can remove an α-fraction of the edges to get connected components of size $O(\alpha^{-2})$. Let us set $\alpha = \mathrm{poly}(\varepsilon)$ and the walk length $\ell \ll 1/\alpha$. We can show that ℓ-length random walks in G encounter the removed edges with very low probability. By and large, the walks behave as if they were performed on the decomposition. Thus, walks in G are "trapped" in the small components of size $O(\alpha^{-2})$. Quantitatively, we can show that most vertices s, $|\mathbf{p}_{s,\ell}\|_2 \geq \mathrm{poly}(\varepsilon)$. (We use $\mathbf{p}_{s,\ell}$ to denote the random walk distribution with starting vertex s.) By the contrapositive: if there are at least $\mathrm{poly}(\varepsilon)$-fraction of vertices s such that $|\mathbf{p}_{s,\ell}\|_2 \leq \mathrm{poly}(\varepsilon)$, then G contains an H-minor. This is easily testable. We get a more convenient, $\mathrm{poly}(\varepsilon^{-1})$-query testable version of Lemma 1.

Clipped norms for local partitioning. We can now express our new condition $\neg \boldsymbol{Q}$ as: for more than a $(1-\mathrm{poly}(\varepsilon))$-fraction of vertices s, $\|\mathbf{p}_{s,\ell}\|_2 \geq \mathrm{poly}(\varepsilon)$. This is a *weakening* of the antecedent. Previously, the condition referred to returning walks, which have smaller norm. Furthermore, the returning walks specifically reference S, the set in which we are performing local partitioning. Thus, we have some conditions on the behavior of random walks within S itself, which is necessary to perform the local partitioning. Our new condition only refers to the l_2-norms of random walks in G.

The new condition appears to be too fragile to get local partitioning within S. It is possible that the l_2-norm of $\mathbf{p}_{s,\ell}$ is dominated by a few vertices outside of S, whose l_1-norm is tiny. In other words, an event of small probability dominates the l_2-norm. The existing proof of Lemma 2 (from KSS) is not sensitive enough to handle such situations.

We overcome this problem by using a more robust version of norm, called the *clipped norm*. We define $\mathrm{cl}(\boldsymbol{x}, \xi)$ for distribution vector $\boldsymbol{x}$ and $\xi \in (0, 1)$ to be the smallest l_2-norm obtained by removing ξ probability mass (l_1-norm) from $\boldsymbol{x}$. In other words, we can measuring the l_2-norm after "clipping" away ξ probability worth of outliers. We can prove a version of Lemma 2 with a lower bound of the clipped norm. We need to now rework Lemma 1 in terms of clipped norms. This turns out to be relatively straightforward.

Putting it all together. Our final tester is as follows. The length ℓ is set to $\mathrm{poly}(\varepsilon^{-1})$. It picks some random vertices, and estimates the l_2-norm of *clipped* probability vectors of ℓ-length random walks from these vertices. If sufficiently many of them have "small" ($\mathrm{poly}(\varepsilon^{-1})$) norms, then the tester rejects. Otherwise, it runs $\mathrm{poly}(\varepsilon^{-1})$ walks to find a superset of a low conductance cut. The tester employs some exact H-minor finding algorithm on the observed subgraph.

2 The algorithm

In the algorithm and analysis, we will use the following notation.

- **Random walks** - Unless stated otherwise, we consider lazy random walks on graphs. If the walk is at a vertex, v, it transitions to each neighbor of v with probability $1/2d$ and remains at v with probability $1 - \frac{d_v}{2d}$ where d_v is the degree of the vertex v. Note that the stationary distribution is uniform. We use M to denote the transition matrix of this random walk.

- $\mathbf{p}_{v,t}$ - the n-dimensional probability vector, where the uth entry is the probability that a length t random walk started from v ends at u. We denote each entry as $\mathbf{p}_{v,t}(u)$.

- $\|\cdot\|_p$ - the usual l_p norm on vectors.

The two parameters to the algorithm are $\varepsilon \in [0, 1/2]$, and a graph H on $r \geq 3$ vertices. We set the walk length $\ell = \alpha r^3 + \lceil \varepsilon^{-20} \rceil$, where α is some absolute constant.

Our algorithm runs as a subroutine the exact quadratic time minor-finding algorithm of Kawarabayashi-Kobayashi-Reed [KKR12]. We denote this procedure by `KKR`.

`IsMinorFree` (G, ε, H)

1. Pick multiset S of ℓ^{21} uniform random vertices.
2. For every $s \in S$, run `EstClip`(s) and `LocalSearch`(S).
3. If any call to `LocalSearch` returns FOUND, REJECT.
4. If more than $2\ell^{20}$ calls to `EstClip` return LOW, REJECT.
5. ACCEPT

`LocalSearch`(s)

1. Perform ℓ^{21} independent random walks of length ℓ^{11} from s. Add all the vertices encountered to set B_s.
2. Determine $G[B_s]$, the subgraph induced by B_s.
3. If `KKR`$(G[B_s], H)$ finds an H-minor, return FOUND.

`EstClip`(s)

1. Perform $w = \ell^{14}$ walks of length ℓ from s.
2. For every vertex v, let $w_v =$ number of walks that end at v.
3. Let $T = \{v \mid w_v \geq \ell^7/2\}$.
4. If $\sum_{v \in T} w_v \geq w/3$, output HIGH, else output LOW.

Theorem 1.2 follows directly from the following theorems.

Theorem 2.1. *If G is H-minor-free,* `IsMinorFree` *outputs ACCEPT with probability at least* $2/3$.

Theorem 2.2. *If G is ε-far from H-minor-freeness, then* `IsMinorFree` *outputs REJECT with probability at least* $2/3$.

Claim 2.3. *There exists an absolute constant, c such that the query complexity of* `IsMinorFree` *is $O(d(r/\varepsilon)^c)$ and time complexity is $O(d^2(r/\varepsilon)^2)$.*

Proof. The entire algorithm is based on performing poly(ℓ) random walks of length poly(ℓ). Note that $\ell = \text{poly}(r/\varepsilon)$. The dependence on d appears because the subgraph $G[B_s]$ is constructed by query the neighborhood of all vertices in B_s. The quadratic overhead in running time is because of KKR. $\square$

3 Random walks do not spread in minor-free graphs

We first define the clipped norm.

Definition 3.1. *Given $\boldsymbol{x} \in (\mathbb{R}^+)^{|V|}$ and parameter $\xi \in [0,1)$, the ξ-clipped vector $\mathrm{cl}(\boldsymbol{x},\xi)$ is the lexicographically least vector $\boldsymbol{y}$ optimizing the program: $\min \|\boldsymbol{y}\|_2$, subject to $\|\boldsymbol{x}-\boldsymbol{y}\|_1 \le \xi$ and $\forall v \in V, \boldsymbol{y}(v) \le \boldsymbol{x}(v)$.*

The clipping operation removes "outliers" from a vector, with the intention of minimizing the l_2-norm. For a probability distribution $\mathbf{p}_{s,\ell}$, a small value of $\|\mathbf{p}_{s,\ell}\|_2^2$ is a measure of the spread of the walk. But this is a crude lens. There may be one large coordinate in $\mathbf{p}_{s,\ell}$ that determines the norm, while all other coordinates are (say) uniform. The clipped norm better captures (for our purposes) the notion of a random walk spreading.

We state the main result of this section. The constant 3/8 below is just for convenience, and can be replaced by any non-zero constant (with a constant drop in the lower bound).

Lemma 3.2. *There is an absolute constant α such that the following holds. Let H be a graph on r vertices. Suppose G is a H-minor-free graph. Then for any $\ell \ge \alpha r^3$, there exists at least $(1-1/\ell)n$ vertices such that $\|\mathrm{cl}(\mathbf{p}_{v,\ell},3/8)\|_2^2 \ge \ell^{-7}$.*

In order to show this lemma, we will use the classic decomposition theorem for minor-free graphs by Alon-Seymour-Thomas [AST90]. It originally appears phrased in terms of a weight function $w : V \to \mathbb{R}^+$. We use the uniform weight function $\forall v \in V \;\; w(v) = 1/n$ to obtain the restatement below.

Lemma 3.3 (Proposition 4.1 of [AST90])**.** *There is an absolute constant α such that the following holds. Let H be a graph on r vertices. Suppose G is an H-minor-free graph with maximum degree d. Then, for all $k \in \mathbb{N}$, there exists a set of at most $\alpha n r^{3/2}/k^{1/2}$ vertices whose removal leaves G will all connected components of size at most k.*

It is convenient to think of the Markov chain on G in terms of a multigraph on G, with $2d$ edges from each vertex. Each edge has probability exactly $1/2d$, and self-loops consist of many such edges. Note that every edge of the original graph is a single edge in this multigraph. For any subset of vertices $C \subseteq V$, let us define the random walk restricted to C. We remove every cut edge (u,v) (where $u \in C$ and $v \notin C$) and add a self-loop of the same probability at u. This produces a Markov chain on C that is symmetric. Given a subset C and $v \in C$, we use $\mathbf{p'}_{v,t}$ to denote the distribution of endpoints of t-length random walk starting from v and restricted to C. (In our use, C will apparent from context, so we will not carry the dependence on C in the notation.)

The following claim relates the clipped norms of the $\mathbf{p}_{v,t}$ and $\mathbf{p'}_{v,t}$ vectors.

Claim 3.4. *Let $C \subset V$ and $v \in C$. Let η be the probability that a t-length random walk from v (in G) leaves C. For any $\sigma > \eta$, $\|\mathrm{cl}(\mathbf{p}_{v,t},\sigma-\eta)\|_2^2 \ge \|\mathrm{cl}(\mathbf{p'}_{v,t},\sigma)\|_2^2$.*

Proof. The random walk restricted to C is obtained by adding some self-loops that are not in the original Markov chain. Color all these self-loops red. Let $\mathbf{r}_{v,t}(u)$ be the probability of a t-length walk from v to u that contains a red edge. Any path without a red edge is a path in G (with the same probability), so $\mathbf{p'}_{v,t}(u) \le \mathbf{p}_{v,t}(u) + \mathbf{r}_{v,t}(u)$.

Note that $\sum_{u\in C} \mathbf{r}_{v,t}(u)$ is the total probability of a random walk from u restricted to C encountering a red self-loop. Red self-loops correspond to cut edges in the original graph, and thus, this is the probability of encountering a cut edge. Hence, $\sum_{u\in C} \mathbf{r}_{v,t}(u) \le \eta$.

Intuitively, we can obtain a σ-clipping of $\mathbf{p'}_{v,t}$ by first clipping at most η probability mass to get $\mathbf{p}_{v,t}$, and then performing a $(\sigma-\eta)$-clipping of $\mathbf{p}_{v,t}$. We formalize this below.

Let $\boldsymbol{q} = \mathrm{cl}(\mathbf{p}_{v,t}, \sigma-\eta)$, and let us define the $|C|$-dimensional vector $\boldsymbol{w}$ by $\boldsymbol{w}(u) = \min\big(\boldsymbol{q}(u), \mathbf{p'}_{v,t}(u)\big)$. Since $\boldsymbol{w}$ is non-negative and $\boldsymbol{w}(u) \leq \boldsymbol{q}(u)$ for all $u \in C$, it follows that $\|\boldsymbol{w}\|_2^2 \leq \|\boldsymbol{q}\|_2^2 = \|\mathrm{cl}(\mathbf{p}_{v,t}, \sigma - \eta)\|_2^2$. By construction, for all $u \in C$, $\boldsymbol{w}(u) \leq \mathbf{p'}_{v,t}(u)$. We will prove that $\|\boldsymbol{w} - \mathbf{p'}_{v,t}\|_1 \leq \sigma$, implying that $\|\mathrm{cl}(\mathbf{p'}_{v,t}, \sigma)\|_2^2 \leq \|\boldsymbol{w}\|_2^2$. This will complete the argument.

Let $D \subseteq C$ be the set of coordinates such that $\boldsymbol{q}(u) < \mathbf{p'}_{v,t}(u)$. Since $\boldsymbol{w}(u) = \min(\boldsymbol{q}(u), \mathbf{p'}_{v,t}(u))$, $\|\mathbf{p'}_{v,t} - \boldsymbol{w}\|_1 = \sum_{u \in D}[\mathbf{p'}_{v,t}(u) - \boldsymbol{q}(u)]$. Combining with the previous observations and noting that $\boldsymbol{q} = \mathrm{cl}(\mathbf{p}_{v,t}, \sigma - \eta)$,

$$\|\mathbf{p'}_{v,t} - \boldsymbol{w}\|_1 \leq \sum_{u \in D}[\mathbf{p}_{v,t}(u) + \mathbf{r}_{v,t}(u) - \boldsymbol{q}(u)] \leq \|\mathbf{p}_{v,t}(u) - \boldsymbol{q}\|_1 + \sum_{u \in C} \mathbf{r}_{v,t}(u) \leq (\sigma - \eta) + \eta = \sigma \quad (1)$$

□

We now prove the main lemma of this section.

Proof of Lemma 3.2. Fix some $\ell \in \mathbb{N}$, $\ell > \alpha r^3$ and use Lemma 3.3 with $k = r^3\ell^6$. There exists a subset R of at most $\alpha dn/\ell^3$ edges whose removal breaks up G into connected components of size at most $r^3\ell^6$. Refer to these as AST components. Now, consider an ℓ-length walk in G starting from the stationary distribution (which is uniform). The probability that this walk encounters an edge in R at any step is exactly $|R|/2dn$. Let the random variable X_v be the number of edges of R encountered in an ℓ-length walk from v. Note that when $X_v = 0$, then the walk remains in the AST component containing v. Thus,

$$(1/n)\sum_v \Pr[\text{walk from } v \text{ leaves AST component}] \leq \mathbf{E}_{v \sim \text{u.a.r.}}[X_v] = \ell|R|/2dn \leq \alpha/(2\ell^2)$$

Since $\ell > \alpha r^3 > 4\alpha$, we can upper bound by $1/8\ell$. By the Markov bound, for at least $(1 - 1/\ell)n$ vertices, the probability that an ℓ-length walk starting at v encounters an edge of R and thus leaves the AST piece containing v is at most $1/8$. Denote the set of these vertices by S.

Consider any $s \in S$. Suppose it is contained in the AST component C. Note that $\|\mathrm{cl}(\mathbf{p'}_{s,\ell}, 1/2)\|_1 \geq 1/2$. Furthermore, it has support at most $|C| \leq r^3\ell^6$. By Jensen's inequality, $\|\mathrm{cl}(\mathbf{p'}_{s,\ell}, 1/2)\|_2^2 \geq 1(4r^3\ell^6)$. As argued earlier, the probability that a random walk (in G) from s leaves C is at most $1/8$. Applying Claim 3.4 for $\sigma = 1/2$ and $\eta = 1/8$, we conclude that $\|\mathrm{cl}(\mathbf{p}_{s,\ell}, 1/2 - 1/8)\|_2^2 \geq 1/(4r^3\ell^6) \geq 1/\ell^7$. (For convenience, we assume that $\alpha > 4$.) □

4 The existence of a discoverable decomposition

If many vertices have large clipped norms, we prove that G can be partitioned into small low conductance cuts. Furthermore, each cut can be discovered by poly(ℓ) ℓ-length random walks. The analysis follows the structure given in [KSS18a].

Lemma 4.1. *Let $c > 1$ be a parameter. Suppose there exists $S \subseteq V$ such that $|S| > n/\ell^{1/5}$ and $\forall s \in S, \|\mathrm{cl}(\mathbf{p}_{s,\ell}, 1/4)\|_2^2 > \ell^{-c}$. Then, there exists $\widetilde{S} \subseteq S$ with $|\widetilde{S}| \geq |S|/4$ such that for each $s \in \widetilde{S}$, there exists a subset $P_s \subseteq S$ where*

- $\forall v \in P_s$, $\sum_{t < 16\ell^{c+1}} p_{s,t}(v) \geq 1/8\ell^{c+1}$.

- $|E(P_s, S \setminus P_s)| \leq 4d|P_s|\sqrt{c\ell^{-1/5}\log \ell}$.

A straightforward application of this lemma leads to the main decomposition theorem.

Theorem 4.2. *Suppose there are at least $(1 - 1/\ell^{1/5})n$ vertices s such that $\|\mathrm{cl}(\mathbf{p}_{s,\ell}, 1/4)\|_2^2 > \ell^{-c}$. Then, there is a partition $\{P_1, P_2, \ldots, P_b\}$ of the vertices such that:*

- *For each P_i, there exists $s \in V$ such that: $\forall v \in P_i$, $\sum_{t<10\ell^{c+1}} p_{s,t}(v) \geq 1/8\ell^{c+1}$.*
- *The total number of edges crossing the partition is at most $8dn\sqrt{c\ell^{-1/5}\log \ell}$.*

Proof. We simply iterate over Lemma 4.1. Let $T = \{s \mid \|\mathrm{cl}(\mathbf{p}_{s,\ell}, 1/4)\|_2^2 \leq \ell^{-c}\}$. By assumption, $|T| \leq n/\ell^{1/5}$. We will maintain a partition of the vertices $\{T, Q_1, Q_2, \ldots, Q_a, S\}$ with the following properties. (1) Each Q_i satisfies the first condition of the theorem. (2) The total number of edges crossing the partition is at most $4d\sqrt{c\ell^{-1/5}\log \ell}\sum_{i\leq a}|Q_i| + d|T|$. We initialize with the trivial partition $\{T, S = V \setminus T\}$.

As long as $|S| > n/\ell^{1/5}$, we invoke Lemma 4.1. We get a new set $Q \subseteq S$ satisfying the first condition of the theorem, and the number of edges from Q to $S \setminus Q$ is at most $4d\sqrt{c\ell^{1/5}\log \ell}|Q|$. We add Q to our partition, reset $S = S \setminus Q$, and iterate.

When this process terminates, $|S| \leq n/\ell^{1/5}$. We get the final partition by removing all edges incident to $S \cup T$. Alternately, every single vertex in $S \cup T$ becomes a separate set. Note that a single vertex trivially satisfies the first condition of theorem, since for all s, $p_{s,s}(1) \geq 1/2$. The total number of edges crossing the partition is at most $4dn\sqrt{c\ell^{-1/5}\log \ell} + 2dn\ell^{-1/5} \leq 8dn\sqrt{c\ell^{-1/5}\log \ell}$. □

4.1 Proving Lemma 4.1

An important tool used to argue about conductances within S is the projected Markov chain. These ideas come from the work of Kale-Peres-Seshadhri to analyze random walks in noisy expanders [KPS13], and were used by the authors in their previous paper on one-sided testers for minor-freeness [KSS18a]. We closely follow the structure and notation of that paper, and explicitly mention the differences.

We define the "projection" of the random walk onto the set S. We define a Markov chain M_S, over the set S. We retain all transitions from the original random walk on G that are within S, and we denote these by $e^{(1)}_{u,v}$ for every u to v transition in the random walk on G. Additionally, for every $u, v \in S$ and $t \geq 2$, we add a transition $e^{(t)}_{u,v}$. The probability of this transition is equal to the total probability of t-length walks in G from u to v, where all internal vertices in the walk lie outside S.

Note that $e^{(t)}_{u,v} = e^{(t)}_{v,u}$. Since G is irreducible and the stationary mass on S is nonzero, all walks eventually reach S. Thus, for any u, $\sum_t \sum_v e^{(t)}_{u,v} = 1$, so M_S is a symmetric Markov chain. The stationary distribution of M_S is uniform on S.

For a transition $e^{(t)}_{u,v}$ in M_S, define the "length" of this transition to be t. For clarity, we use "hops" to denote the number of steps of a walk in M_S, and retain "length" for walks in G. The length of an h hop random walk in M_S is defined to be the sum of the lengths of the transitions it takes.

We use $\boldsymbol{\tau}_{s,h}$ to denote the distribution of the h-hop walk from s, and $\tau_{s,h}(v)$ to denote the corresponding probability of reaching v. We use $\mathcal{W}_h$ to denote the distribution of h-hop walks starting from the uniform distribution.

The following lemma is crucial for relating walks in G with M_S.

Lemma 4.3 (Lemma 6.4 of [KSS18b]). $\mathbf{E}_{W\sim\mathcal{W}_h}[\textit{length of } W] = hn/|S|$

We come to an important lemma. The conditions in Lemma 4.1 are on the clipped norms of random walks in G, but the conclusion (regarding the cut) refers to conductances within the projected Markov chain M_S. The following lemma shows that random walks in M_S must also be sufficiently trapped. This is an analogue of Lemma 6.5 of [KSS18b], but the proof deviates significantly because of the use of clipped norms.

Lemma 4.4. *There exists a subset $S' \subseteq S$, $|S'| \geq |S|/2$, such that $\forall s \in S'$, $\|\boldsymbol{\tau}_{s,\ell^{1/5}}\|_\infty \geq 1/2\ell^{c+1}$.*

Proof. Consider ℓ-length random walks in G starting from $s \in S$. For any such walk, we can define the number of hops it makes as the number of vertices in S encountered minus one.

For $h \in \mathbb{N}$ and $s \in S$, define the event $\mathcal{E}_{s,h}$ that an ℓ-length walk from s makes h hops. We will further split this event into $\mathcal{F}_{s,h}$, when the walk ends at S, and $\mathcal{G}_{s,h}$, when the walk does not end at S. A walk that ends in S directly corresponds to an h-hop walk in M_S. By Lemma 4.3, $|S|^{-1}\sum_{s\in S}\Pr[\mathcal{F}_{s,h}]\ell \leq hn/|S|$. Consider any walk in the event $\mathcal{G}_{s,h}$. If one continued until it ends in S, this gives a walk in M_S with a single additional hop (and a longer length). Thus, the total probability mass $\Pr[\mathcal{G}_{s,h}]$ corresponds to walks in M_S that make $(h+1)$ hops and have length at least ℓ. By Lemma 4.3 again, $|S|^{-1}\sum_{s\in S}\Pr[\mathcal{G}_{s,h}]\ell \leq (h+1)n/|S|$.

Summing these bounds and applying the size bound on S,

$$|S|^{-1}\sum_{s\in S}\Pr[\mathcal{E}_{s,h}]\ell \leq (2h+1)n/|S| \leq \ell^{1/5}(2h+1) \Longrightarrow |S|^{-1}\sum_{s\in S}\Pr[\mathcal{E}_{s,h}] \leq \ell^{-4/5}(2h+1)$$

Now, we sum over h and use the fact that ℓ is a sufficiently large constant.

$$|S|^{-1}\sum_{h\leq\ell^{1/5}}\sum_{s\in S}\Pr[\mathcal{E}_{s,h}] \leq \ell^{-4/5}\sum_{h\leq\ell^{1/5}}(2h+1) \leq 4\ell^{-2/5} < 1/10$$

By the Markov bound, there is a set S', $|S'| \geq |S|/2$ such that $\forall s \in S', \sum_{h\leq\ell^{1/5}}\Pr[\mathcal{E}_{s,h}] < 1/5$.

For $v \in V$, let $y_s(v)$ be the probability that an ℓ-length walk from s to v makes at most $\ell^{1/5}$ hops. Note that $\sum_{v\in S} y_s(v) \leq \sum_{h\leq\ell^{1/5}}\Pr[\mathcal{E}_{s,h}] < 1/5$. We now use the clipped norm definition. Since $\|\mathrm{cl}(\mathbf{p}_{s,\ell}, 1/4)\|_2^2 \geq \ell^{-c}$, $\sum_{v\in V}(p_{s,\ell}(v) - y_s(v))^2 \geq \ell^{-c}$. This is important, since we can "remove" the low hop walks and still have a large norm.

Consider the probability α that a 2ℓ-length walk from s back to s makes at least $\ell^{1/5}$ hops. (Note that this corresponds to walks in M_S.) Clearly, any walk going from s to v in an ℓ-length walk making at least $\ell^{1/5}$ hops and then returning to s in an ℓ-length walk contributes to this probability. Thus, we can lower bound α by $\sum_{v\in V}(p_{s,\ell}(v) - y_s(v))^2 \geq \ell^{-c}$. Note that all walks considered make at most 2ℓ hops.

Thus, $\sum_{h\geq\ell^{1/5}}^{2\ell}\|\boldsymbol{\tau}_{s,\ell^{1/5}}\|_\infty \geq \ell^{-c}$. Since the infinity norm is non-increasing in hops, by averaging, $\|\boldsymbol{\tau}_{s,\ell^{1/5}}\|_\infty \geq 1/2\ell^{c+1}$. □

The remaining proof of Lemma 4.1 is almost identical to analogous calculations in Section 6 of [KSS18b]. Therefore, we move it to the appendix.

5 Proof of main result

Before we show Theorem 2.1 and Theorem 2.2, we argue about the guarantees of `EstClip`. The proofs of the next two claims are relatively routine concentration arguments. Recall that T is the vertex set constructed in a call to $\texttt{EstClip}(s)$.

Claim 5.1. *Consider any vertex s. With probability at least $1 - 2^{-1/\varepsilon^2}$ over the randomness in $\texttt{EstClip}(s)$: all v such that $\mathbf{p}_{s,\ell}(v) \geq 1/\ell^7$ are in T, and no v such that $\mathbf{p}_{s,\ell}(v) \leq 1/\ell^8$ is in T.*

Proof. Consider v such that $\mathbf{p}_{s,\ell}(v) \geq 1/\ell^7$. Recall that the total number of walks is $w = \ell^{14}$. The expected value of w_v is at least $\ell^{14}/\ell^7 = \ell^7$. Note that w_v is a sum of Bernoulli random variables. By a multiplicative Chernoff bound (Theorem 1.1 of [DP09]), $\Pr[w_v \leq \ell^7/2] \leq \exp(-\ell^7/8)$. There are at most ℓ^7 such vertices v. By a union bound over all of them, the probability that some such v is not in T is at most $\ell^7 \cdot \exp(-\ell^7/8) \leq \exp(-\ell^6) \leq 2^{-2/\varepsilon^2}$. (Note that $\ell > \varepsilon^{-20}$.) This proves the first part.

For the second part, consider v such that $\mathbf{p}_{s,\ell}(v) \leq 1/\ell^8$. We split into two cases.

Case 1, $\mathbf{p}_{s,\ell}(v) \geq \exp(-\ell/2)$. The expectation of w_v is at most $\ell^{14}/\ell^8 = \ell^6$. Since $\ell^7/2 \geq 2e\ell^6$, by a Chernoff bound (third part, Theorem 1.1 of [DP09]), $\Pr[w_v \geq \ell^7/2] \leq 2^{-\ell^7/2}$. There are at most $\exp(\ell/2)$ such vertices v. Taking a union bound over all of them, the probability that any such vertex appears in T is at most $\exp(\ell/2)2^{-\ell^7/2} \leq 2^{-\ell^5} \leq 2^{-2/\varepsilon^2}$.

Case 2, $\mathbf{p}_{s,\ell}(v) < \exp(-\ell/2)$. For convenience, set $p = \mathbf{p}_{s,\ell}(v)$. The probability that $w_v \leq 1$ is:

$$(1-p)^w + wp(1-p)^{w-1} \geq (1 - wp) + wp(1 - p(w-1)) = 1 - p^2 w(w-1) \geq 1 - p^2 w^2 \tag{2}$$

(We use the inequality $(1-x)^r \geq 1 - xr$, for $|x| \leq 1, r \in \mathbb{N}$.) Thus, the probability that $w_v > 1$ is at most $p^2 w^2$. Note that $\ell^7/2$ (the threshold to be placed in T) is at least 2.

Let us take a union bound over all such vertices. We note that $w = \ell^{14}$ and $\ell > \varepsilon^{-20}$. The probability that any such v is placed in T is at most

$$\sum_{v:\mathbf{p}_{s,\ell}(v)<\exp(-\ell/2)} \mathbf{p}_{s,\ell}(v)^2 w^2 \leq \ell^{28}\exp(-\ell/2)\sum_v \mathbf{p}_{s,\ell}(v) \leq \exp(-1/\varepsilon^2) \tag{3}$$

We union bound over all errors to complete the proof. □

We can now argue about the main guarantee of `EstClip`.

Claim 5.2. *For all vertices s, with probability at least $1-2^{-1/\varepsilon}$ over the randomness of $\texttt{EstClip}(s)$:*

- *If $\|\mathrm{cl}(\mathbf{p}_{s,\ell}, 1/4)\|_2^2 < \ell^{-8}/400$, then $\texttt{EstClip}(s)$ outputs LOW.*
- *If $\|\mathrm{cl}(\mathbf{p}_{s,\ell}, 3/8)\|_2^2 > \ell^{-7}$, then $\texttt{EstClip}(s)$ outputs HIGH.*

Proof. Consider the first case. Let $H = \{v \mid \mathbf{p}_{s,\ell}(v) \geq \ell^{-8}\}$. We first argue that $\sum_{v\in H}\mathbf{p}_{s,\ell}(v) \leq 1/4 + 1/20$. Suppose not. Then, any clipping of $1/4$ of the probability mass of $\mathbf{p}_{s,\ell}$ leaves at least $1/20$ probability mass on H. The size of H is at most ℓ^8. By Jensen's inequality, $\|\mathrm{cl}(\mathbf{p}_{s,\ell}, 1/4)\|_2^2 \geq 1/400\ell^8$, contradicting the case condition.

Thus, $\sum_{v\in H}\mathbf{p}_{s,\ell}(v) \leq 1/4 + 1/20$. The expected value of $\sum_{v\in H} w_v \leq w(1/4+1/20)$. By an additive Chernoff bound (first part, Theorem 1.1 of [DP09]), $\Pr[\sum_{v\in H} w_v \geq w/3] \leq \exp(-2(1/3 - 1/4 - 1/20)^2 w) \leq \exp(-\ell^{13})$. By Claim 5.1, with probability at least $1 - 2^{-1/\varepsilon^2}$, $T \subseteq H$. By a union bound, with probability at least $1 - 2^{-1/\varepsilon}$, $\sum_{v \in T} w_v \leq \sum_{v \in H} w_v < w/3$, and the output is LOW.

Now for the second case. Let $H' = \{v \mid \mathbf{p}_{s,\ell}(v) \geq \ell^{-7}\}$. We will show that $\sum_{v \in H} \mathbf{p}_{s,\ell}(v) \geq 3/8$. Suppose not. We can clip away all the probability mass of $\mathbf{p}_{s,\ell}$ that is on H, which is at most $3/8$. All remaining probability/entries of the clipped vector are at most ℓ^{-7}. Thus, the squared l_2-norm is at most ℓ^{-7}, implying $\|\mathrm{cl}(\mathbf{p}_{s,\ell}, 3/8)\|_2^2 \leq \ell^{-7}$ (contradiction).

Thus, $\sum_{v \in H'} \mathbf{p}_{s,\ell}(v) \geq 3/8$. By an additive Chernoff bound (first part, Theorem 1.1 of [DP09]), $\Pr[\sum_{v \in H} w_v < w/3] \leq \exp(-2(3/8 - 1/3)^2 w) \leq \exp(-\ell^{13})$. By Claim 5.1, with probability at least $1 - 2^{-1/\varepsilon^2}$, $H' \subseteq T$. By a union bound, with probability at least $1 - 2^{-1/\varepsilon}$, $\sum_{v \in T} w_v \geq \sum_{v \in H'} w_v \geq w/3$, and the output is HIGH. □

We now prove completeness, Theorem 2.1. We will prove that if G is H-minor-free, then the tester `IsMinorFree` accepts with probability $> 2/3$. This follows almost directly from Lemma 3.2.

Proof of Theorem 2.1. Suppose G is H-minor-free. Note that calls to `LocalSearch` can never return FOUND, so rejection can only happen because of the output of calls to `EstClip`.

By Lemma 3.2, there are at least $(1 - 1/\ell)n$ vertices such that $\|\mathrm{cl}(\mathbf{p}_{s,\ell}, 3/8)\|_2^2 \geq \ell^{-7}$. Call these vertices *heavy*. The expected number of light vertices in the multiset S chosen in Step 1 of `IsMinorFree` is at most $1/\ell \times \ell^{21} = \ell^{20}$. By a multiplicative Chernoff bound (Theorem 1 of [DP09]), the number of light vertices in S is strictly less than $2\ell^{20}$ with probability at least $1 - \exp(-\ell^{19}) > 9/10$. Let us condition on this event. The probability that any call to `EstClip`(s) returns HIGH for a heavy $s \in S$ is at least $1 - 2^{-1/\varepsilon}$, by Claim 5.2. By a union bound over the at most ℓ^{21} heavy vertices in S, all calls to `EstClip`(s) for heavy $s \in S$ return HIGH with probability at least $1 - \ell^{21} 2^{-1/\varepsilon} > 9/10$.

We now remove the conditioning. With probability $> (9/10)^2 > 2/3$, there are strictly less than $2\ell^{18}$ calls (for the light vertices) that return LOW. When this happens, `IsMinorFree` accepts. □

Now we prove soundness, Theorem 2.2. We prove that if G is ε-far from H-minor-freeness, the tester rejects with probability $> 2/3$. The main ingredient is the decomposition of Theorem 4.2.

Proof of Theorem 2.2. Assume G is ε-far from being H-minor free. We split into two cases.

Case 1: There are less than $(1 - 1/\ell^{1/5})n$ vertices such that $\|\mathrm{cl}(\mathbf{p}_{s,\ell}, 1/4)\|_2^2 > \ell^{-9}$.

Then, there are at least $n/\ell^{1/5}$ vertices such that $\|\mathrm{cl}(\mathbf{p}_{s,\ell}, 1/4)\|_2^2 \leq \ell^{-9}$. The expected number of such vertices (with repetition) in the multiset S (of Step 1) is at least $\ell^{21}/\ell^{1/5}$. By a multiplicative Chernoff bound, there are at least $\ell^{21}/2\ell^{1/5} > 2\ell^{20}$ such vertices in S, with probability at least $1 - \exp(-\ell^{20}/4)$. For each such vertex s, the probability that `EstClip`(s) outputs LOW is at least $1 - 2^{-1/\varepsilon}$ (Claim 5.2). By a union bound over all vertices in S, with probability $> (1 - \exp(-\ell^{20}))(1 - \ell^{21} 2^{-1/\varepsilon}) > 5/6$, there are at least $2\ell^{20}$ calls to `EstClip`(s) that return LOW. So the tester rejects.

Case 2: There are at least $(1 - 1/\ell^{1/5})n$ vertices such that $\|\mathrm{cl}(\mathbf{p}_{s,\ell}, 1/4)\|_2^2 > \ell^{-9}$.

We apply the decomposition of Theorem 4.2 (with $c = 9$). There is a partition $\{P_1, P_2, \ldots, P_b\}$ of the vertices such that:

- For each P_i, there exists $s \in V$ such that: $\forall v \in P_i$, $\sum_{t < 10\ell^{10}} p_{s,t}(v) \geq 1/8\ell^{10}$. Call s the *anchor* for P_i, noting that multiple sets may have the same anchor.
- The total number of edges crossing the partition is at most $24dn\sqrt{\ell^{-1/5} \log \ell}$.

Among the sets in the partition, let $\{Q_1, Q_2, \dots, Q_a\}$ be the sets of vertices that contain an H-minor (or technically, the subgraphs induced by these sets contain an H-minor). Note that one can remove $d\sum_{i\leq a}|Q_i| + 24dn\sqrt{\ell^{-1/5}\log\ell}$ edges to make G H-minor-free. Since $\ell > \varepsilon^{-20}$, $24dn\sqrt{\ell^{-1/5}\log\ell} \leq \varepsilon nd/2$. Since G is ε-far from being H-minor free, we deduce from the above that $\sum_{i\leq a}|Q_i| \geq \varepsilon n/2$.

Let $Z = \{s \mid s \text{ is anchor for some } Q_i\}$. Let us lower bound $|Z|$. For every Q_i, there is some $s \in Z$ such that $\forall v \in Q_i$, $\sum_{t<10\ell^{10}} p_{s,t}(v) \geq 1/8\ell^{10}$. Thus, for every Q_i, there is some $s \in Z$ such that $\sum_{v\in Q_i}\sum_{t<10\ell^{10}} p_{s,t}(v) \geq |Q_i|/8\ell^{10}$. Let us sum over all $s \in Z$ (and note that $\sum_{v\in V} p_{s,t}(v) = 1$).

$$\sum_{i\leq a}|Q_i|/8\ell^{10} \leq \sum_{s\in Z}\sum_{v\in V}\sum_{t<10\ell^{10}} p_{s,t}(v) \leq \sum_{t<10\ell^{10}}\sum_{s\in Z}\sum_{v\in V} p_{s,t}(v) \leq 10\ell^{10}|Z| \tag{4}$$

Since $\sum_{i\leq a}|Q_i| \geq \varepsilon n/2$, $|Z| \geq \varepsilon n/160\ell^{20} \geq 5n/\ell^{21}$.

Focus on the multiset S in Step 1 of `IsMinorFree` . Note that S contains an element of Z with probability $\geq 1 - (1 - 5/\ell^{21})^{\ell^{21}} \geq 9/10$. Let us condition of this event, and let $s \in S \cap Z$. There exists some Q_i such that $\forall v \in Q_i$, $\sum_{t<10\ell^{10}} p_{s,t}(v) \geq 1/8\ell^{10}$. By averaging over walk length, $\forall v \in Q_i$, $\exists t < 10\ell^{10}$ such that $p_{s,t}(v) \geq 1/80\ell^{20}$.

Now, consider the call to `LocalSearch`(s). The set B_s in Step 1 of `LocalSearch` is constructed by performing ℓ^{21} random walks of length ℓ^{11}. For any $v \in Q_i$, the probability that v is in B_s is at least $1 - (1 - 1/80\ell^{20})^{\ell^{21}} \geq 1 - \exp(-\ell/80)$. Taking a union bound over all $v \in Q_i$, the probability that $Q_i \subseteq B_s$ is at least $1 - \ell^{21}\exp(-\ell/80) \geq 9/10$. When $Q_i \subseteq B_s$, then $G[B_s]$ contains an H-minor and the tester rejects. The probability of this happening is at least $(9/10)^2 > 2/3$. □

References

[AST90] Noga Alon, Paul Seymour, and Robin Thomas. A separator theorem for nonplanar graphs. *Journal of the American Mathematical Society*, 3(4):801–808, 1990. 6

[BSS08] I. Benjamini, O. Schramm, and A. Shapira. Every minor-closed property of sparse graphs is testable. In *Symposium on the Theory of Computing (STOC)*, pages 393–402, 2008. 1, 2

[CGR+14] Artur Czumaj, Oded Goldreich, Dana Ron, C Seshadhri, Asaf Shapira, and Christian Sohler. Finding cycles and trees in sublinear time. *Random Structures & Algorithms*, 45(2):139–184, 2014. 2

[Die10] Reinhard Diestel. *Graph Theory, Fourth Edition*. Springer, 2010. 2

[DP09] D. P. Dubhashi and A. Panconesi. *Concentration of measure for the analysis of randomized algorithms*. Cambridge, 2009. 10, 11

[EHNO11] Alan Edelman, Avinatan Hassidim, Huy N. Nguyen, and Krzysztof Onak. An efficient partitioning oracle for bounded-treewidth graphs. In *Workshop on Randomization and Computation (RANDOM)*, pages 530–541, 2011. 1, 2

[FLVW17] Hendrik Fichtenberger, Reut Levi, Yadu Vasudev, and Maximilian Wötzel. On testing minor-freeness in bounded degree graphs with one-sided error. *CoRR*, abs/1707.06126, 2017. 2

[Gol17] O. Goldreich. *Introduction to Property Testing*. Cambridge University Press, 2017. 1, 2

[GR99] O. Goldreich and D. Ron. A sublinear bipartite tester for bounded degree graphs. *Combinatorica*, 19(3):335–373, 1999. 2

[GR02] O. Goldreich and D. Ron. Property testing in bounded degree graphs. *Algorithmica*, 32(2):302–343, 2002. 1

[HKNO09] A. Hassidim, J. Kelner, H. Nguyen, and K. Onak. Local graph partitions for approximation and testing. In *Foundations of Computer Science (FOCS)*, pages 22–31, 2009. 1, 2

[HT74] John Hopcroft and Robert Tarjan. Efficient planarity testing. *Journal of the ACM (JACM)*, 21(4):549–568, 1974. 1

[KKR12] Ken-ichi Kawarabayashi, Yusuke Kobayashi, and Bruce A. Reed. The disjoint paths problem in quadratic time. *J. Comb. Theory, Ser. B*, 102(2):424–435, 2012. 1, 5

[KPS13] Satyen Kale, Yuval Peres, and C. Seshadhri. Noise tolerance of expanders and sublinear expansion reconstruction. *SIAM J. Comput.*, 42(1):305–323, 2013. 8

[KSS18a] Akash Kumar, C. Seshadhri, and Andrew Stolman. Finding forbidden minors in sublinear time: A $\mathrm{o}(\mathrm{n}^{1/2 + \mathrm{o}(1)})$-query one-sided tester for minor closed properties on bounded degree graphs. In *Foundations of Computer Science (FOCS)*, pages 509–520, 2018. 2, 7, 8

[KSS18b] Akash Kumar, C. Seshadhri, and Andrew Stolman. Finding forbidden minors in sublinear time: a $\mathrm{o}(\mathrm{n}^{1/2 + \mathrm{o}(1)})$-query one-sided tester for minor closed properties on bounded degree graphs. *CoRR*, abs/1805.08187, 2018. 9, 14

[Kur30] K. Kuratowski. Sur le problème des courbes gauches en topologie. *Fundamenta Mathematica*, 15:271–283, 1930. 1

[LR15] Reut Levi and Dana Ron. A quasi-polynomial time partition oracle for graphs with an excluded minor. *ACM Transactions on Algorithms (TALG)*, 11(3):24, 2015. 1, 2

[LS90] László Lovász and Miklós Simonovits. The mixing rate of markov chains, an isoperimetric inequality, and computing the volume. In *Foundations of Computer Science (FOCS)*, pages 346–354, 1990. 14

[NS13] Ilan Newman and Christian Sohler. Every property of hyperfinite graphs is testable. *SIAM Journal on Computing*, 42(3):1095–1112, 2013. 2

[RS95a] N. Robertson and P. D. Seymour. Graph minors. XII. Distance on a surface. *Journal of Combinatorial Theory Series B*, 64(2):240–272, 1995. 1

[RS95b] N. Robertson and P. D. Seymour. Graph minors. XIII. The disjoint paths problem. *Journal of Combinatorial Theory Series B*, 63(1):65–110, 1995. 1

[RS04] N. Robertson and P. D. Seymour. Graph minors. XX. Wagner's conjecture. *Journal of Combinatorial Theory Series B*, 92(1):325–357, 2004. 1

[Spi] D. Spielman. Lecture notes on spectral graph theory. http://www.cs.yale.edu/homes/spielman/eigs/. 14

[ST12] D. Spielman and S.-H. Teng. A local clustering algorithm for massive graphs and its application to nearly-linear time graph partitioning. *SIAM Journal on Computing*, 42(1):1–26, 2012. 3, 14

[Wag37] K. Wagner. Über eine eigenschaft der ebenen komplexe. *Mathematische Annalen*, 114:570–590, 1937. 1

[YI15] Yuichi Yoshida and Hiro Ito. Testing outerplanarity of bounded degree graphs. *Algorithmica*, 73(1):1–20, 2015. 1, 2

A Local partitioning, and completing the proof of Lemma 4.1

We perform local partitioning on M_S, starting with an arbitrary $s \in S'$. We apply the Lovász-Simonovits curve technique. (The definitions are originally from [LS90]. Refer to Lecture 7 of Spielman's notes [Spi] as well as Section 2 in Spielman-Teng [ST12]. This is also a restatement of material in Section 6.1 of [KSS18b], which is needed to state the main lemma.)

- Conductance: for some $T \subseteq S$ we define the conductance of T in M_S to be
$$\Phi(T) = \frac{\sum_{\substack{u \in T \\ v \in S \setminus T}} \tau_{u,1}(v)}{\min\{|S \setminus T|, |T|\}}$$
- Ordering of states at time t: At time t, let us order the vertices in M_S as $v_1^{(t)}, v_2^{(t)}, \ldots$ such that $\tau_{s,t}(v_1^{(t)}) \geq \tau_{s,t}(v_2^{(t)}) \ldots$, breaking ties by vertex id. At $t = 0$, we set $\tau_{s,0}(s) = 1$, and all other values to 0.
- The LS curve h_t: We define a function $h_t : [0, |S|] \to [0, 1]$ as follows. For every $k \in [|S|]$, set $h_t(k) = \sum_{j \leq k} \tau_{s,t}(v_j^{(t)})$. (Set $h_t(0) = 0$.) For every $x \in (k, k+1)$, we linearly interpolate to construct $h(x)$. Alternately, $h_t(x) = \max_{\vec{w} \in [0,1]^{|S|}, \|\vec{w}\|_1 = x} \sum_{v \in S} [\tau_{s,t}(v) - 1/n] w_i$.
- Level sets: For $k \in [0, |S|]$, we define the (k,t)-level set, $L_{k,t}$ to be $\{v_1^{(t)}, v_2^{(t)}, \ldots, v_k^{(t)}\}$. The *minimum probability* of $L_{k,t}$ denotes $\tau_{s,t}(v_k^{(t)})$.

The main lemma of Lovász-Simonovits is the following (Lemma 1.4 of [LS90], also refer to Theorem 7.3.3 of Lecture 7 in [Spi]).

Lemma A.1. *For all k and all t,*
$$h_t(k) \leq \frac{1}{2}[h_{t-1}(k - 2\min(k, n-k)\Phi(L_{k,t})) + h_{t-1}(k + 2\min(k, n-k)\Phi(L_{k,t}))]$$

We employ this lemma to prove a condition of the level set conductances. An analogous lemma was proven in [KSS18b] for specific parameters. We redo the calculation here.

Lemma A.2. *Suppose there exists $\phi \in [0,1]$ and $p > 2/n$ such that for all $t' \leq t$ it is true that for all $k \in [n]$ that if $L_{k,t'}$ has a minimum probability of at least p, then $\Phi(L_{k,t}) \geq \phi$. Then for all $k \in [0,n]$, $h_t(k) \leq \sqrt{k}(1-\phi^2/2)^t + pk$.*

Proof. We will prove by induction over t. For the base case, consider $t = 0$. The RHS is at least 1, proving the bound.

Now for the induction. Note that h_t is a concave, and the RHS is also concave. Thus, it suffices to prove the bound for the integer points ($h_t(k)$ for integer k). If $k \geq 1/p$, then the RHS is at least 1. Thus the bound is trivially true. Let us assume that $k < 1/p < n/2$. We now split the proof into two cases based on the conductance of $L_{k,t}$.

First let us consider the case where $\Phi(L_{k,t}) \geq \phi$. By Lemma A.1 and concavity of h,

$$\begin{aligned} h_t(k) &\leq \frac{1}{2}\Big(h_{t-1}\big(k(1-2\phi)\big) + h_{t-1}\big(k(1+2\phi)\big)\Big) && (5)\\ &\leq \frac{1}{2}\Big(\sqrt{k(1-2\phi)}(1-\phi^2/2)^{t-1} + \sqrt{k(1+2\phi)}(1-\phi^2/2)^{t-1} + 2kp\Big) && (6)\\ &\leq \frac{1}{2}\Big(\sqrt{k}\left(1-\phi^2/2\right)^{t-1}\left(\sqrt{1-2\phi}+\sqrt{1+2\phi}\right) + 2kp\Big) && (7)\\ &\leq \sqrt{k}\left(1-\phi^2/2\right)^{t} + kp && (8)\end{aligned}$$

For the last inequality we use the bound $\left(\sqrt{1+z}+\sqrt{1-z}\right)/2 \leq 1 - z^2/8$.

Now we deal with the case when $\Phi(L_{k,t}) < \phi$. By assumption, $L_{k,t}$ has minimum probability less than p. Let $k' < k$ be the largest index such that $L_{k',t}$ has minimum probability at least p. Note that $\Phi(L_{k',t}) \geq \phi$. Therefore, as proven in the first case, $h_t(k') \leq \sqrt{k'}\left(1-\phi^2/2\right)^t + k'p$. Every vertex we add to $L_{k',t}$ adds less than p probability mass to $L_{k',t}$, and therefore, by the concavity of $h_t(x)$,

$$\begin{aligned} h_t(k) &\leq h_t(k') + (k-k')p && (9)\\ &\leq \sqrt{k'}\left(1-\phi^2/2\right)^t + k'p + (k-k')p && (10)\\ &\leq \sqrt{k'}\left(1-\phi^2/2\right)^t + kp \leq \sqrt{k}\left(1-\phi^2/2\right)^t + kp && (11)\end{aligned}$$

□

For convenience, we restate Lemma 4.1.

Lemma A.3. *Let $c > 1$ be a parameter. Suppose there exists $S \subseteq V$ such that $|S| > n/\ell^{1/5}$ and $\forall s \in S, \|\mathrm{cl}(\mathbf{p}_{s,\ell}, 1/4)\|_2^2 > \ell^{-c}$. Then, there exists $\widetilde{S} \subseteq S$ with $|\widetilde{S}| \geq |S|/4$ such that for each $s \in \widetilde{S}$, there exists a subset $P_s \subseteq S$ where*

- *$\forall v \in P_s$, $\sum_{t<16\ell^{c+1}} p_{s,t}(v) \geq 1/8\ell^{c+1}$.*
- *$|E(P_s, S \setminus P_s)| \leq 4d|P_s|\sqrt{c\ell^{-1/5}\log \ell}$.*

Proof. By Lemma 4.4, there is a set $S' \subseteq S$, $|S'| \geq |S|/2$ such that for all $s \in S'$, $\|\boldsymbol{\tau}_{s,\ell^{1/5}}\|_\infty \geq 1/2\ell^{c+1}$. Consider any $s \in S'$.

Suppose for all $t' \leq \ell^{1/5}$, all level sets $L_{k,t'}$ with minimum probability $1/2\ell^{c+1}$ have conductance at least $\sqrt{4c\ell^{-1/5}\log \ell}$. Lemma A.2 implies that $\|\boldsymbol{\tau}_{s,\ell^{1/5}}\|_\infty = h_{\ell^{1/5}}(1) \leq (1 - 2c\ell^{-1/5}\log \ell)^{\ell^{1/5}} + 1/4\ell^{c+1} < 1/4\ell^{c+1} + 1/4\ell^{c+1} = 1/2\ell^{c+1}$. This is a contradiction.

Thus, for every $s \in S'$, there exists a level set denoted P_s with minimum probability $1/2\ell^{c+1}$ and conductance at most $\sqrt{4c\ell^{-1/5}\log\ell}$. Note that $|P_s| \leq 2\ell^{c+1} < |S|/2$.

$$\sqrt{4c\ell^{-1/5}\log\ell} \geq \Phi(P_s) = \frac{\sum_{\substack{x\in P_s\\ y\in S\setminus P_s}} \tau_{x,1}(y)}{\min(|P_s|, |S\setminus P_s|} \geq \frac{E(P_s, S\setminus P_s)}{2d|P_s|} \tag{12}$$

The inequality is obtained by only considering transitions from S to $S \setminus P_s$ that come from a single edge in G. Each such edge has a traversal probability of $1/2d$. Therefore, $E(P_s, S \setminus P_s) \leq 4d|P_s|\sqrt{c\ell^{-1/5}\log\ell}$.

Set $L = 8\ell^{c+2}$. Define $\widetilde{S} \subseteq S'$ to be the vertices $s \in S'$ with the property that $\forall v \in P_s$, $\sum_{l<L} p_{s,v}(l) \geq 1/8\ell^{c+1}$. Together with the cut bound above, this clearly satisfies the conditions on the lemma. It remains the prove a suitable upper bound of $|S' \setminus \widetilde{S}|$, to show that $\widetilde{S}$ is sufficiently large.

For every $s \in S' \setminus \widetilde{S}$, there exists $v_s \in P_s$ such that $\sum_{l<L} p_{s,l}(v) < 1/8\ell^{c+1}$. Let $\hat{p}_{s,l}(v)$ denote that probability that an $\ell^{1/5}$-hop walk in M_S from s reaches v with length l. Consider $s \in S' \setminus \widetilde{S}$.

$$\tau_{s,\ell^{1/5}}(v_s) = \sum_{l\geq\ell^{1/5}} \hat{p}_{s,l}(v_s) = \sum_{l\geq\ell^{1/5}}^{L-1} \hat{p}_{s,l}(v_s) + \sum_{l\geq L} \hat{p}_{s,l}(v_s) \leq \sum_{l\geq\ell^{1/5}}^{L-1} p_{s,l}(v) + \sum_{l\geq L} \hat{p}_{s,l}(v) \tag{13}$$

Since the minimum probability of P_s is at least $1/4\ell^{c+1}$, $\tau_{s,\ell^{1/5}}(v_s) \geq 1/4\ell^{c+1}$. We argued above that $\sum_{l\geq\ell^{1/5}}^{L-1} p_{s,l}(v) \leq \sum_{l<L} p_{s,l}(v) < 1/8 \leq^{c+1}$. We conclude that $\sum_{l>L} \hat{p}_{s,l}(v) \geq 1/8\ell^{c+1}$. Note that all of this probability mass corresponds to $\ell^{1/5}$-hop walks that have a large length. We now lower bound $\mathbf{E}_{W\sim\mathcal{W}_{\ell^{1/5}}}[\text{length of } W]$.

$$\mathbf{E}_{W\sim\mathcal{W}_{\ell^{1/5}}}[\text{length of } W] \geq \frac{1}{|S|} \sum_{s\in S'\setminus\widetilde{S}} \Big(\sum_{l>L} \hat{p}_{s,l}(v_s)\Big) L \geq \frac{|S'\setminus\widetilde{S}|}{|S|} \cdot \frac{L}{8\ell^{c+1}} \geq \frac{\ell|S'\setminus\widetilde{S}|}{|S|} \tag{14}$$

By Lemma 4.3, $\mathbf{E}_{W\sim\mathcal{W}_{\ell^{1/5}}}[\text{length of } W] = \ell^{1/5}n/|S|$. Combining, $|S' \setminus \widetilde{S}| \leq n/\ell^{4/5} \leq n/4\ell^{1/5} \leq |S|/4$. By Lemma 4.4, $|S'| \geq |S|/2$. By the setting of Lemma 4.1, $|S| > n/\ell^{1/5}$. Thus, $|S' \setminus \widetilde{S}| \leq n/4\ell^{1/5}$, and $|\widetilde{S}| \geq |S|/4$. □

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il