H $\le_m$ G   G contains H.   learning theory ← Rocco Servedio, Oded Goldreich, Yoshida + Bhattacharya

Sps G is a bipartite graph cont. a cycle, want to say, G contains a $K_3 \le_m G$ "subdivision" is a natural notion

TV dist → to do   Birthday paradox idea?

SIAM J. COMPUT.
Vol. 0, No. 0, pp. FOCS18-216–FOCS18-249

RANDOM WALKS AND FORBIDDEN MINORS I: AN $n^{1/2+o(1)}$-QUERY ONE-SIDED TESTER FOR MINOR CLOSED PROPERTIES ON BOUNDED DEGREE GRAPHS*

AKASH KUMAR†, C. SESHADHRI‡, AND ANDREW STOLMAN‡

Abstract. Let G be an undirected, bounded degree graph with n vertices. Fix a finite graph H, and suppose one must remove εn edges from G to make it H-minor-free (for some small constant $\varepsilon > 0$). We give an $n^{1/2+o(1)}$-time randomized procedure that, with high probability, finds an H-minor in such a graph. As an application, suppose one must remove εn edges from a bounded degree graph G to make it planar. This result implies an algorithm, with the same running time, that produces a $K_{3,3}$- or K_5-minor in G. No prior sublinear time bound was known for this problem. By the graph minor theorem, we get an analogous result for any minor-closed property. Up to $n^{o(1)}$ factors, this resolves a conjecture of Benjamini, Schramm, and Shapira [*Adv. Math.*, 223 (2010), pp. 2200–2218] on the existence of one-sided property testers for minor-closed properties. Furthermore, our algorithm is nearly optimal by an $\Omega(\sqrt{n})$ lower bound of Czumaj et al. [*Random Structures Algorithms,* 45 (2014), pp. 139–184]. Prior to this work, the only graphs H for which nontrivial one-sided property testers were known for H-minor-freeness were the following: H being a forest or a cycle [Czumaj et al., *Random Structures Algorithms,* 45 (2014), pp. 139–184], $K_{2,k}$, $(k \times 2)$-grid, and the k-circus [Fichtenberger et al., preprint, arXiv:1707.06126v1, 2017].

Key words. property testing, minor-free graphs, bounded degree graphs

AMS subject classifications. 68Q25, 68R10, 68W20, 05C83

DOI. 10.1137/19M1245463

1. Introduction.

Deciding whether an n-vertex graph G is planar is a classic algorithmic problem solvable in linear time [21]. The Kuratowski–Wagner theorem asserts that any nonplanar graph must contain a K_5- or $K_{3,3}$-minor [27, 36]. Thus, certifying nonplanarity is equivalent to producing such a minor, which can be done in linear time. Can we beat the linear time bound if we know that G is "sufficiently" nonplanar?

Assume random access to an adjacency list representation of a bounded degree graph, G. Suppose, for some constant $\varepsilon > 0$, one had to remove εn edges from G to make it planar. Can one find a forbidden (K_5- or $K_{3,3}$-)minor in $o(n)$ time? It is natural to ask this question for any property expressible through forbidden minors. By the famous Robertson–Seymour graph minor theorem [32], any graph property $\mathcal{P}$ that is closed under taking minors can be expressed by a finite list of forbidden minors. We desire sublinear time algorithms to find a forbidden minor in any G that requires εn edge deletions to make it have $\mathcal{P}$.

This problem was first posed by Benjamini, Schramm, and Shapira [5] in the context of property testing on bounded degree graphs. We follow the model of property testing on bounded degree graphs as defined by Goldreich and Ron [18]. Fix a degree

*Received by the editors February 21, 2019; accepted for publication (in revised form) August 27, 2020; published electronically December 18, 2020.
https://doi.org/10.1137/19M1245463
Funding: The work of the authors was supported by NSF grants CCF-1319080 and CCF-1740850.

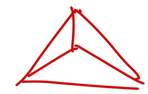
†Department of Computer Science, Purdue University, West Lafayette, IN 47907 USA (akumar@purdue.edu, https://sites.google.com/view/kumarakash/home).
‡Department of Computer Science, University of California, Santa Cruz, Santa Cruz, CA 95062 USA (sesh@ucsc.edu, https://users.soe.ucsc.edu/~sesh/index.html; astolman@ucsc.edu, https://users.soe.ucsc.edu/~astolman).

Section 4: Bulk

Section 5: Simple Algorithm

Thm: (large vertices) expanders are certifiably non-planar

• All planar graphs admit a straight line drawing



K_4

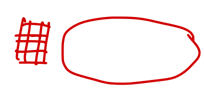
Fix $\phi > 0$. Then $\exists$ an algorithm which on input a graph $G=(V,E)$ on n-vertices, d-regular $\phi(G) \ge \phi$ runs in $\tilde{O}(\sqrt{n})$ time, returns K_3/K_{55} in G w.p. $\ge 2/3$

Stratification
→ only to handle non-expanding case

Key idea: Don't need r-walks to mix in few steps from every vertex.
mixing from $n^{0.99}$ vtx
b'day paradox goes thru.
→ Look at better notion of mixing

Looser defn: Track $\|p(v,t)\|_2^2$

Ideal: $\|p(v,t) - \pi\|_1 \le \varepsilon$



large $\|p(v,t)\|_2^2$ exp.

2nd realisation:

Sps G is bounded degree on b vertices

$\|p(v,t)\|_2^2 \ge \frac{1}{b}$ (cauchy-schwarts)

$\forall t\ \forall v$

Bucket v acc. to $\|p(v,t)\|_2^2$ $t = n^{\delta}$

$B_0 \equiv$ fallen set $B_0 = \{v \in V : n^{-\delta} \le \|p(v,t)\|_2^2 \le 1\}$

B_1

$R_1 = R_0 \setminus B_0$

$B_1 = \{v \in R_1 : n^{-2\delta} \le \|p(v,t)\|^2 \le n^{-\delta}\}$

B_i = expanding vtx. $1/\delta$ rounds to get $= 1/n$

bound d. Consider $G = (V, E)$, where $V = [n]$, and G is represented by an adjacency list. 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 $\perp$).

Given any property $\mathcal{P}$ of graphs with degree bound d, the distance of G to $\mathcal{P}$ is defined as the minimum number of edge additions/removals required to make G have $\mathcal{P}$ divided by dn. This ensures that the distance is in $[0, 1]$. 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 one-sided tester must accept $G \in \mathcal{P}$ with probability 1 and thus must provide a certificate of rejection.

We are interested in properties expressible through *forbidden minors*. Fix a finite graph H. The property $\mathcal{P}_H$ of *H-minor-freeness* is the set of graphs that do not contain H as a minor. Observe that one-sided testers for $\mathcal{P}_H$ have a special significance since they must produce an H-minor whenever they reject. One can cast one-sided property testers for $\mathcal{P}_H$ as sublinear time procedures that find forbidden minors. Our main theorem follows.

THEOREM 1.1. *Fix a finite graph H with $|V(H)| = r$ and arbitrarily small $\delta > 0$. Let $\mathcal{P}_H$ be the property of H-minor-freeness. There is a randomized algorithm that takes as input (oracle access to) a graph G with maximum degree d and a parameter $\varepsilon > 0$. Its running time is $dn^{1/2+O(\delta r^2)} + d\varepsilon^{-2\exp(2/\delta)/\delta}$. If G is ε-far from $\mathcal{P}_H$, then, with probability $> 2/3$, the algorithm outputs an H-minor in G.*

Equivalently, there exists a one-sided property tester for $\mathcal{P}_H$ with the above running time.

The graph minor theorem of Robertson and Seymour [32] asserts the following. Consider any property $\mathcal{Q}$ that is closed under taking minors. There is a finite list $\boldsymbol{H}$ of graphs such that $G \in \mathcal{Q}$ iff G is H-minor-free for all $H \in \boldsymbol{H}$. If G is ε-far from $\mathcal{Q}$, then G is $\Omega(\varepsilon)$-far from $\mathcal{P}_H$ for some $H \in \boldsymbol{H}$. Thus, a direct corollary of Theorem 1.1 is the following.

COROLLARY 1.2. *Let $\mathcal{Q}$ be any minor-closed property of graphs with degree bound d. For any $\delta > 0$, there is a one-sided property tester for $\mathcal{Q}$ with running time $O(dn^{1/2+\delta} + d\varepsilon^{-2\exp(2/\delta)/\delta})$.*

In the following discussion, we suppress dependence on ε and n^δ by $O^*(\cdot)$. Previously, the only graphs H for which an analogue of Theorem 1.1 was known are the following: $O^*(1)$ time for H being a forest, $O^*(\sqrt{n})$ for H being a cycle [7], and $O^*(n^{2/3})$ for H being $K_{2,k}$, the $(k \times 2)$-grid, and the k-circus [13, 14]. No sublinear time bound was known for planarity.

Corollary 1.2 implies that properties such as planarity, series-parallel graphs, embeddability in bounded genus surfaces, and bounded treewidth are all one-sided testable in $O^*(\sqrt{n})$ time.

We note a particularly pleasing application of Theorem 1.1. Suppose a bounded degree graph, G, has more than $(3+\varepsilon)n$ edges. Then it is guaranteed to be ε-far from being planar, and thus there is an algorithm for finding a forbidden minor in G in $O^*(\sqrt{n})$ time. Since all minor-closed properties have constant average degree bounds, analogous statements can be made for all such properties.

(interesting)
$\varepsilon n \leftarrow$ non-planar
delete!
2-norm is not reliable
no collisions

$n/2$ $n/2$
$\|p_{v,t}\|_2^2 \approx \frac{2}{n}$ $\approx 2/n$

walks from w, w' don't necessarily collide w.h.p. if the only thing you know $\|p_{w,t}\|_2^2 \approx \|p_{w',t}\|_2^2$

$w, w' \sim p_{u,t}$

However, walks from $w, w' \sim p_{u,t}$ likely to collide

1) $\|p_{w,t}\|_2 \approx \|p_{w',t}\|_2$ large inner product.

2) $p_{u,t}(w) \geq \alpha$ and $p_{u,t}(w') \geq \frac{\alpha}{10}$ roughly same probability.

You also know, ①

$1_u^T M^t 1_w = p_{u,t}(w) = p_{w,t}(u) = 1_w^T M^t 1_u$ symm

$p_{u,t}(w') = p_{w',t}(u)$ (d-regular)

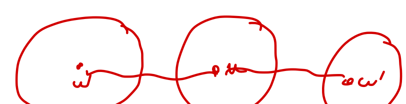
① $\Rightarrow \vec{p}_{u,\frac{t}{2}} \cdot \vec{p}_{w,\frac{t}{2}} \geq \frac{\alpha}{10}$

LHS of ①

$\vec{p}_{w',\frac{t}{2}} \cdot \vec{p}_{u,\frac{t}{2}}$

two vecs with large inner product with same guy, they have large inner prod.

If bigger walk doesn't spread, more penalty.



Binning $\rightarrow$ same length.
we also want large inner product

$w, w' \sim p_{v,t}$

In order to find a minor, I want $p_{w,t}$ to collide with $p_{w',t}$

$\langle p_{w,t}, p_{w',t} \rangle$ large
$\neq p_{w,2t}(w')$ hence $2t$!

We want w, w', u to end up in different bins

1.1. Related work. Graph minor theory is a deep topic, and we refer the reader to Chapter 12 of Diestel's book [11] and to Lovász's survey [29]. For our purposes, we use as a black-box a polynomial time algorithm that finds fixed minors in a graph. A result of Kawarabayashi, Kobayashi, and Reed provides an $O(n^2)$-time algorithm [24].

Property testing on graphs is an immensely rich area of study, and we refer the reader to Goldreich's recent textbook for more details [16]. There is a significant difference between the theory of property testing for dense graphs and that of bounded degree graphs. For the former, there is a complete characterization of properties (one-sided, nonadaptive) testable in query complexity independent of graph size. There is a deep connection between property testing and the Szemeredi regularity lemma [2]. Property testing for bounded degree graphs is much less understood. This study was initiated by Goldreich and Ron, and the first results focused on connectivity properties [18]. Czumaj, Shapira, and Sohler proved that hereditary properties of nonexpanding graphs are testable [9]. A breakthrough result of Benjamini, Schramm, and Shapira (henceforth BSS) proved that all minor-closed (more generally, hyperfinite) properties are two-sided testable in constant time. The dependence on ε was subsequently improved by Hassidim et al. using the concept of local partitioning oracles [20]. A result of Levi and Ron [28] significantly simplified and improved this analysis to get a final query complexity quasi-polynomial in $1/\varepsilon$. Indeed, it is a major open question to get polynomial dependence on $1/\varepsilon$ for two-sided planarity testers. Toward this goal, Yoshida and Ito give such a bound for testing outerplanarity [37], and Edelman et al. generalize the result for bounded treewidth graphs [12].

In contrast to dense graph testing, there is a significant jump in complexity for one-sided testers. BSS first raised the question of one-sided testers for minor-closed properties (especially planarity) and conjectured that the bound is $O(\sqrt{n})$. Czumaj et al. [7] took the first step by giving an $\widetilde{O}(\sqrt{n})$ one-sided tester for the property of being C_k-minor-free [7]. For $k = 3$, this is precisely the class of forests. This tester is obtained by a reduction to a much older result of Goldreich and Ron for one-sided bipartiteness testing for bounded degree graphs [17] (the results in Czumaj et al. are obtained by black-box applications of this result). Czumaj et al. adapt the one-sided $\Omega(\sqrt{n})$ lower bound for bipartiteness and show an $\Omega(\sqrt{n})$ lower bound for one-sided testers for H-minor-freeness when H has a cycle [7]. This is complemented with a constant time tester for H-minor-freeness when H is a forest.

Recently, Fichtenberger et al. give an $\widetilde{O}(n^{2/3})$ tester for H-minor-freeness when H is one of the following graphs: $K_{2,k}$, the $(k \times 2)$-grid, or the k-circus graph (a wheel where spokes have two edges) [13, 14]. This subsumes the properties of outerplanarity and cactus graphs. This result uses a different, more combinatorial (as opposed to random walk based) approach than Czumaj et al. A one step random walk in a graph G that begins at a vertex v consists of picking a neighbor $u \sim N(v)$ uniformly at random (u.a.r.) and going to that vertex.

The use of random walks in property testing was pioneered by Goldreich and Ron [17] and was then (naturally) used in testing expansion properties and clustering structure [19, 10, 23, 31, 22, 8]. Our approach is inspired by the Goldreich–Ron analysis, and we discuss more on this in the next section. A number of previous results have used random walks for routing in expanders [6, 25]. We use techniques from Kale, Peres, and Seshadhri to analyze random walks on projected Markov chains [22]. We also employ the local partitioning method of Spielman and Teng [34], which is in turn derived from the Lovász–Simonovits analysis technique [30].

2. Main ideas. We give an overview of the proof strategy and discuss the various moving parts of the proof. Assume that G is a d-regular graph. It is instructive to understand the method of Goldreich and Ron (henceforth GR) for one-sided bipartiteness testing [17]. The basic idea is to perform $O(\sqrt{n})$ *lazy random walks* of $\mathrm{poly}(\log n)$ length from a u.a.r. vertex s. Recall that a lazy random walk stays at the current vertex with probability 1/2 and moves to a uniform random neighbor (since we assume d-regularity) with probability 1/2. An odd cycle is discovered when two walks end at the same vertex v through paths of differing parity (of length).

The GR analysis first considers the case when G is an expander (and ε-far from bipartite). In this case, the walks from s reach the stationary distribution. One can use a standard collision argument to show that $O(\sqrt{n})$ suffice to hit the same vertex v twice, with different parity paths. The deep insight is that any graph G can be decomposed into pieces where the algorithm works, and each piece P has a small cut to $\overline{P}$. This has connections with decomposing a graph into expander-like pieces [35, 15]. Famously, the Arora–Barak–Steurer algorithm [4] for unique games basically proves such a statement. We note that GR does *not* decompose into expanders but rather into pieces where the expander analysis goes through. So, one might hope to analyze the algorithm by its behavior on each component. Unfortunately, the algorithm cannot produce the decomposition; it can only walk in G and hope that performing random walks in G suffices to simulate the procedure within P. This is extremely challenging and is precisely what GR achieve (this is the bulk of the analysis). The main lemma produces a decomposition into such pieces, such that for each piece P, there exists $s \in P$ wherein short random walks (in G) from s reach all vertices in P with sufficient probability. One can think of this as a simulation argument: we would like to simulate the random walk algorithm running only on P through random walks in G.

The challenge of general minors. With planarity in mind, let us focus on finding K_5-minors. It is highly unlikely that random walks from a single vertex will find a such a minor. Intuitively, we would need to find five different vertices, launch random walks from all of them, and hope these walks produce a minor. Thus, we would need to simulate a much more complex procedure than the (odd) cycle finder of GR. Most significantly, we need to understand the random walk's behavior from multiple sources within P simultaneously. The GR analysis actually constructs the pieces P by a local partitioning looking at the random walk distribution from a single vertex. There is no guarantee on random walk behavior from other vertices in P.

There is a more significant challenge from arbitrary minors. The simulation does not say anything about the specific structure of the paths generated. It only deals with the probability of reaching v from s by a random walk in G when v and s are in the same piece. For bipartiteness, as long as we find two paths of differing parity, we are done. They may intersect each other arbitrarily. For finding a K_5-minor, the actual intersection matters. We would need paths between all pairs of five seed vertices to be "disjoint enough" to give a K_5 minor. This appears extremely difficult using the GR analysis. Even if we did understand the random walk behavior (in G) from all vertices in P, we would have little control over their behavior when they leave P. (Based on the parameters, the walks leave P with high probability.) They may intersect arbitrarily and thus destroy any minor structure.

2.1. When do random walks find minors? Inspired by GR, we start with an algorithm for finding a K_5-minor in an expander G (variants of this idea were present in a result of Kleinberg and Rubinfeld stating that expanders contain an H-minor for any H with $n/\mathrm{poly}(\log n)$ edges [25]). Let ℓ denote the mixing time. Pick u.a.r. a vertex, s, and launch five random walks, each of length ℓ, to reach $v_1, v_2, \ldots, v_5$. From each v_i, launch $\sqrt{n}$ random walks, each of length ℓ. With high probability, a walk from v_i and a walk from v_j will "collide" (end at the same vertex). We can collect these collisions to get paths between all v_i, v_j, and one can, with some effort, show that these form a K_5-minor.

Our main insight is showing that this algorithm, with minor modifications, works even when random walks have extremely slow mixing properties. When the random walks mix even more slowly than the requisite bound, we can essentially perform local partitioning to pull out very small (n^δ for arbitrarily small $\delta > 0$) pieces that have low conductance cuts. We can simply query all edges in this piece and run a planarity test.

There is a parameter $\delta > 0$ that can be set to an arbitrarily small constant. Let us set the random walk length ℓ to n^δ, and let $\mathbf{p}_{s,\ell}$ be the random walk distribution after ℓ steps from s. Our proof splits into two cases, where $\alpha = c\delta$ for explicit constant $c > 1$:

- Case 1 (the leaky case): For at least εn vertices s, $\|\mathbf{p}_{s,\ell}\|_2^2 \leq 1/n^\alpha$.
- Case 2 (the trapped case): For at least $(1-\varepsilon)n$ vertices s, $\|\mathbf{p}_{s,\ell}\|_2^2 > 1/n^\alpha$.

In the leaky case, the upper bound is too weak to argue about convergence of the random walk. We merely have the property that a random walk of length n^δ (roughly speaking) spreads to a set of size $n^{c\delta}$.

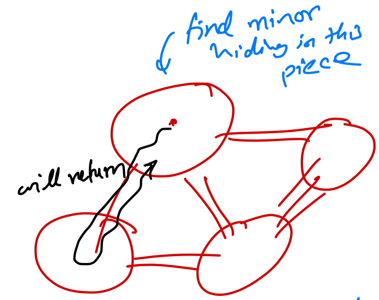
We prove that, in the leaky case, the procedure described in the first paragraph of this subsection succeeds in finding a K_5 with high probability. We give an outline of this proof strategy.

Let us assume that $\mathbf{p}_{v,\ell/2} = \mathbf{p}_{v,\ell}$ (so $\ell/2$-length walks have "stabilized"). Let us make a slight modification to the algorithm. We pick $v_1, \ldots, v_5$ as before, with ℓ-length random walks from s. We will perform $O(\sqrt{n})$ $\ell/2$-length random walks from each v_i to produce the K_5-minor. By symmetry of the random walks, the probability that a single walk from v_i and one from v_j collide (to produce a path) is exactly $\mathbf{p}_{v_i,\ell/2} \cdot \mathbf{p}_{v_j,\ell/2}$. Thus, we would like these dot products to be large. By the symmetry of the random walk, the probability of an ℓ-length random walk starting from s and ending at v is $\mathbf{p}_{s,\ell/2} \cdot \mathbf{p}_{v,\ell/2}$. In other words, the entries of $\mathbf{p}_{s,\ell}$ are precisely these dot products, and $\|\mathbf{p}_{s,\ell}\|_2^2 = \sum_{v \in V} (\mathbf{p}_{s,\ell/2} \cdot \mathbf{p}_{v,\ell/2})^2 = \mathbf{E}_{v \sim \mathbf{p}_{s,\ell/2}}[\mathbf{p}_{s,\ell/2} \cdot \mathbf{p}_{v,\ell/2}]$. Since $\mathbf{p}_{s,\ell/2} = \mathbf{p}_{s,\ell}$, we rewrite to get $\mathbf{p}_{s,\ell/2} \cdot \mathbf{p}_{s,\ell/2} = \mathbf{E}_{v \sim \mathbf{p}_{s,\ell/2}}[\mathbf{p}_{s,\ell/2} \cdot \mathbf{p}_{v,\ell/2}]$.

Think of the dot products as correlations between distributions. We are saying that the average correlation (over some distribution on vertices) of $\mathbf{p}_{v,\ell/2}$ with $\mathbf{p}_{s,\ell/2}$ is exactly the self-correlation of $\mathbf{p}_{s,\ell/2}$. If the distributions mostly had low ℓ_2-norm (as in the leaky case), we might hope that these distributions were reasonably correlated with each other. Indeed, this is what we prove in Lemma 4.11. Under some conditions, we show that $\mathbf{E}_{v_i,v_j \sim \mathbf{p}_{s,\ell/2}}[\mathbf{p}_{v_i,\ell/2} \cdot \mathbf{p}_{v_j,\ell/2}]$ can be lower bounded, where $\mathbf{p}_{s,\ell/2}$ is exactly the distribution from which the algorithm picks v_i and v_j. This is evidence that $\ell/2$-length random walks will connect the v_i's through collisions.

There are four difficulties in increasing order of worry:

1. We only have a lower bound of the average $\mathbf{p}_{v_i,\ell/2} \cdot \mathbf{p}_{v_j,\ell/2}$. We would need bounds for all (or most) pairs to produce a minor.
2. $\mathbf{p}_{v,\ell}$ might be very different from $\mathbf{p}_{v,\ell/2}$.
3. The expected number of collisions between walks from v_i and v_j is controlled by the dot product above, but the variance (which really controls the probability of getting a collision) can be large. There are instances where the dot product is high, but the collision probability is extremely low.
4. There is no guarantee that these paths will produce a minor since we do not have any obvious constraints on the intermediate vertices in the path.



ABS decomp. — don't compute

The first problem is surmounted by a technical trick. It turns out to be cleaner to analyze the probability of getting a biclique minor. Observe that $K_{r,r}$ has K_r as a minor. (Contract all the edges in any perfect matching of $K_{r,r}$ to obtain K_r.) So, we perform 10 random walks from s to get sets $A = \{a_1, a_2, \ldots, a_5\}$ and an analogous B. We launch $\ell/2$-length random walks from each vertex in $A \cup B$. The average lower bound on the dot product suffices to get a lower bound on the probability of getting a $K_{5,5}$-minor, which contains a K_5-minor.

For the second problem, it turns out that the weaker bound of $\|\mathbf{p}_{v,\ell}\|_2 = \Omega(n^{-\delta}\|\mathbf{p}_{v,\ell/2}\|_2)$ suffices. We could try to search for some value of ℓ where this happens. If there is no (small) value of ℓ where this bound held, then this suggests that $\|\mathbf{p}_{v,n^\delta}\|_2$ is extremely small (say $\Theta(1/n)$). This reasoning is discussed in more detail in the next subsection.

$p_1^T = p_0^T M'$

disapper proximan-

The third problem requires bounds on the variance, or higher norms, of $\mathbf{p}_{v,\ell/2}$. Unfortunately, there appears to be no handle on these. At a high level, our idea is to truncate $\mathbf{p}_{v,\ell/2}$ by ignoring large entries. This truncated vector is not a probability vector anymore, but we can hope to redo the analysis for such vectors.

gamble $\{$ $q^{t\ell}_{v,s} = 1_v^T (I_S M^\ell)$ — $S \times 1$ block — 1_u

outside S, mass $= 0$.

Now for the fourth problem. Naturally, if the vertices $v_1, \ldots, v_5$ are close to one another, we do not expect to get a minor by connecting them. Suppose they were sufficiently "spread out"; one could hope that the paths connecting the v_i, v_j pairs would only intersect "near" the v_i's. The portion of the paths near the v_i's could be contracted to get a K_5-minor. We can roughly quantify how far the v_i's will be by the variance of $\mathbf{p}_{v,\ell/2}$. Thus, the third and fourth problems are coupled.

Binning again !

2.2. *R*-returning walks. The main technical contribution of our work is in defining R-returning walks. These are walks that periodically return to a given set of vertices, R. A careful analysis of these walks provides the tools for handling the various problems discussed above.

Fix ℓ as before. Formally, an R-returning walk of length $j\ell$ (for $j \in \mathbb{N}$) is a walk that encounters R at every $i\ell$ step for all $i \in [j]$. While random walk distributions can have poor variance, we can carefully choose R to ensure that the distribution of R-returning walks is well behaved. We will quantify this as approximate "support uniformity" (being approximately uniform on the support).

In the leaky case, there is some (large) set, R, such that for all $s \in R$, $\|\mathbf{p}_{s,\ell/2}\|_2^2 \leq 1/n^\alpha$. Let $\mathbf{p}_{[R],s,\ell}$ be the random walk distribution restricted to R. Suppose for some $s \in R$, $\|\mathbf{p}_{[R],s,\ell}\|_2^2 \geq 1/n^{\alpha+\delta}$. Observe that each entry in $\mathbf{p}_{[R],s,\ell}$ is $\mathbf{p}_{s,\ell/2} \cdot \mathbf{p}_{v,\ell/2}$ for $s, v \in R$. By the Cauchy–Schwarz inequality, this is at most $1/n^\alpha$. For any distribution $\mathbf{v}$, the condition $\|\mathbf{v}\|_2^2 = \|\mathbf{v}\|_\infty$ is equivalent to support uniformity. While $\mathbf{p}_{[R],s,\ell}$ is not a distribution, if we assume that $\|\mathbf{p}_{[R],s,\ell}\|_1$ is sufficiently large, we deduce that $\mathbf{p}_{[R],s,\ell}$ is approximately support uniform. (When R is sufficiently large, one can prove that $\|\mathbf{p}_{[R],s,\ell}\|_1$ is large.) The math discussed in the previous subsection goes through for any such $s \in R$. In other words, if the random walk algorithm started from s, it succeeds in finding a K_5-minor.

Suppose only a negligible fraction of vertices satisfies this condition, and in this case our algorithm would not actually find such a vertex. Let us remove all these vertices from R (abusing notation, let R be the resulting set). Now, for all $s \in R$, $\|\mathbf{p}_{[R],s,\ell}\|_2^2 \leq 1/n^{\alpha+\delta}$. So, the bound on the l_2-norm has fallen by an n^δ factor. What does $\mathbf{p}_{[R],s,\ell} \cdot \mathbf{p}_{[R],v,\ell}$ signify? This is the probability of a 2ℓ-length random walk starting from s, ending at v, and encountering R at the ℓth step. This is an R-returning walk of length 2ℓ. Let $\mathbf{q}_{[R],s,2\ell}$ denote the vector of R-returning walk probabilities. Suppose that for some s, $\|\mathbf{q}_{[R],s,2\ell}\|_2^2 \geq 1/n^{\alpha+2\delta}$. By Cauchy–

Schwarz, $\|\mathbf{q}_{[R],s,2\ell}\|_\infty \leq 1/n^{\alpha+\delta}$, implying that $\mathbf{q}_{[R],s,2\ell}$ is approximately support uniform. Again, the math of the previous subsection goes through for such an s.

We remove all vertices that have this property and end up with R such that for all $s \in R$, $\|\mathbf{q}_{[R],s,2\ell}\|_2^2 \leq 1/n^{\alpha+2\delta}$. Observe that $\mathbf{q}_{[R],s,2\ell} \cdot \mathbf{q}_{[R],v,2\ell}$ is the probability of a 4ℓ R-returning walk. We then iterate this argument.

In general, this argument goes through phases. In the ith phase, we find all $s \in R$ that satisfy $\|\mathbf{q}_{[R],s,2^i\ell}\|_2^2 \geq 1/n^{\alpha+i\delta}$. We show that the random walk procedure of the previous section (with some modifications) finds a K_5-minor starting from such vertices. We remove all such vertices from R, increment i, and continue the argument. The vertices removed at the ith phase are called the *ith stratum*, and we refer to this entire process as stratification. Intuitively, for vertices in the ith stratum, the R-returning (for the setting of R at that phase) walk probabilities roughly form a uniform distribution of support $n^{\alpha+i\delta}$. Thus, for vertices in higher strata, the random walks are spreading to larger sets.

There is a major problem. The $\mathbf{q}$ vectors are *not* distributions, and the vast majority of walks are not R-returning. Indeed, the reduction in norm as we increase strata might simply be an artifact of the lower probability of a longer R-returning walk (note that the walk lengths are increasing exponentially in the phase number). We prove a spectral lemma asserting that this is not the case. As long as R is sufficiently large, the probabilities of R-returning walks are sufficiently high. Unfortunately, these probabilities (must) decrease exponentially in the number of returns. In the ith phase, the walk length is $2^i\ell$ and must return to R 2^i times. Here is where the n^δ decay in the l_2-norm condition saves us. After $1/\delta$ phases, the $\|\mathbf{q}_{[R],s,2^i\ell}\|_2^2$ is basically $1/n$. The spectral lemma tells us that if R is still large, the probability that a $2^{1/\delta}\ell$-length walk is R-returning is sufficiently large. Thus, the norm cannot decrease, and almost all vertices end up in the very next stratum. If R were small, then there would be an earlier stratum containing $\Omega(\delta\varepsilon n)$ vertices. Regardless of the case, there exists an $i \leq 1/\delta + O(1)$ such that the ith stratum contains $\Omega(\delta\varepsilon n)$ vertices. For all these vertices, the random walk algorithm for finding minors succeeds with nontrivial probability.

2.3. The trapped case: Local partitioning to the rescue. In this case, for almost all vertices, $\|\mathbf{p}_{s,\ell}\|_2^2 \geq 1/n^\alpha$. The proof of the (contrapositive of the) Cheeger inequality basically implies the existence of a set of low conductance cut P_s "around" s. By local partitioning methods such as those of Spielman and Teng and Andersen, Chung, and Lang [34, 3], we can actually find P_s in roughly n^α time. We expect our graph to basically decompose into $O(n^\alpha)$-sized components with few edges between them. Our algorithm can simply find these pieces P_s and run a planarity test on them. We refer to this as the *local search* procedure.

While the intuition is correct, the analysis is difficult. The main problem is that actual partitioning of the graph (into small components connected by low conductance cuts) is fundamentally iterative. It starts by finding a low conductance set P_{s_1}; it then finds a low conductance set P_{s_2} in $\overline{P_{s_1}}$, a set P_{s_3} in $\overline{P_{s_1} \cup P_{s_2}}$, and so on. In general, this requires conditions on the random walk behavior inside $\overline{\bigcup_{j<i} P_{s_j}}$. On the other hand, our algorithm and the trapped case condition only refer to random walk behavior in all of G. Furthermore, $\overline{\bigcup_{j<i} P_{s_j}}$ can be as small as $\Theta(\varepsilon n)$, and so we can expect the random walk behavior to be quite different.

The bipartiteness analysis of GR surmounts this problem and performs such a decomposition, but their parameters do not work for us. Starting from a source vertex s, their analysis discovers P_s such that probability of reaching any vertex in

P_s (from s) is roughly uniform *and* smaller than $1/\sqrt{n}$. On the other hand, we would like to discover all of P_s in $n^{O(\delta)}$ time so that we can run a full planarity test.

We employ a collection of tools, and use the methods of Kale, Peres, and Seshadhri to analyze "projected" Markov chains [22]. In the analysis above, we have some set S ($\overline{\bigcup_{j<i} P_{s_j}}$) and want to find a low conductance set P completely contained in S. Moreover, we wish to discover P using random walks in G. We construct a Markov chain, M_S, with vertex set S, and include new transitions that correspond to walks in G whose intermediate vertices are not in S. Each such transition has an associated "cost," corresponding to the actual length in G. (GR also have a similar idea, although their Markov chain introduces extra vertices to track the length of the walk in G. This makes the analysis somewhat unwieldy, since low conductance cuts in M_S may include these extra vertices.)

Using bounds on the return time of random walks, we have relationships between the average length of a walk in G, whose endpoints are in S, and the corresponding length when "projected" to M_S. On average, an ℓ-length walk in G with endpoints in S corresponds to an $\ell|S|/n$-length walk in M_S. Roughly speaking, we hope that for many vertices s, an $\ell|S|/n$-length walk in M_S is trapped in a set of size n^α.

We employ the Lovász–Simonovits curve technique to produce a low conductance cut P_s in M_S [30]. We can guarantee that all vertices in P_s are reachable with roughly $n^{-\alpha}$ probability from s through $\ell|S|/n$-length random walks in M_S. Using the average length correspondence between walks in M_S to G, we can make a similar statement in G—albeit with a longer length. We basically iterate over this entire argument to produce the decomposition into low conductance pieces.

In our analysis, we use the stratification itself to (implicitly) distinguish between the leaky and trapped cases. Stratification peels the graph into $1/\delta + O(1)$ strata. If a vertex s lies in a stratum numbered as at least some fixed constant b, we can show that the algorithm finds a K_r-minor with s as the starting vertex. Thus, if at least (say) $n^{1-\delta}$ vertices lie in stratum b or higher, we are done. If s is in a low stratum, we have a lower bound on the random walk norm. This allows for local partitioning around s.

$B_{1/\delta}$ = higher bins

2 step algorithm :

1. Handles situation

// $|B_{HIGH}| \geq n^{0.99}$

Pick $s \sim V$

1) Launch $\sqrt{}$ walks of length n^δ from s

call destinations $w_1, \dots, w_{\sqrt{}}$

2) For $\ell = 1$ to n^δ

Run ℓ-length walks from each w_i ($\sqrt{n}$)

- If pairwise collision found b/w $w_1, \dots w_{\sqrt{}}$ then if the colliding walks contain a minor, return it
- else continue

→ If minor is never found, goto PHASE 2

3. The algorithm.

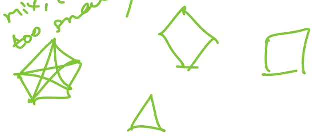
We are given a bounded degree graph $G = (V, E)$, with max degree d. We assume that $V = [n]$. We follow the standard adjacency list model of GR for (random) access to the graph. This model allows an algorithm to sample u.a.r. vertices and perform *edge queries*. Given a pair $(v, i) \in [n] \times [d]$, the output of an edge query is the ith neighbor of v according to the adjacency list ordering. If the degree of v is smaller than i, the output is $\perp$.

In the algorithm, the phrase "random walk" refers to a lazy random walk on G. Given a current vertex, v, with probability $1/2$, the walk remains at v. With probability $1/2$, the procedure generates u.a.r. $i \in [d]$. It performs the edge query for (v, i). If the output is $\perp$, the walk remains at v; otherwise, the walk visits the output vertex. This is a symmetric, ergodic Markov chain with a uniform stationary distribution.

Our main procedure, `FindMinor`(G, ε, H), tries to find an H-minor in G. We prove that it succeeds with high probability if G is ε-far from being H-minor-free. There are three subroutines:

- `LocalSearch`(s): This procedure performs a small number of short random walks to find the piece described in section 2.3. This produces a small subgraph of G, where an exact H-minor-finding algorithm is used.
- `FindPath`(u, v, k, i): This procedure tries to find a path from u to v. The

(don't mix, cuts too small)



// PHASE 2 : $|B_{HIGH}| \leq n^{0.99}$

① Repeat n^{δ} times

(a) Pick $s \sim V$

if not

(b) Use LS to find low conductance cut around s, call this set F_s

(c) Run minor finding algo on graph induced $G[F_s]$

parameter i decides the length of the walk, and the procedure performs k walks from u and v. If any pair of these walks collide, this path is output.

- `FindBiclique`(s): This is the main procedure mostly as described in section 2.1. It attempts to find a sufficiently large biclique minor. First, it generates seed sets A and B by performing random walks from s. Then, it calls `FindPath` on all pairs in $A \times B$.

We fix a collection of parameters.

- δ: An arbitrarily small constant.
- r: The number of vertices in H.
- ℓ: The random walk length. This will be $n^{5\delta}$.
- $\varepsilon_{\mathsf{CUTOFF}}$: $\varepsilon_{\mathsf{CUTOFF}} = n^{\frac{-\delta}{exp(2/\delta)}}$. If $\varepsilon < \varepsilon_{\mathsf{CUTOFF}}$, the algorithm just queries the whole graph.
- `KKR`(F, H): This refers to an exact H-minor-finding process (in F). For concreteness, we use the $O(|V(G)|^2)$ procedure of Kawarabayashi, Kobayashi, and Reed [24].

`FindMinor`(G, ε, H)

1. If $\varepsilon < \varepsilon_{\mathsf{CUTOFF}}$, query all of G, and output `KKR`(G, H)
2. Else
   (a) Repeat $\varepsilon^{-2} n^{35\delta r^2}$ times:
      i. Pick u.a.r. $s \in V$
      ii. Call `LocalSearch`(s) and `FindBiclique`(s).

`LocalSearch`(s)

1. Initialize set $B = \emptyset$.
2. For $h = 1, \ldots, n^{7\delta r^2}$:
   (a) Perform $\varepsilon^{-1} n^{30\delta r^2}$ independent random walks of length h from s. Add all destination vertices to B.
3. Determine $G[B]$, the subgraph induced by B.
4. Run `KKR`$(G[B], H)$. If it returns an H-minor, output that and terminate.

`FindBiclique`(s)

1. For $i = 5r^2, \ldots, 1/\delta + 4$:
   (a) Perform $2r$ independent random walks of length $2^{i+1}\ell$ from s. Let the destinations of the first r walks be multiset A, and let the destinations of the remaining walks be B.
   (b) For each $a \in A$, $b \in B$:
      i. Run `FindPath`$(a, b, n^{\delta(i+18)/2}, i)$.
   (c) If all calls to `FindPath` return a path, then let the collection of paths be the subgraph F. Run `KKR`(F, H). If it returns an H-minor, output that and terminate.

`FindPath`(u, v, k, i)

1. Perform k random walks of length $2^i\ell$ from u and v.
2. If walks from u and v terminate at the same vertex, return these paths. (Otherwise, return nothing.)

THEOREM 3.1. *If G is ε-far from being H-minor-free, then* `FindMinor`(G, ε, H) *finds an H-minor of G with probability at least $2/3$. Furthermore,* `FindMinor` *has a running time of $dn^{1/2+O(\delta r^2)} + d\varepsilon^{-2\exp(2/\delta)/\delta}$.*

The query complexity is fairly easy to compute. The total queries made in the `LocalSearch` calls is $dn^{O(\delta r^2)}$. The main work happens in the calls of `FindPath`, within `FindBiclique`. Observe that k is set to $n^{\delta(i+18)/2}$, where $i \leq 1/\delta+4$. This leads to the $\sqrt{n}$ in the final complexity. (In general, a setting of $\delta < 1/\log(\varepsilon^{-1}\log\log n)$ suffices for an $n^{1/2+o(1)}$ running time.)

Outline. There are a number of moving parts in the proof, which we relegate to their own subsections. We first develop the notion of R-returning walks and the stratification process in section 4. In section 5, we use these techniques to prove that `FindBiclique` discovers a sufficiently large biclique-minor in the leaky case. In section 6, we prove a local partitioning lemma that will be used to handle the trapped case. Finally, in section 7 we put the tools together to complete the proof of Theorem 3.1.

4. Returning walks and stratification.

We introduce the concept of R-returning random walks for any $R \subseteq V$. These definitions are with respect to a fixed length ℓ. (See Table 1.)

TABLE 1
Stratification notation.

Notation	Meaning	Where defined
$\boldsymbol{q}^{(i)}_{[R],s}$	R-returning probability vector	section 4, Def. 4.1
$q^{(i)}_{[R],s}(u)$	Probability of R-returning walk ending at $u \in R$	section 4, Def. 4.1
$\hat{q}^{(i)}_{[R_i],s}(u)$	Distribution on R_i induced by $\boldsymbol{q}^{(i)}_{[R_i],s}$	Def. 4.10

DEFINITION 4.1. *For any set of vertices R, $s \in R$, $u \in R$, and $i \in \mathbb{N}$, we define the* R-returning probability *as follows. We denote by $q^{(i)}_{[R],s}(u)$ the probability that a $2^i\ell$-length random walk from s ends at u, and encounters a vertex in R at every $j\ell$th step for all $1 \leq j \leq 2^i$. The R-returning probability vector, denoted by $\boldsymbol{q}^{(i)}_{[R],s}$, is the $|R|$-dimensional vector of returning probabilities.*

PROPOSITION 4.2. $q^{(i+1)}_{[R],s}(u) = \boldsymbol{q}^{(i)}_{[R],s} \cdot \boldsymbol{q}^{(i)}_{[R],u}$.

Proof. We use the symmetry of (returning) random walks in G.

$$q^{(i+1)}_{[R],s}(u) = \sum_{w\in S} q^{(i)}_{[R],s}(w) q^{(i)}_{[R],w}(u) = \sum_{w\in R} q^{(i)}_{[R],s}(w) q^{(i)}_{[R],u}(w) = \boldsymbol{q}^{(i)}_{[R],s} \cdot \boldsymbol{q}^{(i)}_{[R],u}. \qquad \square$$

Let M be the transition matrix of the lazy random walk on G. Let $\mathbb{P}_R$ be the $n \times |R|$ matrix on R, where each column is the unit vector for some $s \in R$. For any set U, we use $\mathbf{1}_U$ for the indicator vector on U. If no subscript is given, it is the all-ones vector for the appropriate dimension.

PROPOSITION 4.3. $\boldsymbol{q}^{(i)}_{[R],s} = (\mathbb{P}_R^T M^\ell \mathbb{P}_R)^{2^i} \mathbf{1}_s$.

Now we introduce a critical lemma. We can lower bound the total probability of an R-returning random walk. If R contains at least a β-fraction of vertices, the average R-returning walk probability for t returns is at least β^t.

LEMMA 4.4. $|R|^{-1} \sum_{s\in R} \|\boldsymbol{q}^{(i)}_{[R],s}\|_1 \geq (|R|/n)^{2^i}$.

Proof. We will express $\sum_{s\in R} \|\boldsymbol{q}^{(i)}_{[R],s}\|_1 = \mathbf{1}^T(\mathbb{P}_R^T M^\ell \mathbb{P}_R)^{2^i}\mathbf{1}$. Let us first prove the lemma for $i = 0$. Observe that

$$\sum_{s \in R} \|\boldsymbol{q}^{(0)}_{[R],s}\|_1 = \mathbf{1}_R^T M^\ell \mathbf{1}_R = ((M^T)^{\ell/2}\mathbf{1}_R)^T (M^{\ell/2}\mathbf{1}_R) = \|M^{\ell/2}\mathbf{1}_R\|_2^2.$$

Since $M^{\ell/2}$ is a stochastic matrix, $\|M^{\ell/2}\mathbf{1}_R\|_1 = \|\mathbf{1}_R\|_1 = |R|$. By a standard norm inequality, $\|M^{\ell/2}\mathbf{1}_R\|_2^2 \geq \|M^{\ell/2}\mathbf{1}_R\|_1^2/n = |R|^2/n$. This completes the proof for $i = 0$.

Let $N = \mathbb{P}_R^T M^\ell \mathbb{P}_R$, which is a symmetric matrix. We have just proven that $\mathbf{1}^T N \mathbf{1} \geq |R|^2/n$. Let the eigenvalues of N be $1 \geq \lambda_1 \geq \lambda_2, \ldots, \lambda_{|R|}$, with corresponding eigenvectors $\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_s}$. We can express $\mathbf{1} = \sum_{k \leq |R|} \alpha_k \mathbf{u_k}$, where $\sum_k \alpha_k^2 = |R|$. Observe that $N^{2^i}\mathbf{1} = \sum_{k \leq |R|} \alpha_k \lambda_k^{2^i} \mathbf{u_k}$

Let $\mu_k = \alpha_k^2 / \sum_j \alpha_j^2$, noting that $\sum_k \mu_k = 1$. We apply Jensen's inequality below.

$$\frac{\mathbf{1}^T N^{2^i} \mathbf{1}}{|R|} = \frac{\sum_k \alpha_k^2 \lambda_k^{2^i}}{\sum_j \alpha_j^2} = \sum_k \mu_k \lambda_k^{2^i} \geq \left(\sum_k \mu_k \lambda_k\right)^{2^i}.$$

For $i = 0$, we already proved that $\mathbf{1}^T N \mathbf{1}/|R| = \sum_k \mu_k \lambda_k \geq |R|/n$. We plug in this bound to complete the proof for general i. ∎

4.1. Stratification. Stratification results in a collection of disjoint sets of vertices denoted by $S_0, S_1, \ldots$ which are called *strata*. The corresponding *residue* sets are denoted by $R_0, R_1, \ldots$. The zeroth residue R_0 is initialized before stratification, and subsequent residues are defined by the recurrence $R_i = R_0 \setminus \bigcup_{j<i} S_j$. The definitions and claims may seem technical, and the proofs are mostly norm manipulations. But these provide the tools for analyzing our main algorithm.

DEFINITION 4.5. *Suppose R_i has been constructed. A vertex $s \in R_i$ is placed in S_i if $\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_2^2 \geq 1/n^{\delta i}$.*

We have an upper bound for the length of R_i-returning walk vectors.

CLAIM 4.6. *For all $s \in R_i$ and $1 \leq j \leq i$, $\|\boldsymbol{q}^{(j)}_{[R_i],s}\|_2^2 \leq 1/n^{\delta(j-1)}$.*

Proof. Suppose $\exists j \leq i, \|\boldsymbol{q}^{(j)}_{[R_i],s}\|_2^2 > 1/n^{\delta(j-1)}$. By assumption, $s \in R_i \subseteq R_{j-1}$. An R_i-returning walk from s is also an R_{j-1}-returning walk. Thus, every entry of $\boldsymbol{q}^{(j)}_{[R_{j-1}],s}$ is at least that of $\boldsymbol{q}^{(j)}_{[R_i],s}$. So $\|\boldsymbol{q}^{(j)}_{[R_{j-1}],s}\|_2^2 \geq \|\boldsymbol{q}^{(j)}_{[R_i],s}\|_2^2 > 1/n^{\delta(j-1)}$. This implies that $s \in S_{j-1}$ or an earlier stratum, contradicting the assumption that $s \in R_i$. ∎

We prove an ℓ_∞ bound on the returning probability vectors. Note that we allow j to be $i+1$ in the following bound.

CLAIM 4.7. *For all $s \in R_i$ and $2 \leq j \leq i+1$, $\|\boldsymbol{q}^{(j)}_{[R_i],s}\|_\infty \leq 1/n^{\delta(j-2)}$.*

Proof. By Proposition 4.2, for any $v \in R_i$, $q^{(j)}_{[R_i],s}(v) = \boldsymbol{q}^{(j-1)}_{[R_i],s} \cdot \boldsymbol{q}^{(j-1)}_{[R_i],v}$. Note that $1 \leq j-1 \leq i$. By Cauchy–Schwarz and Claim 4.6, $q^{(j)}_{[R_i],s}(v) \leq 1/n^{\delta(j-2)}$. ∎

As a consequence of these bounds, we are able to bound the amount of probability mass retained by R_i-returning walks.

CLAIM 4.8. *For all $s \in S_i$, $||\boldsymbol{q}^{(i+1)}_{[R_i],s}||_1 \geq n^{-\delta}$.*

Proof. Since $s \in S_i$, $||\boldsymbol{q}^{(i+1)}_{[R_i],s}||_2^2 \geq n^{-i\delta}$, and by Claim 4.7, $||\boldsymbol{q}^{(i+1)}_{[R_i],s}||_\infty \leq n^{-\delta(i-1)}$. Since $||\boldsymbol{q}^{(i+1)}_{[R_i],s}||_2^2 \leq ||\boldsymbol{q}^{(i+1)}_{[R_i],s}||_1 ||\boldsymbol{q}^{(i+1)}_{[R_i],s}||_\infty$, we conclude that $||\boldsymbol{q}^{(i+1)}_{[R_i],s}||_1 \geq n^{-i\delta} n^{\delta(i-1)} = n^{-\delta}$. ∎

We prove that most vertices lie in "early" strata.

LEMMA 4.9. *Suppose $\varepsilon \geq \varepsilon_{\mathsf{CUTOFF}}$. At most $\varepsilon n/\log n$ vertices are in $R_{1/\delta+3}$.*

Proof. The proof is by contradiction. Suppose that $R_{1/\delta+3}$ has at least $\varepsilon n/\log n$ vertices. The previous residue, $R_{1/\delta+2}$, is bigger, and thus $|R_{1/\delta+2}| \geq \varepsilon n/\log n$ as well. By Lemma 4.4,

$$|R_{1/\delta+2}|^{-1} \sum_{s \in R_{1/\delta+2}} \|\boldsymbol{q}^{(1/\delta+3)}_{[R_{1/\delta+2}],s}\|_1 \geq \left(\frac{\varepsilon}{\log n}\right)^{2^{1/\delta+3}}. \tag{4.1}$$

By averaging and Cauchy–Schwarz (to relate the l_1- and l_2-norms),

$$||\boldsymbol{q}^{(1/\delta+3)}_{[R_{1/\delta+2}],s}\|_2^2 \geq n^{-1} \left(\frac{\varepsilon}{\log n}\right)^{2^{1/\delta+4}}. \tag{4.2}$$

By assumption, $\varepsilon \geq \varepsilon_{\mathsf{CUTOFF}} \geq n^{-\delta/\exp(1/\delta)}$. For sufficiently small δ, $\delta/\exp(1/\delta) < 2\delta/2^{1/\delta+4}$. Thus, $\varepsilon \geq (\log n) n^{-2\delta/(2^{1/\delta+4})}$. Plugging into the RHS of the previous equation, we see that $||\boldsymbol{q}^{(1/\delta+3)}_{[R_{1/\delta+2}],s}\|_2^2 \geq 1/n^{1+2\delta} = 1/n^{\delta(1/\delta+2)}$. This implies that $v \in S_{1/\delta+2}$—a contradiction. ∎

Tf ?

what contradiction ?

4.2. The correlation lemma. The following lemma is an important tool in our analysis. Here is an intuitive explanation. Fix some $s \in S_i$. By Proposition 4.2, the probability $q^{(i+1)}_{[R_i],s}(v)$ is the correlation between the vectors $\boldsymbol{q}^{(i)}_{[R_i],s}$ and $\boldsymbol{q}^{(i)}_{[R_i],v}$. If many of these probabilities are large, then there are many v such that $\boldsymbol{q}^{(i)}_{[R_i],v}$ is correlated with $\boldsymbol{q}^{(i)}_{[R_i],s}$. We then expect many of these vectors to be correlated among themselves.

DEFINITION 4.10. *For $s \in R_i$, the distribution $\mathcal{D}_{s,i}$ has support R_i, and the probability of $u \in R_i$ is $\hat{q}^{(i+1)}_{[R_i],s}(v) = q^{(i+1)}_{[R_i],s}(v)/\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1$.*

LEMMA 4.11. *Fix arbitrary $s \in R_i$; then the following inequality holds:*

$$\mathbf{E}_{u_1,u_2 \sim \mathcal{D}_{s,i}}[\boldsymbol{q}^{(i)}_{[R_i],u_1} \cdot \boldsymbol{q}^{(i)}_{[R_i],u_2}] \geq \frac{1}{\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1^2} \cdot \frac{\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_2^4}{\|\boldsymbol{q}^{(i)}_{[R_i],s}\|_2^2}.$$

Proof.

(4.3)

$$
\begin{aligned}
&\mathbf{E}_{u_1,u_2\sim\mathcal{D}_{s,i}}[\boldsymbol{q}^{(i)}_{[R_i],u_1}\cdot\boldsymbol{q}^{(i)}_{[R_i],u_2}] \\
&= \sum_{u_1,u_2\in R_i} \|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1^{-2} q^{(i+1)}_{[R_i],s}(u_1) q^{(i+1)}_{[R_i],s}(u_2) \boldsymbol{q}^{(i)}_{[R_i],u_1}\cdot\boldsymbol{q}^{(i)}_{[R_i],u_2} \\
&= \|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1^{-2} \sum_{u_1,u_2\in R_i} (\boldsymbol{q}^{(i)}_{[R_i],s}\cdot\boldsymbol{q}^{(i)}_{[R_i],u_1})(\boldsymbol{q}^{(i)}_{[R_i],s}\cdot\boldsymbol{q}^{(i)}_{[R_i],u_2})(\boldsymbol{q}^{(i)}_{[R_i],u_1}\cdot\boldsymbol{q}^{(i)}_{[R_i],u_2}) \\
&= \|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1^{-2} \sum_{u_1,u_2\in R_i} (\boldsymbol{q}^{(i)}_{[R_i],s}\cdot\boldsymbol{q}^{(i)}_{[R_i],u_1})(\boldsymbol{q}^{(i)}_{[R_i],s}\cdot\boldsymbol{q}^{(i)}_{[R_i],u_2}) \sum_{w\in R_i} q^{(i)}_{[R_i],u_1}(w) q^{(i)}_{[R_i],u_2}(w) \\
&= \|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1^{-2} \sum_{\substack{w\in R_i\\ u_1,u_2\in R_i}} [(\boldsymbol{q}^{(i)}_{[R_i],s}\cdot\boldsymbol{q}^{(i)}_{[R_i],u_1}) q^{(i)}_{[R_i],u_1}(w)][(\boldsymbol{q}^{(i)}_{[R_i],s}\cdot\boldsymbol{q}^{(i)}_{[R_i],u_2}) q^{(i)}_{[R_i],u_2}(w)] \\
&= \|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1^{-2} \sum_{w\in R_i} \left[\sum_{u\in R_i} (\boldsymbol{q}^{(i)}_{[R_i],s}\cdot\boldsymbol{q}^{(i)}_{[R_i],u}) q^{(i)}_{[R_i],u}(w)\right]^2.
\end{aligned}
$$

We now write out $\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_2^2 = \sum_{u\in R_i} q^{(i+1)}_{[R_i],s}(u)^2 = \sum_{u\in R_i}(\boldsymbol{q}^{(i)}_{[R_i],s}\cdot\boldsymbol{q}^{(i)}_{[R_i],u})^2$ by Proposition 4.2. We expand this equation further below. The only inequality is Cauchy–Schwarz.

$$
\begin{aligned}
\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_2^2 &= \sum_{u\in R_i} (\boldsymbol{q}^{(i)}_{[R_i],s}\cdot\boldsymbol{q}^{(i)}_{[R_i],u}) \sum_{w\in R_i} q^{(i)}_{[R_i],s}(w) q^{(i)}_{[R_i],u}(w) \\
&= \sum_{w\in R_i} q^{(i)}_{[R_i],s}(w) \Big[\sum_{u\in R_i} (\boldsymbol{q}^{(i)}_{[R_i],s}\cdot\boldsymbol{q}^{(i)}_{[R_i],u}) q^{(i)}_{[R_i],u}(w)\Big] \\
&\le \sqrt{\sum_{w\in R_i} q^{(i)}_{[R_i],s}(w)^2}\sqrt{\sum_{w\in R_i}\Big[\sum_{u\in R_i} (\boldsymbol{q}^{(i)}_{[R_i],s}\cdot\boldsymbol{q}^{(i)}_{[R_i],u}) q^{(i)}_{[R_i],u}(w)\Big]^2} \\
&= \|\boldsymbol{q}^{(i)}_{[R_i],s}\|_2 \|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1 \sqrt{\mathbf{E}_{u_1,u_2\sim\mathcal{D}_{s_i}}[\boldsymbol{q}^{(i)}_{[R_i],u_1}\cdot\boldsymbol{q}^{(i)}_{[R_i],u_2}]} \qquad \text{(by (4.3))}.
\end{aligned}
$$

We rearrange and take squares to complete the proof. ∎

We can apply previous norm bounds to get an explicit lower bound. To see the significance of the following lemma, note that by Claim 4.6 and Cauchy–Schwarz, for all $u_1, u_2 \in R_i$, $\boldsymbol{q}^{(i)}_{[R_i],u_1}\cdot\boldsymbol{q}^{(i)}_{[R_i],u_2} \le 1/n^{\delta(i-1)}$ (fairly close to the lower bound below).

LEMMA 4.12. *Fix arbitrary $s \in S_i$; then the following inequality holds:*

$$\mathbf{E}_{u_1,u_2\sim\mathcal{D}_{s,i}}[\boldsymbol{q}^{(i)}_{[R_i],u_1}\cdot\boldsymbol{q}^{(i)}_{[R_i],u_2}] \ge 1/n^{\delta(i+1)}.$$

Proof. By Lemma 4.11, the LHS is at least $\frac{1}{\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1^2}\cdot\frac{\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_2^4}{\|\boldsymbol{q}^{(i)}_{[R_i],s}\|_2^2}$. Note that $\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1 \le 1$. By Definition 4.5, $\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_2^2 \ge 1/n^{\delta i}$. Since $s \in S_i \subseteq R_i$, by Claim 4.6, $\|\boldsymbol{q}^{(i)}_{[R_i],s}\|_2^2 \le 1/n^{\delta(i-1)}$. ∎

5. Analysis of `FindBiclique`. This is the central theorem of our analysis. It shows that the `FindBiclique`(s) procedure discovers a $K_{r,r}$-minor with nontrivial probability when s is in a sufficiently high stratum.

THEOREM 5.1. *Suppose $s \in S_i$ for $5r^2 \le i \le 1/\delta + 3$. The probability that the paths discovered in* `FindBiclique`(s) *contain a $K_{r,r}$-minor is at least $n^{-4\delta r^2}$.*

Theorem 5.1 is proved in section 5.5. Toward the proof, we will need multiple tools. In section 5.1, we perform a standard calculation to bound the success probability of `FindPath`. In section 5.2, we use this bound to show that the sets A and B sampled by `FindBiclique` are successfully connected by paths as discovered by `FindPath`. In section 5.3, we argue that the intersections between these paths are "well behaved" enough to induce a $K_{r,r}$-minor.

We note that the $\sqrt{n}$ in the final running time comes from the calls to `FindPath` in `FindBiclique`.

5.1. The procedure `FindPath`. For convenience, we reproduce the procedure `FindPath`. It is a relatively straightforward application of a birthday paradox argument for bidirectional path finding.

`FindPath`(u, v, k, i)
1. Perform k random walks of length $2^i\ell$ from u and v.
2. If walks from u and v terminate at the same vertex, return these paths.

LEMMA 5.2. *Let c be a sufficiently large constant. Consider $u, v \in R_i$. Suppose there exist $\alpha \le \beta$ such that $\max(\|\boldsymbol{q}^{(i)}_{[R_i],u}\|_2^2, \|\boldsymbol{q}^{(i)}_{[R_i],v}\|_2^2) \le 1/n^\alpha$ and $\boldsymbol{q}^{(i)}_{[R_i],u} \cdot \boldsymbol{q}^{(i)}_{[R_i],v} \ge 1/2n^\beta$. Then, with $k \ge cn^{\beta/2+4(\beta-\alpha)}$,* `FindPath`$(u, v, k, i)$ *returns an R_i-returning path of length $2^{i+1}\ell$ with probability $\ge 2/3$.*

Proof. First, define $W = \{w | q^{(i)}_{[R_i],u}(w)/q^{(i)}_{[R_i],v}(w) \in [1/(8n^{\beta-\alpha}), 8n^{\beta-\alpha}]\}$.

$$\begin{aligned}\sum_{w \notin W} q^{(i)}_{[R_i],u}(w) q^{(i)}_{[R_i],v}(w) &\le (8n^{\beta-\alpha})^{-1} \sum_{w\notin W} \max(q^{(i)}_{[R_i],u}(w), q^{(i)}_{[R_i],v}(w))^2 \\ &\le (8n^{\beta-\alpha})^{-1}(\|\boldsymbol{q}^{(i)}_{[R_i],u}\|_2^2 + \|\boldsymbol{q}^{(i)}_{[R_i],v}\|_2^2) \le 1/4n^\beta.\end{aligned}$$

Therefore, $\sum_{w\in W} q^{(i)}_{[R_i],u}(w) q^{(i)}_{[R_i],v}(w) \ge 1/4n^\beta$.

For $a, b \le k$, let $X_{a,b}$ be the indicator for the following event: the ath $2^i\ell$-length random walk from u is an R_i-returning walk that ends at some $w \in W$, and the bth random walk from v is also R_i-returning, ending at the same w. Let $X = \sum_{a,b\le k} X_{a,b}$. Observe that the probability that `FindPath`(u, v, k, i) returns a path is at least $\Pr[X > 0]$.

We can bound

$$\mathbf{E}\left[\sum_{a,b\le k} X_{a,b}\right] = k^2 \sum_{w\in W} q^{(i)}_{[R_i],u}(w) q^{(i)}_{[R_i],v}(w) \ge k^2/4n^\beta \ge (c^2/4)n^{8(\beta-\alpha)}.$$

Let us now bound the variance. First, let us expand out the expected square.

$$\begin{aligned}&\mathbf{E}\left[\left(\sum_{a,b} X_{a,b}\right)^2\right] \\ &= \sum_{a,b} \mathbf{E}[X_{a,b}^2] + 2\sum_{a\ne a',b} \mathbf{E}[X_{a,b}X_{a',b}] + 2\sum_{a,b\ne b'} \mathbf{E}[X_{a,b}X_{a',b}] + 2\sum_{a\ne a',b\ne b'} \mathbf{E}[X_{a,b}X_{a',b'}].\end{aligned} \tag{5.1}$$

Observe that $X_{a,b}^2 = X_{a,b}$. Furthermore, for $a \neq a', b \neq b'$, by independence of the walks, $\mathbf{E}[X_{a,b}X_{a',b'}] = \mathbf{E}[X_{a,b}]\mathbf{E}[X_{a',b'}]$. (This term will cancel out in the variance.) By symmetry, $\sum_{a\neq a',b} \mathbf{E}[X_{a,b}X_{a',b}] \leq k^3\mathbf{E}[X_{1,1}X_{2,1}]$ (and analogously for the third term in (5.1)). Plugging these in and expanding out the $\mathbf{E}[X]^2$, we obtain

$$\mathbf{var}[X] \leq \mathbf{E}[X] + 2k^3\mathbf{E}[X_{1,1}X_{2,1}] + 2k^3\mathbf{E}[X_{1,1}X_{1,2}].$$

Note that $X_{1,1}X_{2,1} = 1$ when the first and second walks from u end at the same vertex where the first walk from v ends. Thus, $\mathbf{E}[X_{1,1}X_{2,1}] = \sum_{w\in W} q^{(i)}_{[R_i],u}(w)^2 q^{(i)}_{[R_i],v}(w)$. Since $w \in W$, we have $q^{(i)}_{[R_i],u}(w)/8n^{\beta-\alpha} \leq q^{(i)}_{[R_i],v}(w) \leq 8n^{\beta-\alpha}q^{(i)}_{[R_i],u}(w)$. Plugging in this bound, we obtain

$$\begin{aligned}
2k^3\mathbf{E}[X_{1,1}X_{2,1}] &\leq 16k^3n^{\beta-\alpha}\sum_{w\in W} q^{(i)}_{[R_i],u}(w)^3 \\
&\leq 16k^3n^{\beta-\alpha}\left[\sum_{w\in W} q^{(i)}_{[R_i],u}(w)^2\right]^{3/2} \qquad (l_2\text{- and } l_3\text{-norm inequalities}) \\
&= 16n^{\beta-\alpha}\left[k^2\sum_{w\in W} q^{(i)}_{[R_i],u}(w)^2\right]^{3/2}.
\end{aligned}$$

We can apply the bound $q^{(i)}_{[R_i],u}(w) \leq 8n^{\beta-\alpha}q^{(i)}_{[R_i],v}(w)$.

$$\begin{aligned}
2k^3\mathbf{E}[X_{1,1}X_{2,1}] &\leq 16n^{\beta-\alpha}\left[k^2\sum_{w\in W} q^{(i)}_{[R_i],u}(w)q^{(i)}_{[R_i],v}(w)\cdot 8n^{\beta-\alpha}\right]^{3/2} \\
&\leq 512n^{5(\beta-\alpha)/2}\left[k^2\sum_{w\in W} q^{(i)}_{[R_i],u}(w)q^{(i)}_{[R_i],v}(w)\right]^{3/2}.
\end{aligned} \tag{5.2}$$

Note that $k^2\sum_{w\in W} q^{(i)}_{[R_i],u}(w)q^{(i)}_{[R_i],v}(w)]^{3/2}$ is exactly $\mathbf{E}[X]$. We have previously bounded $\mathbf{E}[X] \geq (c^2/4)n^{8(\beta-\alpha)}$. Thus, $512n^{5(\beta-\alpha)/2} \leq \mathbf{E}[X]^{1/2}/(c/100)$. Applying the bounds in (5.2), we deduce that $2k^3\mathbf{E}[X_{1,1}X_{2,1}] \leq (\mathbf{E}[X]^{1/2}/(c/100))(\mathbf{E}[X]^{3/2}) = \mathbf{E}[X]^2/(c/100)$. We get an identical bound for $2k^3\mathbf{E}[X_{1,1}X_{1,2}]$. Putting this all together, we can prove that $\mathbf{var}[X] \leq 4\mathbf{E}[X]^2/c'$ for $c' = \Theta(c)$. An application of the Chebyshev inequality proves that $\Pr[X > 0] > 2/3$. ∎

5.2. The procedure `IdealFindBiClique`. We describe an "ideal" variant of `FindBiclique` below. It is not possible to directly implement `IdealFindBiClique`. On the other hand, we can directly analyze the probability that it produces a $K_{r,r}$-minor. We will eventually prove that the `FindBiclique` procedure can efficiently simulate `IdealFindBiClique`.

`IdealFindBiClique`(s)

1. For $i = 5r^2, \dots, 1/\delta + 4$:
   (a) Perform $2r$ independent draws from the distribution $\mathcal{D}_{s,i}$ (Definition 4.10). Let the first r draws be multiset A, and let the remaining draws be B.
   (b) For each $a \in A$, $b \in B$:
      i. Run `FindPath`$(a, b, n^{\delta(i+18)/2}, i)$
   (c) If all calls to `FindPath` return a path, then let the collection of paths be the subgraph F. Run `KKR`(F, H). If it returns an H-minor, output that and terminate.

The next lemma asserts that `IdealFindBiClique` finds (with nontrivial probability) paths between all pairs of vertices between two sets of r vertices. Ignoring the gnarly problem of the paths intersecting internally, we see that this structure looks like a $K_{r,r}$-minor. Lemma 5.2 gives bounds for finding a single path between one pair of vertices, whose truncated random walk vectors satisfy some technical conditions. Somewhat naively, we could hope to find two sets of vertices A, B where all pairs in $A \times B$ satisfy these conditions. Then, on applying Lemma 5.2 for each pair, we could find a subgraph that looks like a $K_{r,r}$-minor.

It turns out that if A and B are themselves generated by performing random walks from a fixed vertex, the probability conditions of Lemma 5.2 hold "on average" for the pairs in $A \times B$. This crucially uses the correlation lemma. The proof of the next lemma further shows that this average condition suffices to lower bound the success probability of `FindBiclique`. We can basically assume that each call of `FindPath` within `IdealFindBiClique` has an independent success probability.

LEMMA 5.3. *Suppose $s \in S_i$ for some $i \leq 1/\delta + 4$. With probability $(4n^{2\delta})^{-r^2}$, the calls to* `FindPath` *in* `IdealFindBiClique`(s) *output paths from every $a \in A$ to every $b \in B$, where each path is an R_i-returning walk of length $2^{i+1}\ell$.*

Proof. Recall that each element in $A \cup B$ is drawn according to $\hat{q}^{(i+1)}_{[R_i],s}(u)$. For any $a, b \in V$, let $\tau_{a,b}$ be the probability that `FindPath`$(a, b, n^{\delta(i+18)/2}, i)$ succeeds in finding an R_i-returning walk between a and b (of length $2^{i+1}\ell$). The probability of success for `FindBiclique`(s) conditioned on $A, B \subseteq R_i$ is at least

$$
\begin{aligned}
&\sum_{A \in R_i^r} \sum_{B \in R_i^r} \prod_{a \in A} \hat{q}^{(i+1)}_{[R_i],s}(a) \prod_{b \in B} \hat{q}^{(i+1)}_{[R_i],s}(b) \tau_{a,b} \\
&= \sum_{A \in R_i^r} \sum_{B \in R_i^r} \prod_{a \in A} \hat{q}^{(i+1)}_{[R_i],s}(a) \Big(\prod_{b \in B} \hat{q}^{(i+1)}_{[R_i],s}(b) \Big) \Big(\prod_{b \in B} \tau_{a,b} \Big) \\
&= \sum_{B \in R_i^r} \Big(\prod_{b \in B} \hat{q}^{(i+1)}_{[R_i],s}(b) \Big) \sum_{A \in R_i^r} \prod_{a \in A} \Big[\hat{q}^{(i+1)}_{[R_i],s}(a) \Big(\prod_{b \in B} \tau_{a,b} \Big) \Big] \\
&= \sum_{B \in R_i^r} \prod_{b \in B} \hat{q}^{(i+1)}_{[R_i],s}(b) \Big(\sum_{a \in R_i} \hat{q}^{(i+1)}_{[R_i],s}(a) \prod_{b \in B} \tau_{a,b} \Big)^r.
\end{aligned}
$$

Observe that $\prod_{b \in B} \hat{q}^{(i+1)}_{[R_i],s}(b)$ is a probability distribution over R_i^r. By Jensen's inequality, we lower bound.

$$\sum_{B\in R_i^r}\prod_{b\in B}\hat{q}^{(i+1)}_{[R_i],s}(b)\Big(\sum_{a\in R_i}\hat{q}^{(i+1)}_{[R_i],s}(a)\prod_{b\in B}\tau_{a,b}\Big)^r \tag{5.3}$$
$$\geq \Big[\sum_{B\in R_i^r}\Big(\prod_{b\in B}\hat{q}^{(i+1)}_{[R_i],s}(b)\Big)\sum_{a\in R_i}\hat{q}^{(i+1)}_{[R_i],s}(a)\prod_{b\in B}\tau_{a,b}\Big]^r.$$

We manipulate and expand the RHS of (5.3) further.

$$\begin{aligned}
\text{RHS of (5.3)} &= \Big[\sum_{a\in R_i}\sum_{B\in R_i^r}\hat{q}^{(i+1)}_{[R_i],s}(a)\Big(\prod_{b\in B}\hat{q}^{(i+1)}_{[R_i],s}(b)\Big)\Big(\prod_{b\in B}\tau_{a,b}\Big)\Big]^r\\
&= \Big[\sum_{a\in R_i}\hat{q}^{(i+1)}_{[R_i],s}(a)\sum_{B\in R_i^r}\prod_{b\in B}\hat{q}^{(i+1)}_{[R_i],s}(b)\tau_{a,b}\Big]^r\\
&= \Big[\sum_{a\in R_i}\hat{q}^{(i+1)}_{[R_i],s}(a)\big(\sum_{b\in R_i}\hat{q}^{(i+1)}_{[R_i],s}(b)\tau_{a,b}\big)^r\Big]^r\\
&\geq \Big[\sum_{a\in R_i}\sum_{b\in R_i}\hat{q}^{(i+1)}_{[R_i],s}(a)\hat{q}^{(i+1)}_{[R_i],s}(b)\tau_{a,b}\Big]^{r^2} \quad \text{(Jensen)}\\
&= \Big[\mathbf{E}_{a,b\sim\mathcal{D}_{s,i}}[\tau_{a,b}]\Big]^{r^2}.
\end{aligned}$$

Toward lower bounding $\tau_{a,b}$, we first lower bound $\boldsymbol{q}^{(i)}_{[R_i],a}\cdot\boldsymbol{q}^{(i)}_{[R_i],b}$. By Lemma 4.12, $\mathbf{E}_{a,b}[\boldsymbol{q}^{(i)}_{[R_i],a}\cdot\boldsymbol{q}^{(i)}_{[R_i],b}]\geq 1/n^{\delta(i+1)}$. Claim 4.6 states that $\|\boldsymbol{q}^{(i)}_{[R_i],a}\|_2^2\leq 1/n^{\delta(i-1)}$ and $\|\boldsymbol{q}^{(i)}_{[R_i],b}\|_2^2\leq 1/n^{\delta(i-1)}$. By Cauchy–Schwarz, $\boldsymbol{q}^{(i)}_{[R_i],a}\cdot\boldsymbol{q}^{(i)}_{[R_i],b}\leq 1/n^{\delta(i-1)}$. Let p be the probability (over a,b) that $\boldsymbol{q}^{(i)}_{[R_i],a}\cdot\boldsymbol{q}^{(i)}_{[R_i],b}\geq 1/2n^{\delta(i+1)}$.

$$1/n^{\delta(i+1)}\leq \mathbf{E}_{a,b}[\boldsymbol{q}^{(i)}_{[R_i],a}\cdot\boldsymbol{q}^{(i)}_{[R_i],b}]\leq (1-p)/2n^{\delta(i+1)}+p/n^{\delta(i-1)}.$$

Thus, $p\geq 1/2n^{2\delta}$.

By Claim 4.6, for every $a\in R_i$, $\|\boldsymbol{q}^{(i)}_{[R_i],a}\|_2^2\leq 1/n^{\delta(i-1)}$ (and similarly for $b\in R_i$). Suppose $\boldsymbol{q}^{(i)}_{[R_i],a}\cdot\boldsymbol{q}^{(i)}_{[R_i],b}\geq 1/2n^{\delta(i+1)}$. Let us apply Lemma 5.2, with $\alpha=\delta(i-1)$ and $\beta=\delta(i+1)$. The number of walks performed in calls made to `FindPath` by the `FindBiclique` procedure (the value of k) is $n^{\delta(i+18)/2}$. Note that $\delta(i+18)/2>\delta(i+1)/2+8\delta=\beta/2+4(\alpha-\beta)$. By Lemma 5.2, $\tau_{a,b}\geq 1/2$. As argued in the previous paragraph, this will happen with probability $1/2n^{2\delta}$ (over the choice of $a,b\sim\mathcal{D}_{s,i}$). We plug in (10) and deduce that the probability of success is at least $(1/4n^{2\delta})^{r^2}$. ◻

5.3. Criteria for `IdealFindBiClique` to reveal a minor. Fix $s\in S_i$ as in Lemma 5.3. This lemma only asserts that all pairs in $A\times B$ are connected by `IdealFindBiClique` (with nontrivial probability). We need to argue that these paths will actually induce a $K_{r,r}$-minor.

The overall proof is highly technical, and we provide Table 2 for reference.

As in Lemma 5.3, let us focus on the ith iteration within `IdealFindBiClique`. For every $a\in A, b\in B$, there is a call to `FindPath`$(a,b,n^{\delta(i+18)/2},i)$. Within each such call, a set of walks is performed from both a and b, with the hope of connecting a to b. We use a,a' (resp., b,b') to refer to elements in A (resp., B).

TABLE 2
Bad intersections notation.

Notation	Meaning	Where defined
$\boldsymbol{W}_a^b$	R_i-returning walks from a in $\texttt{FindPath}(a, b, \cdot, \cdot)$	Def. 5.4
$\boldsymbol{W}_a$	$\bigcup_{b\in B} \boldsymbol{W}_a^b$	Def. 5.4
$P_{a,b}$	A path from a to b discovered in FindPath	Def. 5.4
τ	Middle step of the walks in $\boldsymbol{W}_a$ or $\boldsymbol{W}_b$	Def. 5.5
$\boldsymbol{\sigma}_{s,S,t}$	Prob. vector of S-returning t-length walk from s	section 5.4, Def. 5.7

DEFINITION 5.4.

- *Let $\boldsymbol{W}_a^b$ denote the set of R_i-returning walks from a performed in the call to* $\texttt{FindPath}(a, b, n^{\delta(i+18)/2}, i)$.
- *Let $\boldsymbol{W}_a$ denote the set of all vertices in $\bigcup_{b\in B} \boldsymbol{W}_a^b$.*
- *Let $P_{a,b}$ be a single path from a to b discovered by* $\texttt{FindPath}(a, b, n^{\delta(i+18)/2}, i)$ *that consists of a walk in $\boldsymbol{W}_a^b$ and a walk $\boldsymbol{W}_b^a$ that end at the same vertex. If there are many possible such paths, pick the lexicographically smallest.*

We stress that walks in $\boldsymbol{W}_a^b$ do not necessarily end at b and come from a distribution independent of b (but we wish to track the specific call of FindPath where these walks were performed). Note that $\boldsymbol{W}_b^a$ is the set of R_i-returning walks starting from b performed in the same call.

Note that any of the paths/sets described above could be empty. We will think of paths as sequences, rather than sets, since the order in which the path is constructed is relevant. For any path, P, we use $P(t)$ to denote the tth element in the sequence. We use $P(\geq t)$ to denote the sequence of elements with index at least t. When we refer to intersections of paths being empty/nonempty, we mean sets induced by the corresponding sequences.

For $s \in S_i$, conditioned on $A, B \subseteq R_i$, Lemma 5.3 gives a lower bound on $\Pr[\text{for all } a \in A, b \in B, (P_{a,b} \neq \emptyset)]$. We will now define some *bad* events that interfere with minor structure.

Recall that A and B are multisets (it is convenient to think of them as sequences). The same vertex may appear multiple times in $A \cup B$, but we think of each occurrence as a distinct multiset element. Therefore, when we speak of equality of vertices, we mean vertices at the same index in A (or B). By definition, elements in A are disjoint from B.

DEFINITION 5.5. *The following events are referred to as* bad events *of Type* 1, 2, *or* 3. *We set $\tau = 2^{i-1}\ell$.*

1. *$\exists a, b, c \in A \cup B$, $c \notin \{a, b\}$, such that $\boldsymbol{W}_c \cap P_{a,b} \neq \emptyset$.*
2. *$\exists a, b, b'$ (all distinct) such that $\exists W \in \boldsymbol{W}_a^b$ where $W(\geq \tau) \cap P_{a,b'} \neq \emptyset$. (Or, $\exists a, a' \in A, b \in B$, all distinct, such that $\exists W \in \boldsymbol{W}_b^a$ where $W(\geq \tau) \cap P_{a',b} \neq \emptyset$.)*
3. *$\exists a, b, W_a \in \boldsymbol{W}_a^b, W_b \in \boldsymbol{W}_b^a$ such that W_a, W_b end at the same vertex, and $\exists t_1, t_2$ such that $\min(t_1, t_2) \leq \tau$ and $W_a(t_1) = W_b(t_2)$.*

For clarity, let us express the above bad events in nontechnical terms. Note that τ is the index of the midpoint of the walks, so it splits walks into halves.

1. A walk from $c \in A \cup B$ intersects $P_{a,b}$, where $c \neq a, b$.
2. The second half of a walk in $\boldsymbol{W}_a^b$, which starts from a, intersects $P_{a,b'}$ for $b \neq b'$. Or, the second half of a walk in $\boldsymbol{W}_b^a$, which starts at b, intersects $P_{a',b}$ for $a \neq a'$.
3. A walk in $\boldsymbol{W}_a^b$ and a walk in $\boldsymbol{W}_b^a$ intersect at least twice. Note that this is

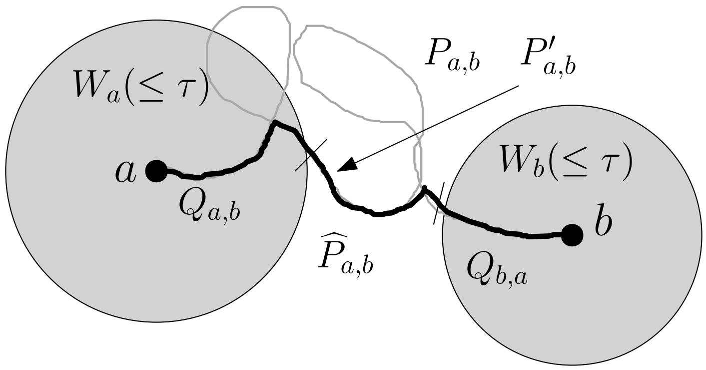


FIG. 1. *This figure shows the various subpaths defined in the proof of Claim* 5.6. *The simple path $P'_{a,b}$ (bold line) from a to b induced by $P_{a,b}$ (gray line). This path is broken into three contiguous subpaths: the portion "close to" a, the portion close to b, and the remainder.*

a pair of walks, one from a and the other from b. The first intersection is in the first half of either of these walks. The walks also end at the same vertex.

CLAIM 5.6. *If all $P_{a,b}$ sets are nonempty and there is no bad event, then $\bigcup_{a,b} P_{a,b}$ contains a $K_{r,r}$-minor.*

Proof. Each $P_{a,b}$ is formed by $W_a \in \boldsymbol{W}_a^b$ and $W_b \in \boldsymbol{W}_b^a$ that end at the same vertex. Since there is no Type 3 bad event, $W_a(\leq \tau)$ is disjoint from W_b (and vice versa). Since $P_{a,b}$ is not necessarily a simple path, let us remove all self-intersections to form a simple path $P'_{a,b}$. We construct three vertex disjoint contiguous subpaths of $P'_{a,b}$ (also refer to Figure 1).

1. $Q_{a,b}$ is the contiguous subpath of $P'_{a,b}$ contained in $W_a(\leq \tau)$. This is the "first half" of the walk from a that forms $P_{a,b}$.

2. $Q_{b,a}$ is the contiguous subpath of $P'_{a,b}$ contained in $W_b(\leq \tau)$. This is the "first half" of the walk from b that forms $P_{a,b}$.

3. $\widehat{P_{a,b}}$ is the contiguous subpath of $P'_{a,b}$ formed by removing vertices of $Q_{a,b} \cup Q_{b,a}$. Note that $\widehat{P_{a,b}} \subseteq W_a(> \tau) \cup W_b(> \tau)$.

In each bullet below, we state a disjointness condition on all the paths defined above. Each statement is followed by its corresponding proof. Subsequently, we will show how these disjointness conditions imply the existence of the desired minor.

We consider $a, a' \in A$ and $b, b' \in B$, where the elements in A (or B) might be equal. In what follows, when we say paths intersect, we mean that they share a common vertex. (Thus, disjoint paths always means vertex disjoint paths.) The idea is to show that whenever the disjointness condition fails, a bad intersection must occur.

- If $a \neq a'$, $Q_{a,b} \cap Q_{a',b'} = \emptyset$. If $b \neq b'$, $Q_{b,a} \cap Q_{b',a'} = \emptyset$.
  Consider the first statement. (Note that we allow $b = b'$.) Suppose $Q_{a,b} \cap Q_{a',b'} \neq \emptyset$. Observe that $Q_{a,b} \subseteq \boldsymbol{W}_a$ and $Q_{a',b'} \subseteq P_{a',b'}$. So $\boldsymbol{W}_a \cap P_{a',b'} \neq \emptyset$, implying a Type 1 bad event. This is a contradiction. The second statement follows analogously.
- $Q_{a,b} \cap Q_{b',a'} = \emptyset$.
  First, suppose that $a = a'$ and $b = b'$. As argued in the beginning of the proof, $Q_{a,b}$ and $Q_{b,a}$ are vertex disjoint. Now, consider the case $a \neq a'$. As before, $Q_{a,b} \subseteq \boldsymbol{W}_a$ and $Q_{b',a'} \subseteq P_{a',b'}$.

Since no Type 1 bad events occur, $\boldsymbol{W}_a \cap P_{a',b'} = \emptyset$. The case $b \neq b'$ is analogous.

- If $a \neq a'$ or $b \neq b'$, $\widehat{P_{a,b}} \cap P_{a',b'} = \emptyset$. For the sake of contradiction, suppose $\widehat{P_{a,b}} \cap P_{a',b'} \neq \emptyset$. We will show that some bad intersection occurs. Without loss of generality, assume $a \neq a'$ (and b may or not be the same as b'). Note that $\widehat{P_{a,b}} \subseteq W_a(>\tau) \cup W_b(>\tau)$, where $W_a \in \boldsymbol{W}^b_a$ and $W_b \in \boldsymbol{W}^a_b$. Either $W_a(>\tau)$ or $W_b(>\tau)$ intersects $P_{a',b'}$, leading to the following two cases:
  – If $W_a(>\tau) \cap P_{a',b'} \neq \emptyset$, then $\boldsymbol{W}_a \cap P_{a',b'} \neq \emptyset$. This is a Type 1 bad event.
  – If $W_b(>\tau) \cap P_{a',b'} \neq \emptyset$: If $b \neq b'$, then $\boldsymbol{W}_b \cap P_{a',b'} \neq \emptyset$. This is a Type 1 bad event. If $b = b'$, then we have $W_b(>\tau) \cap P_{a',b} \neq \emptyset$. Since $W_b \in \boldsymbol{W}^a_b$ (for $a \neq a'$), $\boldsymbol{W}_b \cap P_{a'}b \neq \emptyset$. This is a Type 2 bad event.

We construct the minor. Let $C(a) = \bigcup_{b \in B} Q_{a,b}$ and $C(b) = \bigcup_{a \in A} Q_{b,a}$. Each $C(a), C(b)$ forms a connected subgraph. By the disjointness properties of the $Q_{a,b}$ sets, all the $C(a), C(b)$ sets/subgraphs are vertex disjoint. Note that $\widehat{P_{a,b}}$ is disjoint from all other $P_{a',b'}$ paths *and* all the $C(a), C(b)$ sets. (We construct $\widehat{P}_{a,b}$ to be disjoint from $Q_{a,b}$ and $Q_{b,a}$ in the first paragraph. Every other $Q_{a',b'}$ is contained in $P_{a',b'}$.) Moreover, each $\widehat{P}_{a,b}$ has an edge to $C(a)$ and $C(b)$, since it is contained in $P_{a,b}$. Thus, we have disjoint paths from each $C(a)$ to $C(b)$, which gives a $K_{r,r}$-minor. ∎

5.4. The probabilities of bad events. In this section, we bound the probability of bad events, as detailed in Definition 5.5. As before, we fix $s \in S_i$.

We require some technical definitions of random walk probabilities.

DEFINITION 5.7. *Let $\sigma_{s,S,t}(v)$ be the probability of a walk from s to v of length t being S-returning. (We allow ℓ not to divide t and require that the walk encounters S at every $j\ell$th step for $j \leq \lfloor t/\ell \rfloor$.)*

We use $\boldsymbol{\sigma}_{s,S,t}$ to denote the vector of these probabilities. More generally, given any distribution vector $\boldsymbol{x}$ on V, $\boldsymbol{\sigma}_{\boldsymbol{x},S,t}$ denotes the vector of S-returning walk probabilities at time t.

We stress that this is not a conditional probability. Note that if $t = 2^i\ell$, then $\boldsymbol{\sigma}_{s,S,t} = \boldsymbol{q}^{(i)}_{[S],s}$. We show some simple propositions on these vectors. Let $\mathbb{I}_S$ denote the $n \times n$ matrix that preserves all coordinates in S and zeros out other coordinates.

PROPOSITION 5.8. *The vector $\boldsymbol{\sigma}_{\boldsymbol{x},S,t}$ evolves according to the following recurrence. First, $\boldsymbol{\sigma}_{\boldsymbol{x},S,0} = \boldsymbol{x}$. For $t \geq 1$ such that $\ell \nmid t$, $\boldsymbol{\sigma}_{\boldsymbol{x},S,t} = M\boldsymbol{\sigma}_{\boldsymbol{x},S,t-1}$. For $t \geq 1$ such that $\ell \mid t$, $\boldsymbol{\sigma}_{\boldsymbol{x},S,t} = \mathbb{I}_S M \boldsymbol{\sigma}_{\boldsymbol{x},S,t-1}$.*

PROPOSITION 5.9. *For all $\boldsymbol{x}$ and all $t \geq 1$, $\|\boldsymbol{\sigma}_{\boldsymbol{x},S,t}\|_\infty \leq \|\boldsymbol{\sigma}_{\boldsymbol{x},S,t-1}\|_\infty$.*

Proof. Since M is a symmetric random walk matrix, it computes the "new" value at a vertex by averaging the values of the neighbors (and itself). This can never increase the maximum value. Furthermore, $\mathbb{I}_S$ only zeros out some coordinates. This proves the proposition. ∎

In what follows, we fix the walk length to $2^i\ell$. To reduce clutter, we drop notational dependence on this length.

DEFINITION 5.10. *The distribution of $2^i\ell$-length walks from u is denoted $\mathcal{W}_u$. For any walk W, $W_u(t)$ denotes the tth vertex of the walk. The Boolean predicate $\rho(W_u)$ is true if W_u is R_i-returning.*

Recall that $\mathcal{D}_{s,i}$ is the distribution with support R_i, where the probability of $u \in R_i$ is $\hat{q}^{(i+1)}_{[R_i],s}(v)$ (Definition 4.10). This is the distribution that A, B are drawn from in `IdealFindBiClique`. We set i to be such that $s \in R_i$. Since i is fixed, we will simply write this as $\mathcal{D}_s$.

CLAIM 5.11. *For any $F \subseteq V$, we have the following:*

1.
$$\Pr_{a\sim\mathcal{D}_s, W_a\sim\mathcal{W}_a}[\rho(W_a) \wedge W_a \cap F \neq \emptyset] \leq 2^i\ell|F|/(n^{\delta(i-1)}\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1).$$

2. *For any $a \in R_i$,*
$$\Pr_{W_a\sim\mathcal{W}_a}[\exists t \geq \tau \mid \rho(W_a) \wedge W_a(t) \in F] \leq 2^i\ell|F|/n^{\delta(i-2)}.$$

Proof. We prove the first part. Let $\boldsymbol{x} = \boldsymbol{\sigma}_{s,R_i,i+1}$ be the probability vector corresponding to $\mathcal{D}_s$. So $\|\boldsymbol{x}\|_\infty = \|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_\infty/\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1$. We proceed with a union bound over F and the walk length and write

$$\Pr_{a\sim\mathcal{D}_s, W_a\sim\mathcal{W}_a}[\rho(W_a)\wedge W_a\cap F \neq \emptyset] \leq \sum_{t\leq 2^i\ell}\sum_{v\in F}\Pr_{a\sim\mathcal{D}_s, W_a\sim\mathcal{W}_a}[\rho(W_a)\wedge W_a(t) = v]. \tag{5.4}$$

Note that for any $v \in V$ and $t \leq 2^i\ell$, we can write

$$\begin{aligned}
\Pr_{a\sim\mathcal{D}_s, W_a\sim\mathcal{W}_a}[\rho(W_a) \wedge W_a(t) = v] &= \sum_{u\in R_i}\Pr_{a\sim\mathcal{D}_s}[a = u]\Pr_{W_a\sim\mathcal{W}_a}[\rho(W_a) \wedge W_a(t) = v | a = u] \\
&\overset{(\mathbf{1})}{\leq} \sum_{u\in R_i}\sigma_{s,R_i,2^{i+1}\ell}(u)\cdot\sigma_{u,R_i,t}(v) \\
&\overset{(\mathbf{2})}{=} \sigma_{s,R_i,2^{i+1}\ell+t}(v) \\
&\overset{(\mathbf{3})}{\leq} \|\boldsymbol{x}\|_\infty.
\end{aligned}$$

Inequality ($\mathbf{1}$) follows because we can upper bound the probability that a $2^i\ell$-length walk from a random a is R_i-returning *and* hits v at the tth step by the probability that just the t-step "prefix" of the walk from a that ends at v is R_i-returning. Equality ($\mathbf{2}$) is immediate from the definition of $\mathcal{D}_s$ and $\boldsymbol{\sigma}_{s,R_i,t}$. For ($\mathbf{3}$), note that by Proposition 5.9, for all $t \geq 1$, $\|\boldsymbol{\sigma}_{\boldsymbol{x},R_i,t}\|_\infty \leq \|\boldsymbol{x}\|_\infty$. Further, using Claim 4.7, this is at most $1/(n^{\delta(i-1)}\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1)$. Together with (5.4), this gives

$$\begin{aligned}
\Pr_{a\sim\mathcal{D}_s, W_a\sim\mathcal{W}_a}[\rho(W_a) \wedge W_a \cap F \neq \emptyset] &\leq \sum_{t\leq 2^i\ell}\sum_{v\in F}\|\boldsymbol{x}\|_\infty \\
&\leq 2^i\ell|F|/(n^{\delta(i-1)}\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1).
\end{aligned}$$

Now we prove the second part. By the union bound, the probability is bounded above by

$$\sum_{t\geq 2^{i-1}\ell}\sum_{u\in F}\Pr_{W_a\sim\mathcal{W}_a}[\rho(W_a) \wedge W_a(t) = u] \leq \sum_{t\geq 2^{i-1}\ell}\sum_{u\in F}\|\boldsymbol{\sigma}_{a,R_i,t}\|_\infty. \tag{5.5}$$

By Proposition 5.9, the infinity norm is bounded above by $\|\boldsymbol{\sigma}_{a,R_i,2^{i-1}\ell}\|_\infty = \|\boldsymbol{q}^{(i-1)}_{[R_i],a}\|_\infty$. By Claim 4.7, the latter is at most $1/n^{\delta(i-2)}$. Plugging in (5.5), we get an upper bound of $2^{i-1}\ell|F|/n^{\delta(i-2)}$. ∎

CLAIM 5.12. *For any $a \in R_i$, we have*

$$\Pr\left[\rho(W_a) \wedge \rho(W_b) \wedge W_a(2^i\ell) = W_b(2^i\ell) \wedge \right. \\ \left. \exists t_a, t_b, \min(t_a, t_b) \leq \tau, W_a(t_a) = W_b(t_b)\right] \qquad \leq \frac{2^{2i}\ell^2}{(n^{\delta(2i-2)}\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1)}.$$

Here, the probability is taken over the following joint distribution: $b \sim \mathcal{D}_s, W_a \sim \mathcal{W}_a, W_b \sim \mathcal{W}_b$.

Proof. We write the main event in nontechnical terms. We fix an arbitrary a and pick $b \sim \mathcal{D}_s$. We perform walks of length $2^i\ell$ from both a and b. We are bounding the probabilities that these walks are R_i-returning and that the "initial half" (less than $2^{i-1}\ell$ steps) of one of the walks intersects with the other. Subsequently, both walks end at the same vertex (and both of these walks happen to be R_i-returning).

To this end, let us define two vertices w_1, w_2. We want to bound the probability that both walks first encounter w_1 and then end at w_2. It will be very useful to treat the latter part simply as two walks from w_1, where one of them is at least of length $2^{i-1}\ell$. Note that w_1 might not be in R_i.

Let $Z_{a,t}$ be the random variable denoting the tth vertex of a random walk from a. Let us also define R_i-returning walks with an offset g starting from w_1. Basically, such a walk starts from w (that may not be in R_i) and performs g steps to end up in R_i. Subsequently, it behaves as an R_i-returning walk. Observe that the second parts of the walks are R_i-returning walks from w_1, with offsets of $\ell - [t_a(\bmod \ell)]$, $\ell - [t_b(\bmod \ell)]$. Let $Y_{w,t}$ be the random variable denoting the tth vertex of an R_i-returning walk from w, with the offset $\ell - [t(\bmod \ell)]$. We use primed notation (e.g., something like $Y'_{w,t}$) for an independent copy of such variables.

Let us fix values for t_a, t_b such that $\min(t_a, t_b) \leq \tau = 2^{i-1}\ell$. (We will eventually union bound over all such values.) The probability we wish to bound is the following. We use independence of the walks to split the probabilities. There are four independent walks under consideration: one from a, one from b, and two from w.

$$\sum_{w_1 \in V}\sum_{w_2 \in V} \Pr_{\substack{b \sim \mathcal{D}_s \\ \mathcal{W}_a, \mathcal{W}_b, \mathcal{W}_{w_1}}}[Z_{a,t_a} = w_1 \wedge Z_{b,t_b} = w_1 \wedge Y_{w_1,2^i\ell - t_a} = w_2 \wedge Y'_{w_1,2^i\ell-t_b} = w_2]$$
$$= \sum_{w_1 \in V}\sum_{w_2 \in V} \Pr_{\mathcal{W}_a}[Z_{a,t_a} = w_1] \Pr_{\substack{b \sim \mathcal{D}_s \\ \mathcal{W}_b}}[Z_{b,t_b} = w_1] \Pr_{\mathcal{W}_{w_1}}[Y_{w_1,2^i\ell-t_a} = w_2] \Pr_{\mathcal{W}_{w_1}}[Y_{w_1,2^i\ell-t_b} = w_2]. \tag{5.6}$$

Consider $\Pr_{b\sim\mathcal{D}_s,\mathcal{W}_b}[Z_{b,t_b} = w_1]$. This is exactly the w_1th entry in $\boldsymbol{\sigma}_{\boldsymbol{x},R_i,t_b}$ where $\boldsymbol{x}$ is the distribution given by $\mathcal{D}_s$. By Proposition 5.9, this is at most $\|\boldsymbol{x}\|_\infty$, which is at most $1/(n^{\delta(i-1)}\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1)$ (as argued in the second part of the proof of Claim 5.11).

Since $\min(t_a, t_b) \leq \tau$, at least one of $2^i\ell - t_a$ or $2^i\ell - t_b$ is at least $2^{i-1}\ell$. Thus, one of $\Pr_{\mathcal{W}_{w_1}}[Y_{w_1,2^i\ell-t_a} = w_2]$ or $\Pr_{\mathcal{W}_{w_1}}[Y_{w_1,2^i\ell-t_b} = w_2]$ refers to a walk of length at least $2^{i-1}\ell$. Let us bound $\Pr_{\mathcal{W}_{w_1}}[Y_{w_1,t} = w_2]$ for $t \geq 2^i\ell$. We can break such a walk into two parts: the first $\ell - [t(\bmod \ell)]$ steps lead to some $v \in R_i$, and the second part is an R_i-returning walk of length at least $2^i\ell$ from v to w. Recall that $p_{x,d}(y)$ is the standard random walk probability of starting from x and ending at y after d steps.

For some $t' \geq 2^i \ell$,

$$\begin{aligned}\Pr_{\mathcal{W}_{w_1}}[Y_{w_1,t} = w_2] &= \sum_{v \in R_i} p_{w_1,\ell-[t (\bmod \ell)]}(v)\sigma_{v,R_i,t'}(w_2) \\ &\leq \sum_{v \in R_i} p_{w_1,\ell-[t (\bmod \ell)]}(v)\|\boldsymbol{\sigma}_{v,R_i,t'}\|_\infty \\ &\leq \sum_{v \in R_i} p_{w_1,\ell-[t (\bmod \ell)]}(v)\|\boldsymbol{q}^{(i)}_{[R_i],v}\|_\infty \\ &\leq \sum_{v \in R_i} p_{w_1,\ell-[t (\bmod \ell)]}(v) n^{-\delta(i-1)} \\ &= n^{-\delta(i-1)}.\end{aligned}$$

Plugging these bounds into (5.6), for fixed t_a, t_b, we see that there exists $t \in \{2^i\ell - t_a, 2^i\ell - t_b\}$ such that the probability of the main event is at most

$$\begin{aligned}&\frac{1}{n^{\delta(i-1)}\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1} \cdot \frac{1}{n^{\delta(i-1)}} \sum_{w_1 \in V} \sum_{w_2 \in V} \Pr_{\mathcal{W}_a}[Z_{a,t_a} = w_1] \Pr_{\mathcal{W}_{w_1}}[Y_{w_1,t} = w_2] \\ &\leq \frac{1}{n^{\delta(2i-2)}\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1} \sum_{w_1 \in V} \Pr_{\mathcal{W}_a}[Z_{a,t_a} = w_1] \sum_{w_2 \in V} \Pr_{\mathcal{W}_{w_1}}[Y_{w_1,t} = w_2] = \frac{1}{n^{\delta(2i-2)}\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1}.\end{aligned}$$

A union bound over all pairs of t_a, t_b completes the proof. ∎

We now bound the total probability of bad events. Most of the technical work has already been done in the previous lemmas; we only need to perform some union bounds.

LEMMA 5.13. *Total probability of bad events (in a call to* `IdealFindBiClique`(s)*) is at most*

$$\frac{2^{2i+4} r^4 n^{30\delta}}{n^{\delta i/2}\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1}. \tag{5.7}$$

Proof. We bound the bad events by type. Recall that $\ell = n^{5\delta}$.

Type 1: $\exists a, b, c \in A \cup B$, $c \neq a, b$, such that $\boldsymbol{W}_c \cap P_{a,b} \neq \emptyset$.

Fix a choice of $a \in A, b \in B$. Any $c \neq a, b$ is drawn from $\mathcal{D}_s$. In Claim 5.11, set $F = P_{a,b}$. By the first part of Claim 5.11, the probability that a single walk drawn from $\mathcal{W}_c$ is R_i-returning and intersects $P_{a,b}$ is at most $2^i\ell(2^{i+1}\ell)/n^{\delta(i-1)}\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1$. The set $\boldsymbol{W}_c$ consists of at most $rn^{\delta(i+18)/2}$ such walks. We union bound over all these walks, and all r^2 choices of a, b, and plug in $\ell = n^{5\delta}$ to get an upper bound of

$$\frac{2^{2i+1}\ell^2 r^3 n^{\delta(i+18)/2}}{n^{\delta(i-1)}\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1} = \frac{2^{2i+1} r^3 n^{20\delta}}{n^{\delta i/2}\|\boldsymbol{q}^{(i+1)}_{[R_i],s}\|_1}.$$

Type 2: $\exists a, b, b'$ (all distinct) such that $\exists W \in \boldsymbol{W}^b_a$ where $W(\geq \tau) \cap P_{a,b'} \neq \emptyset$. (Or, $\exists a, a' \in A, b \in B$ with analogous conditions.)

Fix a, b, b'. Set $F = P_{a,b'}$ in Claim 5.11. By the second part of Claim 5.11, the probability that a single walk from $\mathcal{W}_a$ is R_i-returning and intersects F at step $\geq \tau$ is at most $2^i\ell(2^{i+1}\ell)/n^{\delta(i-2)}$. We union bound over all the $rn^{\delta(i+18)/2}$ walks in $\boldsymbol{W}_a$ and all r^3 choices of a, b, b'. (We also union bound over choosing b, b' or a, a'.) The upper bound is $2^{2i+1} r^3 n^{21\delta}/n^{\delta i/2}$.

Type 3: $\exists a, b, W_a \in \boldsymbol{W}_a^b, W_b \in \boldsymbol{W}_b^a$ such that W_a, W_b end at the same vertex, and $\exists t_1, t_2$ such that $\min(t_1, t_2) \leq \tau$ and $W_a(t_1) = W_b(t_2)$.

This case is qualitatively different. We will take a union bound over *pairs* of walks and require the stronger bound of Claim 5.12.

Fix $a \in A$. Observe that $b \sim \mathcal{D}_s$. For a single walk $W_a \sim \mathcal{W}_a$ and a single walk $W_b \sim \mathcal{W}_b$, the probability of a Type 3 bad event is bounded by Claim 5.12. The upper bound is $2^{2i}\ell^2/(n^{\delta(2i-2)}\|\boldsymbol{q}_{[R_i],s}^{(i+1)}\|_1)$. We union bound over the $r^2 n^{\delta(i+18)}$ *pairs* of walks from a and b and then over the r^2 choices of a, b. The final bound is

$$\frac{2^{2i} r^4 \ell^2 \ n^{\delta(i+18)}}{n^{\delta(2i-2)}\|\boldsymbol{q}_{[R_i],s}^{(i+1)}\|_1} = \frac{2^{2i} r^4 n^{30\delta}}{n^{\delta i}\|\boldsymbol{q}_{[R_i],s}^{(i+1)}\|_1}.$$

We complete the proof by taking a union bound over the three types. Note that $\|\boldsymbol{q}_{[R_i],s}^{(i+1)}\|_1 \leq 1$, so we can upper bound the probability of each type of bad event by

$$\frac{2^{2i+1} r^4 n^{30\delta}}{n^{\delta i/2}\|\boldsymbol{q}_{[R_i],s}^{(i+1)}\|_1}.$$

5.5. Proof of Theorem 5.1.

Proof. Fix $s \in S_i$. We consider a run of the `FindBiclique`(s) procedure and do not consider the procedure `IdealFindBiClique`(s). Nonetheless, we can argue that the former efficiently simulates the latter. Let $\mathcal{C}$ be the event that the multisets A and B generated in `FindBiclique`(s) come from $2r$ independent draws from $\mathcal{D}_{s,i}$. In this case, `FindBiclique`(s) behaves exactly like `IdealFindBiClique`(s).

Let $\mathcal{E}$ be the event $\bigcap_{a \in A, b \in B} P_{a,b} \neq \emptyset$, and let $\mathcal{F}$ be the union of bad events. By Claim 5.6, the probability that `FindBiclique`(s) finds a minor is at least $\Pr[\mathcal{E} \cap \overline{\mathcal{F}}]$. We lower bound as follows: $\Pr[\mathcal{E} \cap \overline{\mathcal{F}}] \geq \Pr[\mathcal{C}] \Pr[\mathcal{E} \cap \overline{\mathcal{F}} | \mathcal{C}] \geq \Pr[\mathcal{C}] \ (\Pr[\mathcal{E}|\mathcal{C}] - \Pr[\mathcal{F}|\mathcal{C}])$.

The probability of a single random walk from s of length $2^i \ell$ being R_i-returning is $\|\boldsymbol{q}_{[R_i],s}^{(i+1)}\|_1$. Thus, $\Pr[\mathcal{C}] = \|\boldsymbol{q}_{[R_i],s}^{(i+1)}\|_1^{2r}$. By Claim 4.8, $\|\boldsymbol{q}_{[R_i],s}^{(i+1)}\|_1 \geq n^{-\delta}$, so $\Pr[\mathcal{C}] \geq n^{-2\delta r}$.

Lemma 5.3 provides a lower bound for $\Pr[\mathcal{E}|\mathcal{C}]$, and Lemma 5.13 provides an upper bound for $\Pr[\mathcal{F}|\mathcal{C}]$. We plug in these bounds below:

$$\Pr[\mathcal{E}|\mathcal{C}] - \Pr[\mathcal{F}|\mathcal{C}] \geq \frac{1}{(4n^{2\delta})^{r^2}} - \frac{2^{2i+4} r^4 n^{30\delta}}{n^{\delta i/2}\|\boldsymbol{q}_{[R_i],s}^{(i+1)}\|_1}. \tag{5.8}$$

Observe how the positive term is independent of i, while the negative term decays exponentially in i. This is crucial for arguing that for a sufficiently large (constant) i, the lower bound is nontrivial.

When $i \geq 5r^2$, $n^{i\delta/2} \geq n^{2\delta r^4 + \delta r^4/2} \geq n^{2\delta r^2 + 40\delta}$ (note that r, the number of vertices in H, is at least 3). By Claim 4.8, $\|\boldsymbol{q}_{[R_i],s}^{(i+1)}\|_1 \geq n^{-\delta}$. Thus, for sufficiently large n, $\Pr[\mathcal{F}|\mathcal{C}] \leq 1/(2(4n^{2\delta})^{r^2})$. Putting this all together, we see that the probability of finding a $K_{r,r}$-minor is at least $n^{-4\delta r^2}$. $\square$

6. Local partitioning in the trapped case. Theorem 5.1 tells us that if there are $\Omega(n^{1-\delta})$ vertices in strata numbered $5r^2$ and above, then `FindMinor` finds a biclique minor with high probability. We deal with the case when most vertices lie in low strata, i.e., random walks from most vertices are trapped in a very small subset.

We will argue that (almost) all vertices in low strata can be partitioned into "pieces" $P_1, P_2, \ldots, P_b$ such that each piece is a low conductance cut that is "easily discoverable." We mean that a superset of each piece $P_i, i \in [b]$, can be found by performing short random walks in G. If `FindMinor` fails to find a minor, this lemma can be iteratively applied to make G H-minor-free by removing few edges (this argument is given in section 7).

We use $p_{s,t}(v)$ to denote the probability that a t-length random walk from s ends at v.

LEMMA 6.1. *Let $i \leq 5r^2$ and $\delta < 1/20r^2$. Let $\alpha \geq n^{-\delta/2}$. Consider some subset $S \subseteq V$ and $i \in \mathbb{N}$ such that for all $s \in S, \|\boldsymbol{q}^{(i)}_{[S],s}\|_2^2 \leq 1/n^{\delta(i-1)}$. Define $S' \subseteq S$ to be $\{s \in S \mid \|\boldsymbol{q}^{(i+1)}_{[S],s}\|_2^2 \geq 1/n^{\delta i}\}$.*

Suppose $|S'| \geq \alpha n$. Then, there is a subset $\widetilde{S} \subseteq S'$, $|\widetilde{S}| \geq \alpha n/8$, such that for all $s \in \widetilde{S}$, there exists a subset $P_s \subseteq S$ where

- *(low conductance cut) $|E(P_s, S \setminus P_s)| \leq 2n^{-\delta/4} d|P_s|$,*
- *(easily discoverable) for all $v \in P_s$, $\exists t \leq 160n^{\delta(i+7)}/\alpha$ such that $p_{s,t}(v) \geq \alpha/n^{\delta(2i+14)}$.*

6.1. Proof overview. The rich literature on local partitioning gives tools for proving variants of the following statement. If there is a vertex s such that short random walks from s do not "spread much," then there exists a low conductance cut "around s" that can be discovered by performing random walks from s. The notion of spreading is measured by the l_2-norm of the random walk distribution.

In Lemma 6.1, if we set S to be V, then the lemma is exactly local partitioning. Our aim is to iteratively apply this lemma to partition the entire graph. Suppose, starting with the entire graph, we found low conductances pieces $P_1, P_2, \ldots$. Our aim is to find the next low conductance piece in $S = V \setminus \bigcup_j P_j$. If we could perform a random walk restricted to S, then we could simply apply existing local partitioning results. Therein lies the main challenge. For our sublinear application, we need to discover the pieces by doing random walks in the *original graph.* (Let us take a detour to understand why. We are analyzing the scenario where most vertices lie in low strata, which are defined by random walk norms in the original graph. We wish to argue that the graph can be partitioned into small pieces of low conductance, and we only assume strata conditions.)

Why is this hard? Note that we need to successfully partition at least $(1-\varepsilon)n$ vertices. Thus, we will have situations where $|S| = \Theta(\varepsilon n)$. In this case, random walks from vertices in S will leave S with high probability. The strata conditions only give us norm guarantees on such walks (that leave S), yet we need to locally partition completely within S. There is no clear correspondence between random walks in the original graph and random walks restricted to S.

Our main tool for addressing this conundrum is the *projected Markov chain M_S*. Consider the Markov chain with vertex set S that contains transitions for all walks in G that start and end at S but have all internal vertices outside S. Thus, M_S captures all walks in G but focuses on only the vertices in S. Our aim is to perform local partitioning on M_S.

Now for the main technical problem. Lemma 6.1 requires us to partition around a vertex s with large $\|\boldsymbol{q}^{(i+1)}_{[S],s}\|_2$ value. We need to get a lower bound for a walk norm, where walks are done in M_S. For any reasonable application of local partitioning in M_S, we need some guarantee on walks (in M_S) of length at least, say, n^δ (or something superconstant). Note that $\boldsymbol{q}^{(i+1)}_{[S],s}$ is looking at walks of length $2^{i+1}\ell$ that return 2^{i+1} times to S. These walks clearly correspond to walks in M_S. But since i is set to a constant, the lower bound on $\|\boldsymbol{q}^{(i+1)}_{[S],s}\|_2$ only seems to give guarantees on constant length walks in M_S.

We need to argue that, at least on average, walks of length $2^{i+1}\ell$ correspond to sufficiently long walks in M_S. This is precisely what happens in the proof of Lemma 6.5. This lemma asserts that for most vertices in S' (as defined in Lemma 6.1), random walks in M_S of length n^δ induce a fairly large norm (of the corresponding probability vector). An important tool for proving this lemma is a classic result in random walks, called *Kac's lemma*, that bounds average return times to sets of vertices.

In section 6.2, we give basic properties of the projected Markov chain which we build upon to prove Lemma 6.5. With Lemma 6.5 at hand, we can bring in the machinery of local partitioning to find a low conductance cut. We specifically use the Lovász–Simonovits curve technique to perform local partitioning on M_S [30]. Our analysis closely follows the notation and methods used by Spielman and Teng [34].

6.2. Projected Markov chain. We define the "projection" of the random walk onto the set S. This uses a construction of [22]. 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.

Since G is irreducible and the stationary mass on S is nonzero, all walks eventually reach S. Thus the outgoing transition probabilities from each v in M_S sum to 1, and hence M_S is a valid Markov chain. Furthermore, by the symmetry of the original random walk, $e^{(t)}_{u,v} = e^{(t)}_{v,u}$. Therefore the transition matrix of M_S remains symmetric, and the stationary distribution is uniform on S.

For a transition $e^{(t)}_{u,v}$ in M_S, we define the length of this transition to be t. For clarity, we use "hops" to denote the length 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 note that these ideas come from the work of Kale, Peres, and Seshadhri for analyzing random walks in noisy expanders [22].

We use $\boldsymbol{\tau}_{s,h}$ to denote the distribution of the h-hop walk from s and use $\tau_{s,h}(v)$ 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 in S.

We state Kac's formula (Corollary 24 in Chapter 2 of [1] restated).

LEMMA 6.2 (Kac's formula). *The expected return time (in G) to S of a random walk starting from S is the reciprocal of the fractional stationary mass of S, i.e., $n/|S|$.*

The following is a direct corollary.

LEMMA 6.3. $\mathbf{E}_{W\sim\mathcal{W}_h}[\textit{length of } W] = hn/|S|$.

Proof. Since the walk starts at the stationary distribution, it remains in this distribution at all hops. By linearity of expectation, it suffices to get the expected length for the first hop (and multiply with h). This is precisely the expected return time to S if we perform random walks in G. By Kac's formula above, the expected return time to S equals the reciprocal of the stationary mass of S, which is just $n/|S|$. ∎

6.2.1. Random walks in M_S do not spread. We begin with an important warmup. Using norm bounds in the premise of Lemma 6.1, we show that for every vertex $s \in S'$, there is a large set of destination vertices that are all reached with high probability through random walks of length $2^{i+1}\ell$.

CLAIM 6.4. *For every $s \in S'$, there exists a set $U_s \subseteq S$, $|U_s| \geq n^{\delta(i-2)}/2$, such that $\forall u \in U_s$, $p_{s,2^{i+1}\ell}(u) \geq 1/2n^{\delta i}$.*

Proof. By Proposition 4.2, for any $u \in S$, $q^{(i+1)}_{[S],s}(u) = \boldsymbol{q}^{(i)}_{[S],s} \cdot \boldsymbol{q}^{(i)}_{[S],u}$. By the property of S and Cauchy–Schwarz, $q^{(i+1)}_{[S],s}(u) \leq 1/n^{\delta(i-1)}$.

Since $s \in S'$, $\sum_{u\in S} q^{(i+1)}_{[S],s}(u)^2 \geq 1/n^{\delta i}$. Let us simply define U_s to be $\{u|u \in S, q^{(i+1)}_{[S],s}(u) \geq 1/(2n^{\delta i})\}$. Note that $p_{s,2^{i+1}\ell}(u) \geq q^{(i+1)}_{[S],s}(u)$.

$$\begin{aligned}1/n^{\delta i} \leq \sum_{u\in S} q^{(i+1)}_{[S],s}(u)^2 &= \sum_{u\in U_s} q^{(i+1)}_{[S],s}(u)^2 + \sum_{u\notin U_s} q^{(i+1)}_{[S],s}(u)^2\\ &\leq |U_s|/n^{2\delta(i-1)} + (1/2n^{\delta i}) \sum_{u\notin U_s} q^{(i+1)}_{[S],s}(u)\\ &\leq |U_s|/n^{2\delta(i-1)} + 1/2n^{\delta i}.\end{aligned}$$

We rearrange to bound the size of U_s. ◻

The next lemma is an analogue of Claim 6.4 for $\boldsymbol{\tau}$ vectors (that is, for M_S). Recall that $\ell = n^{5\delta}$.

LEMMA 6.5. *There exists a subset $S'' \subseteq S'$, $|S''| \geq |S'|/2$, such that for all $s \in S''$, $\|\boldsymbol{\tau}_{s,n^\delta}\|_\infty \geq 1/n^{\delta(i+6)}$.*

Proof. Define event $\mathcal{E}_{s,v,h}$ as follows. The event $\mathcal{E}_{s,v,h}$ occurs when an h-hop random walk from s has length $2^{i+1}\ell$ and ends at v. Observe that $p_{s,2^{i+1}\ell}(v) = \sum_{h\leq 2^{i+1}\ell} \Pr[\mathcal{E}_{s,v,h}]$ (because the number of hops is always at most the length). Since $\boldsymbol{\tau}_{s,h}$ is a random walk vector in a symmetric Markov chain, the infinity norm is nonincreasing in h. Thus, it suffices to find a subset $S'' \subseteq S'$, $|S''| \geq |S'|/2$ such that for all $s \in S''$, $\exists v \in S, h \geq n^\delta$, $\Pr[\mathcal{E}_{s,v,h}] \geq 1/n^{\delta(i+6)}$.

We define U_s as given in Claim 6.4. For all $v \in U_s$, by Claim 6.4, $p_{s,2^{i+1}\ell}(v) \geq 1/2n^{\delta i}$. Therefore, for all $v \in U_s$,

$$\sum_{h\leq 2^{i+1}\ell} \Pr[\mathcal{E}_{s,v,h}] \geq 1/2n^{\delta i}. \tag{6.1}$$

We will construct S'' by finding s where for some $v \in U_s$, $\sum_{h\leq n^\delta} \Pr[\mathcal{E}_{s,v,h}]$ is sufficiently small.

For any h,

$$\frac{1}{|S|}\sum_{s\in S'}\sum_{v\in U_s} \Pr[\mathcal{E}_{s,v,h}](2^{i+1}\ell) \leq \mathbf{E}_{W\sim\mathcal{W}_h}[\text{length of } W] = hn/|S|.$$

Suppose $h \leq 2^{i+1}\ell/n^{4\delta}$. (This is true for all $h \leq n^\delta$.) Then $\sum_{s\in S'}\sum_{v\in U_s}\Pr[\mathcal{E}_{s,v,h}] \leq n^{1-4\delta}$, and $\sum_{h\leq n^\delta}\sum_{s\in S'}\sum_{v\in U_s}\Pr[\mathcal{E}_{s,v,h}] \leq n^{1-3\delta}$.

We rearrange to get

$$\sum_{s\in S'}\sum_{v\in U_s}\sum_{h\leq n^\delta} \Pr[\mathcal{E}_{s,v,h}] \leq n^{1-3\delta}.$$

By the Markov bound, there is a set $S'' \subseteq S'$, $|S''| \geq |S'|/2$, such that for all $s \in S''$, $\sum_{v \in U_s} \sum_{h \leq n^\delta} \Pr[\mathcal{E}_{s,v,h}] \leq 2n^{1-3\delta}/|S'|$. By averaging, for all $s \in S''$, $\exists v \in U_s$, such that $\sum_{h \leq n^\delta} \Pr[\mathcal{E}_{s,v,h}] \leq 2n^{1-3\delta}/(|S'| \cdot |U_s|)$. By the assumptions of Lemma 6.1, $|S'| \geq \alpha n \geq n^{1-\delta/2}$. Claim 6.4 bounds $|U_s| \geq n^{\delta(i-2)}/2$. Plugging these in, we get

$$\sum_{h \leq n^\delta} \Pr[\mathcal{E}_{s,v,h}] \leq \frac{2n^{1-3\delta}}{n^{1-\delta/2} n^{\delta(i-2)}/2} \leq \frac{4}{n^{\delta(i+1/2)}}.$$

Subtracting this bound from (6.1), $\sum_{h \in [n^\delta, 2^{i+1}\ell]} \Pr[\mathcal{E}_{s,v,h}] \geq 1/4n^{\delta i}$. By averaging, for some $h \in [n^\delta, 2^{i+1}\ell]$, $\Pr[\mathcal{E}_{s,v,h}] \geq 1/(2^{i+3} n^{\delta i} \ell) \geq 1/n^{\delta(i+6)}$. This completes the proof. ∎

6.2.2. Local partitioning on M_S: Obtaining the low conductance cut. We perform local partitioning on M_S, starting with arbitrary $s \in S''$. We apply the Lovász–Simonovits curve technique. (The definitions are originally from [30]. Refer to Lecture 7 of Spielman's notes [33] as well as to section 2 in Spielman and Teng [34].) Before getting into the series of definitions, we give an overview of this technique.

The main idea is to represent the distribution at time t through a one-dimensional function h_t. This function is best thought of as a 2D-plot. Define $h_t(k)$ to be the sum of the k largest probabilities at time t. Note that $h_t(|S|)$ is simply 1 (the sum of all probabilities). Alternately, we sort the probabilities at time t in decreasing order and consider all prefix sums. These prefix sums are "plotted" by the function h_t. We can linearly interpolate between these values to get a piecewise linear curve, which is the function h_t.

Crucially, this curve is a *concave* function, since we take prefix sums of a nonincreasing list. As t increases, the curve h_t flattens out into a straight line from $(0,0)$ to $(|S|,1)$ (which represents the stationary distribution). The rate of flattening is the rate of convergence. The remarkable insight of Lovász and Simonovits is a relation between the flattening rate, the curve structure, and the conductance. They showed that to upper bound $h_{t+1}(x)$, one can draw a chord between $(x', h_t(x'))$ and $(x'', h_t(x''))$. Here $x' < x < x''$, and these points x', x'' depend on the conductances.

Now we give the technical setup.

- 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 identity.
- The LS curve h_t: We define a function $h_t : [0, |S|] \to [0,1]$ as follows. For every $k \in [|S|]$, set $h_t(k) = \sum_{j \leq k} [\tau_{s,t}(v_j^{(t)}) - 1/|S|]$. (Set $h_t(0) = 0$.) For every $x \in (k, k+1)$, we linearly interpolate to construct $h_t(x)$. Alternately, $h_t(x) = \max_{\vec{w} \in [0,1]^{|S|}, \|\vec{w}\|_1 = x} \sum_{v \in S} [\tau_{s,t}(v) - 1/|S|] w_i$.
- 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 $\tau_{s,t}(v_k^{(t)})$.
- Conductance: for some $T \subseteq S$ we define the conductance of T in M_S to be
$$\Phi(T) = \frac{\sum_{\substack{u \in T \\ v \in S \setminus T}} \tau_{u,1}(v)}{\min(|T|, |S \setminus T|)}.$$

The main lemma of Lovász and Simonovits is the following (Lemma 1.4 of [30]).

LEMMA 6.6. *For all k and all t,*

$$h_t(k) \leq \frac{1}{2}[h_{t-1}(k - 2\min(k, |S| - k)\Phi(L_{k,t})) + h_{t-1}(k + 2\min(k, |S| - k)\Phi(L_{k,t}))].$$

The typical purpose of the Lovász–Simonovits technique is to argue about rapid mixing when all conductances (or conductances of sufficiently large sets) are lower bounded. We consider a scenario in which only sets with minimum probability at least (say) p have high conductance. In this case, we can guarantee that the largest probability will converge to p.

LEMMA 6.7. *Suppose there exists $\phi \in [0,1]$ and $p > 2/|S|$ such that for all $t' \leq t$ and all $k \in [n]$ the following implication is true: if $L_{k,t'}$ has a minimum probability of at least p, then $\Phi(L_{k,t}) \geq \phi$. Then for all $k \in [0, |S|]$, $h_t(k) \leq \sqrt{k}(1 - \phi^2/2)^t + pk$.*

Proof. We will prove by induction over t. If $k \geq 1/p$, then the RHS is at least 1. Thus the bound is trivially true. Let us assume that $k < 1/p < |S|/2$. We split 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 6.6, $h_t(k) \leq (1/2)[h_{t-1}(k - 2\min(k, |S| - k)\Phi(L_{k,t})) + h_{t-1}(k + 2\min(k, |S| - k)\Phi(L_{k,t}))]$. Since $k < |S|/2$, $\min(k, |S| - k) = k$. By concavity of h_k, we can replace $\Phi(L_{k,t})$ in the above inequality by any lower bound. (We can always upper bound by a chord "above" the given chord.) Thus,

$$\begin{aligned}
h_t(k) &\leq \frac{1}{2}\Big(h_{t-1}\big(k(1-2\phi)\big) + h_{t-1}\big(k(1+2\phi)\big)\Big) \\
&\overset{(\mathbf{1})}{\leq} \frac{1}{2}\Big(\sqrt{k(1-2\phi)}(1-\phi^2/2)^{t-1} + \sqrt{k(1+2\phi)}(1-\phi^2/2)^{t-1} + 2kp\Big) \\
&\overset{(\mathbf{2})}{=} \frac{1}{2}\Big(\sqrt{k}\left(1-\phi^2/2\right)^{t-1}\Big(\sqrt{1-2\phi} + \sqrt{1+2\phi}\Big) + 2kp\Big) \\
&\overset{(\mathbf{3})}{\leq} \sqrt{k}\left(1-\phi^2/2\right)^t + kp.
\end{aligned}$$

Here ($\mathbf{1}$) follows by the inductive hypothesis. In the end, for ($\mathbf{3}$) 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 $\Phi(L_{k,t}) < \phi$. $L_{k,t}$ must have minimum probability less than p by assumption. 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 in $L_{k,t} \setminus L_{k',t}$ has a probability at most p. By the concavity of $h_t(x)$,

$$h_t(k) \leq h_t(k') + (k-k')p \tag{6.2}$$

$$\leq \sqrt{k'}\left(1-\phi^2/2\right)^t + k'p + (k-k')p \tag{6.3}$$

$$\leq \sqrt{k'}\left(1-\phi^2/2\right)^t + kp. \tag{6.4}$$

□

The following lemma is a direct corollary.

LEMMA 6.8. *Consider a vertex s such that $\|\boldsymbol{\tau}_{s,n^\delta}\|_\infty \geq 1/n^{\delta(i+6)}$. Then, there exists a level set for some $t \leq n^\delta$ with minimum probability at least $1/10n^{\delta(i+6)}$ and conductance $< n^{-\delta/4}$.*

Proof. Suppose the contrary. So, for all $t \leq n^\delta$, if a level set has minimum probability at least $1/10n^{\delta(i+6)}$, it has conductance at least $n^{-\delta/4}$. Since $|S| \geq \alpha n \geq$

$n^{1-\delta/2}$, by choosing δ sufficiently small we get that the minimum probability satisfies $1/10n^{\delta(i+6)} \geq 2/|S|$. Thus, we can apply Lemma 6.7. For all $k \in [0, |S|]$, $h_{n^\delta}(k) \leq \sqrt{k}(1-n^{-\delta/2}/2)^{n^\delta} + k/10n^{\delta(i+6)}$. Setting $k = 1$, $h_{n^\delta}(1) \leq \exp(-n^{\delta/2}/2) + 1/10n^{\delta(i+6)} < 1/n^{\delta(i+6)}$. Note that $h_{n^\delta}(1)$ is the largest probability in $\boldsymbol{\tau}_{s,n^\delta}$, which by assumption is at least $1/n^{\delta(i+6)}$. This is a contradiction. ◻

6.2.3. Wrapping up by producing $\widetilde{S}$.

Proof of Lemma 6.1. Define S'' as given in Lemma 6.5. By Lemma 6.8, for every $s \in S''$, there exists some level set for $t_s \leq n^\delta$ with minimum probability at least $1/10n^{\delta(i+6)}$ and conductance at most $n^{-\delta/4}$. Let us call this level set P_s. Note that $|P_s| \leq 10n^{i(\delta+6)}$, and according to Lemma 6.1 we have $|S| \geq |S'| \geq \alpha n \geq n^{1-\delta/2}$. This implies that $|P_s| < |S|/2$ (for sufficiently small δ and $i \leq 5r^2$). By the construction of M_S, we have

$$\Phi(P_s) \geq \frac{\sum_{\substack{x \in P_s \\ y \in S \setminus P_s}} \tau_{x,1}(y)}{\min(|P_s|, |S \setminus P_s|)} = \frac{E(P_s, S \setminus P_s)}{2d|P_s|}.$$

The first inequality follows because we restrict the numerator to length one transitions in the Markov chain M_S (which correspond to edges in G). Rearranging, we get $E(P_s, S \setminus P_s) \leq n^{-\delta/4}(2d|P_s|)$.

For all $s \in S''$ and $v \in P_s$, $\tau_{s,n^\delta}(v) \geq 1/10n^{\delta(i+6)}$. Set $L = 160n^{\delta(i+7)}/\alpha$. Let $\widetilde{S}$ be the subset of S'' such that for all $s \in \widetilde{S}$, P_s is such that for all $v \in P_s$, $\sum_{l \leq L} p_{s,l}(v) \geq 1/20n^{\delta(i+6)}$. By averaging, $\exists l \leq L$ such that $p_{s,l}(v) \geq \alpha/n^{\delta(2i+14)}$.

We have seen that $\widetilde{S}$ satisfies the two desired properties: for all $s \in \widetilde{S}$, $E(P_s, S \setminus P_s) \leq 2n^{-\delta/4}d|P_s|/\alpha$, and for all $v \in P_s$, $\exists t \leq 160n^{\delta(i+7)}$ such that $p_{s,t}(v) \geq \alpha/n^{\delta(2i+14)}$. It only remains to prove an upper bound on $|S'' \setminus \widetilde{S}|$.

Consider any $s \in S'' \setminus \widetilde{S}$. There exists some $v_s \in P_s$ such that $\tau_{s,n^\delta}(v_s) \geq 1/10n^{\delta(i+6)}$ but $\sum_{l \leq L} p_{s,l}(v_s) < 1/20n^{\delta(i+6)}$. Let us use $\hat{p}_{s,l}(v_s)$ to denote the probability of reaching v_s from s in an l-length walk that makes n^δ hops. Observe that

$$\begin{aligned}
\tau_{s,n^\delta}(v_s) &= \sum_{l \geq n^\delta} \hat{p}_{s,l}(v_s) \\
&= \sum_{l=n^\delta}^{L} \hat{p}_{s,l}(v_s) + \sum_{l>L} \hat{p}_{s,l}(v_s) \\
&\leq \sum_{l=n^\delta}^{L} p_{s,l}(v_s) + \sum_{l>L} \hat{p}_{s,l}(v_s) \\
&< 1/20n^{\delta(i+6)} + \sum_{l>L} \hat{p}_{s,l}(v_s).
\end{aligned}$$

The last inequality follows from the fact that $s \in S'' \setminus \widetilde{S}$, and hence $\sum_{l=n^\delta}^{L} p_{s,l}(v_s) < 1/20n^{\delta(i+6)}$. Since $\tau_{s,n^\delta}(v_s) \geq 1/10n^{\delta(i+6)}$, the above calculation shows that $\sum_{l>L} \hat{p}_{s,l}(v_s) > 1/20n^{\delta(i+6)}$. Thus,

$$\frac{1}{|S|} \sum_{s \in S'' \setminus \widetilde{S}} \sum_{l>L} \hat{p}_{s,l}(v_s)L > \frac{|S'' \setminus \widetilde{S}| \cdot L}{|S|20n^{\delta(i+6)}} = \frac{160\alpha^{-1}n^{\delta(i+7)} \cdot |S'' \setminus \widetilde{S}|}{20|S|n^{\delta(i+6)}} = \frac{8n^\delta|S'' \setminus \widetilde{S}|}{\alpha|S|}.$$

By Lemma 6.3,

$$\frac{1}{|S|}\sum_{s\in S''\setminus\widetilde{S}}\sum_{l>L}\hat{p}_{s,l}(v_s)L \leq \mathbf{E}_{W\sim\mathcal{W}_n^\delta}[\text{length of } W] = \frac{n^{1+\delta}}{|S|}.$$

Combining the above, we get $|S''\setminus\widetilde{S}| \leq \alpha n/8$. By Lemma 6.5, $|S''| \geq |S'|/2 \geq \alpha n/2$, yielding the bound $|\widetilde{S}| \geq \alpha n/4$. ∎

7. Wrapping it all up: The proof of Theorem 3.1. We have all the tools required to complete the proof of Theorem 3.1. Our aim is to show that whenever `FindMinor`(G,ε,H) outputs an H-minor with probability $< 2/3$, then G is ε-close to being H-minor-free. Henceforth in this section, we will simply assume the "if" condition.

Below, we again produce the decomposition procedure used in the proof from section 6. We set $\alpha = \varepsilon/(50r^2\log n)$.

`Decompose`(G)

1. Initialize $S = V$ and $\mathcal{P} = \emptyset$.
2. For $i = 1,\ldots,5r^2$:
   (a) Assign $S' := \left\{s \in S : ||\boldsymbol{q}^{(i+1)}_{[S],s}||_2^2 \geq 1/n^{\delta i}\right\}$
   (b) While $|S'| \geq \alpha n$:
      i. Let $S'' = \{s \in S' : \|\boldsymbol{\tau}_{s,n^\delta}\|_\infty \geq 1/n^{\delta(i+6)}\}$ be as in Lemma 6.5.
      ii. Choose arbitrary $s \in S''$, and let P_s be as in Lemma 6.1.
      iii. Add P_s to $\mathcal{P}$ and assign $S := S \setminus P_s$
      iv. Assign $S' := \left\{s \in S : ||\boldsymbol{q}^{(i+1)}_{[S],s}||_2^2 \geq 1/n^{\delta i}\right\}$
   (c) Assign $S := S \setminus S'$
   (d) Assign $X_i := S'$
3. Let $X = \bigcup_i X_i$.
4. Output the partition $\mathcal{P}, X, S$

The procedure `Decompose` repeatedly employs Lemma 6.1 for values of $i \leq 5r^2$. In the ith iteration, eventually $|S'|$ becomes too small for Lemma 6.1. Then, S' is moved (from S) to an "excess" set X_i, and the next iteration begins. `Decompose` ends with a partition $\mathcal{P}, X, S$ where each set in $\mathcal{P}$ is a low conductance cut, X is fairly small, and `FindBiclique` succeeds with high probability on every vertex in S.

This is formalized in the next lemma.

LEMMA 7.1. *Assume $\varepsilon > \varepsilon_{\mathsf{CUTOFF}}$. Suppose* `FindMinor`$(G,\varepsilon,H)$ *outputs an H-minor with probability $< 2/3$. Then, the output of* `Decompose` *satisfies the following conditions:*

- $|X| \leq \varepsilon n/10$.
- $|S| \leq \varepsilon n/10$.
- *For all $P_s \in \mathcal{P}, v \in P_s$, $\exists t \leq 160n^{6\delta r^2}/\alpha$ such that $p_{s,t}(v) \geq \frac{\alpha}{n^{11\delta r^2}}$.*
- *There are at most $\varepsilon n/10$ edges that go between different P_s sets.*

Proof. Consider the X_i's formed by `Decompose`. Each of these has size at most $\alpha n = \varepsilon n/50r^2\log n$, and there are at most $5r^2$ of these. Clearly, their union has size at most $\varepsilon n/10$.

The third condition holds directly from Lemma 6.1. Consider the number of edges that go between P_s and the rest of S, when P_s was constructed (in `Decompose`). By

Lemma 6.1 again, the number of these edges is at most

$$2n^{-\delta/4}d|P_s|/\alpha = 100r^2(\log n)\varepsilon^{-1}n^{-\delta/4}d|P_s|.$$

Note that $\varepsilon > \varepsilon_{\mathsf{CUTOFF}}$. For sufficiently small constant δ, the number of edges between P_s and $S \setminus P_s$ (at the time of removal) is at most $\varepsilon|P_s|/10$. The total number of such edges is at most $\varepsilon n/10$ (since P_s are all disjoint).

Suppose, for the sake of contradiction, that $|S| > \varepsilon n/10$. Consider the stratification process with $R_0 = S$. By construction of S, for all $s \in S$, $||\boldsymbol{q}^{(5r^2+1)}_{[S],s}|| \leq 1/n^{5\delta r^2}$. Thus, all of these vertices will lie in strata numbered $5r^2$ or above. Since $\varepsilon > \varepsilon_{\mathsf{CUTOFF}}$, by Lemma 4.9, at most $\varepsilon n/\log n$ vertices are in strata numbered higher than $1/\delta + 3$. By Theorem 5.1, for at least $\varepsilon n/10 - \varepsilon n/\log n \geq \varepsilon n/20$ vertices, the probability that the paths discovered by `FindBiclique`(s) contain a $K_{r,r}$-minor is at least $n^{-4\delta r^2}$. A $K_{r,r}$-minor contains a K_r-minor (simply contract any perfect matching), and thus it contains an H-minor. The algorithm succeeds in finding an H-minor with probability at least $n^{-4\delta r^2}$.

All in all, this implies that the probability that a single call to `FindBiclique` finds an H-minor is at least $n^{-4\delta r^2}$. Since `FindMinor` makes $n^{35\delta r^2}$ calls to `FindBiclique`, an H-minor is found with probability at least 5/6. This is a contradiction, and we conclude that $|S| \leq \varepsilon n/10$. ∎

Now we can prove the correctness guarantee of `FindMinor`.

Claim 7.2. *Suppose* `FindMinor`(G, ε, H) *outputs an H-minor with probability* $<$ 2/3*. Then G is ε-close to being H-minor-free.*

Proof. If $\varepsilon \leq \varepsilon_{\mathsf{CUTOFF}}$, then `FindMinor` runs an exact procedure. So the claim is clearly true. Henceforth, assume $\varepsilon > \varepsilon_{\mathsf{CUTOFF}}$. Apply Lemma 7.1 to partition V into $\mathcal{P}, X, S$.

Call $s \in V$ bad if there is a corresponding $P_s \in \mathcal{P}$ and P_s induces an H-minor. By Lemma 7.1, for all $v \in P_s$, $\exists t \leq 160n^{6\delta r^2}/\alpha$ such that $p_{s,t}(v) \geq \alpha/n^{11\delta r^2}$. Note that $160n^{6\delta r^2}/\alpha \leq n^{7\delta r^2}$ and $\alpha/n^{11\delta r^2} \geq n^{-12\delta r^2}$. Also, $|P_s| \leq 160(n^{6\delta r^2}/\alpha) \times (n^{11\delta r^2}/\alpha) \leq n^{18\delta r^2}$. Note that `LocalSearch`$(s)$ performs walks of all lengths up to $n^{7\delta r^2}$ and performs $n^{30\delta r^2}$ walks of each length. For any $v \in P_s$, the probability that `LocalSearch`(s) does not add v to B (the set of "discovered" vertices in `LocalSearch`(s)) is at most $(1 - n^{-12\delta r^2})^{n^{30\delta r^2}} \leq 1/n^2$. Taking a union bound over P_s, the probability that P_s is not contained in B is at most $1/n$. Consequently, for bad s, `LocalSearch`(s) outputs an H-minor with probability $> 1 - 1/n$.

Suppose there are more than $n^{1-30\delta r^2}$ bad vertices. The probability that a u.a.r. $s \in V$ is bad is at least $n^{-30\delta r^2}$. Since `FindMinor`(G, ε, H) invokes `LocalSearch` $n^{35\delta r^2}$ times, the probability that `LocalSearch`(s) is invoked for a bad vertex is at least $1 - 1/n$. Thus, `FindMinor`(G, ε, H) outputs an H-minor with probability $> 1 - 2/n$, contradicting the claim's assumption.

We conclude that there are at most $n^{1-30\delta r^2}$ bad vertices. Each P_s has at most $n^{18\delta r^2}$ vertices, and $|\bigcup_{s \text{ bad}} P_s| \leq n^{1-12\delta r^4} \leq \varepsilon n/10$.

We can make G H-minor-free by deleting all edges incident to X, all edges incident to S, all edges incident to vertices in any bad P_s sets, and all edges between P_s sets. By Lemma 7.1 and the bound given above, the total number of edges deleted is at most $4\varepsilon dn/10 < \varepsilon dn$. ∎

Finally, we bound the running time.

CLAIM 7.3. *The running time of* FindMinor(G, ε, H) *is*

$$dn^{1/2+O(\delta r^2)} + d\varepsilon^{-2\exp(2/\delta)/\delta}.$$

Proof. If $\varepsilon < \varepsilon_{\mathtt{CUTOFF}}$, then the running time is simply $O(n^2)$. Since $\varepsilon < n^{-\delta/\exp(2/\delta)}$, this can be expressed as $\varepsilon^{-2\exp(2/\delta)/\delta}$.

Assume $\varepsilon \geq \varepsilon_{\mathtt{CUTOFF}}$. The total number of vertices encountered by all the LocalSearch calls is $n^{O(\delta r^2)}$. There is an extra d factor for determining all incident edges through vertex queries. Thus, the total running time is $dn^{O(\delta r^2)}$, because of the quadratic overhead of KKR. Consider a single iteration for the main loop of FindBiclique. First, FindBiclique performs $2r$ random walks of length $2^{i+1}n^{5\delta}$, and then for each of these, FindPath performs $n^{\delta i/2+9\delta}$ walks of length $2^i n^{5\delta}$. Hence, the total steps (and thus queries) in all walks performed by a single call to FindBiclique is

$$\sum_{i=5r^2}^{1/\delta+3} \left(2r2^{i+1}n^{5\delta} + 2rn^{\delta i/2+9\delta}2^i n^{5\delta}\right) = rn^{1/2+O(\delta)}. \tag{7.1}$$

While this is the total number of vertices encountered, we note that the calls made to KKR(F, H) are for much smaller graphs. The output of FindPath has size $O(2^{1/\delta}n^{5\delta})$, and the subgraph F constructed has at most $O(2^{1/\delta}n^{5\delta})$ vertices. We incur an extra d factor to determine the induced subgraph through vertex queries. Thus, the time for each call to KKR(F, H) is $n^{O(\delta)}$. There are $n^{O(\delta r^2)}$ calls to FindBiclique, and we can bound the total running time by $dn^{1/2+O(\delta r^2)}$. ∎

Acknowledgments. We would like to acknowledge Madhur Tulsiani for his improvements to Lemma 4.4. We also thank the anonymous reviewers for their helpful comments.

REFERENCES

[1] D. Aldous and J. A. Fill, *Reversible Markov Chains and Random Walks on Graphs*, unfinished monograph, 2002, recompiled 2014, available at http://www.stat.berkeley.edu/~aldous/RWG/book.html.

[2] N. Alon, E. Fischer, I. Newman, and A. Shapira, *A combinatorial characterization of the testable graph properties: It's all about regularity*, SIAM J. Comput., 39 (2009), pp. 143–167, https://doi.org/10.1137/060667177.

[3] R. Andersen, F. R. K. Chung, and K. Lang, *Local graph partitioning using PageRank vectors*, in Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06), Berkeley, CA, 2006, pp. 475–486.

[4] S. Arora, B. Barak, and D. Steurer, *Subexponential algorithms for unique games and related problems*, J. ACM, 62 (2015), 42.

[5] I. Benjamini, O. Schramm, and A. Shapira, *Every minor-closed property of sparse graphs is testable*, Adv. Math., 223 (2010), pp. 2200–2218.

[6] A. Z. Broder, A. M. Frieze, and E. Upfal, *Static and dynamic path selection on expander graphs: A random walk approach*, Random Structures Algorithms, 14 (1999), pp. 87–109.

[7] A. Czumaj, O. Goldreich, D. Ron, C. Seshadhri, A. Shapira, and C. Sohler, *Finding cycles and trees in sublinear time*, Random Structures Algorithms, 45 (2014), pp. 139–184.

[8] A. Czumaj, P. Peng, and C. Sohler, *Testing cluster structure of graphs*, in Proceedings of the 47th Annual ACM Symposium on Theory of Computing (STOC), 2015, pp. 723–732, https://doi.org/10.1145/2746539.2746618.

[9] A. Czumaj, A. Shapira, and C. Sohler, *Testing hereditary properties of nonexpanding bounded-degree graphs*, SIAM J. Comput., 38 (2009), pp. 2499–2510, https://doi.org/10.1137/070681831.

[10] A. Czumaj and C. Sohler, *Testing expansion in bounded-degree graphs*, Combin. Probab. Comput., 19 (2010), pp. 693–709.

[11] R. Diestel, *Graph Theory,* 4th ed., Springer, 2010.

[12] A. Edelman, A. Hassidim, H. N. Nguyen, and K. Onak, *An efficient partitioning oracle for bounded-treewidth graphs*, in Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques (APPROX 2011, RANDOM 2011), L. A. Goldberg et al., eds., Lecture Notes in Comput. Sci. 6845, Springer, 2011, pp. 530–541.

[13] H. Fichtenberger, R. Levi, Y. Vasudev, and M. Wötzel, *On Testing Minor-Freeness in Bounded Degree Graphs with One-Sided Error*, preprint, https://arxiv.org/abs/1707.06126v1, 2017.

[14] H. Fichtenberger, R. Levi, Y. Vasudev, and M. Wötzel, *A sublinear tester for outerplanarity (and other forbidden minors) with one-sided error*, in Proceedings of the 45th International Colloquium on Automata, Languages, and Programming (ICALP 2018), Prague, Czech Republic, LIPIcs 107, I. Chatzigiannakis et al., eds., Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018, 52.

[15] S. O. Gharan and L. Trevisan, *Approximating the expansion profile and almost optimal local graph clustering*, in Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2012), IEEE, 2012, pp. 187–196.

[16] O. Goldreich, *Introduction to Property Testing*, Cambridge University Press, 2017.

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

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

[19] O. Goldreich and D. Ron, *On testing expansion in bounded-degree graphs*, in Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation, Lecture Notes in Comput. Sci. 6650, O. Goldreich, ed., Springer, 2011, pp. 68–75.

[20] A. Hassidim, J. Kelner, H. Nguyen, and K. Onak, *Local graph partitions for approximation and testing*, in Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2009), IEEE, 2009, pp. 22–31.

[21] J. Hopcroft and R. Tarjan, *Efficient planarity testing*, J. ACM, 21 (1974), pp. 549–568.

[22] S. Kale, Y. Peres, and C. Seshadhri, *Noise tolerance of expanders and sublinear expansion reconstruction*, SIAM J. Comput., 42 (2013), pp. 305–323, https://doi.org/10.1137/110837863.

[23] S. Kale and C. Seshadhri, *An expansion tester for bounded degree graphs*, SIAM J. Comput., 40 (2011), pp. 709–720, https://doi.org/10.1137/100802980.

[24] K. Kawarabayashi, Y. Kobayashi, and B. A. Reed, *The disjoint paths problem in quadratic time*, J. Combin. Theory, Ser. B, 102 (2012), pp. 424–435.

[25] J. M. Kleinberg and R. Rubinfeld, *Short paths in expander graphs*, in Proceedings of 37th Annual IEEE Symposium on Foundations of Computer Science (FOCS), IEEE, 1996, pp. 86–95.

[26] A. Kumar, C. Seshadhri, and A. Stolman, *Finding Forbidden Minors in Sublinear Time: A $o(n^{1/2+o(1)})$-Query One-Sided Tester for Minor Closed Properties on Bounded Degree Graphs*, preprint, https://arxiv.org/abs/1805.08187v1, 2018.

[27] K. Kuratowski, *Sur le problème des courbes gauches en topologie*, Fund. Math., 15 (1930), pp. 271–283.

[28] R. Levi and D. Ron, *A quasi-polynomial time partition oracle for graphs with an excluded minor*, ACM Trans. Algorithms, 11 (2015), 24.

[29] L. Lovász, *Graph minor theory*, Bull. Amer. Math. Soc., 43 (2006), pp. 75–86.

[30] L. Lovász and M. Simonovits, *The mixing rate of Markov chains, an isoperimetric inequality, and computing the volume*, in Proceedings of the 31st Annual Symposium on Foundations of Computer Science (FOCS), IEEE, 1990, pp. 346–354.

[31] A. Nachmias and A. Shapira, *Testing the expansion of a graph*, Inf. Comput., 208 (2010), pp. 309–314.

[32] N. Robertson and P. D. Seymour, *Graph minors* XX: *Wagner's conjecture*, J. Combin. Theory Ser. B, 92 (2004), pp. 325–357.

[33] D. Spielman, *Spectral Graph Theory and Its Applications,* lecture notes, http://www.cs.yale.edu/homes/spielman/eigs/, 2004.

[34] D. A. Spielman and S.-H. Teng, *A local clustering algorithm for massive graphs and its application to nearly linear time graph partitioning*, SIAM J. Comput., 42 (2013), pp. 1–26, https://doi.org/10.1137/080744888.

[35] L. Trevisan, *Approximation algorithms for unique games*, in Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS), IEEE, 2005, pp. 197–205.

[36] K. Wagner, *Über eine Eigenschaft der ebenen Komplexe*, Math. Ann., 114 (1937), pp. 570–590.

[37] Y. Yoshida and H. Ito, *Testing outerplanarity of bounded degree graphs*, Algorithmica, 73 (2015), pp. 1–20.