Electronic Colloquium on Computational Complexity, Report No. 197 (2026)

Testing Bipartite in the Bounded-Degree Graph Model, Revisited (A digest of the paper of Fei and Rubinfeld (2026))

Oded Goldreich
Department of Computer Science
Weizmann Institute of Science, Rehovot, Israel.

September 20, 2026

Abstract

We provide a digest of the paper of Fei and Rubinfeld (arXiv 2026), which provides an alternative analysis of the Bipartite Tester of Goldreich and Ron (*Combinatorica*, 1999), which operates in the bounded-degree graph model. Loosely speaking, on input an n-vertex graph G, the tester selects few vertices, conducts $\widetilde{O}(n^{1/2})$ random walks of polylogarithmic length from each selected vertex, and rejects if and only if an odd cycle is formed by a pair of walks.

While the analysis of the foregoing tester in the rapid-mixing case is quite appealing, the original analysis of the general case is quite imposing; it involves the introduction and analysis of Markov Chains that capture the behavior of random walks on a sequence of residual subgraphs that are iteratively defined by the analysis. In contrast, Fei and Rubinfeld avoid this iterative process, and present an analysis that only refers to the random walks on the input graph.

More specifically, the original analysis derives a sequence of (non-overlapping) partial 2-partitions of the graph, and stitches them together. In contrast, the new analysis combines a set of "fractional" 2-partitions of the entire graph, where the combination is obtained by defining adequate vectors that represent these fractional 2-partitions and employing randomized rounding (a la Goemans and Williamson (*JACM*, 1995)).

In addition, the new analysis allows for presenting an extremely efficient interactive proof of proximity for Bipartiteness. Such an interactive proof was known before for the rapid-mixing case (Rothblum, Vadhan, and Wigderson (*STOC*, 2013)).

Contents

ISSN 1433-8092

1 Introduction

The bounded-degree graph model, introduced by Goldreich and Ron [3], is the second most popular model within the context of testing graph properties (cf. [2, Chap. 9]). Unlike the dense graph model (cf. [2, Chap. 8]), which is meaningful mainly for dense graphs, the bounded-degree model is meaningful for sparse graphs, alas not as much as the general graph model (cf. [2, Chap. 10]). On the other hand, the bounded-degree model offers many positive results and serves as a good starting point towards studying the general graph model. In addition, both models are more aligned (than the dense graph model) with techniques that can be applied in the context of traditional algorithms.

One of the foregoing positive results is the Bipartite tester of Goldreich and Ron [4]. In fact, it is fair to say that this tester is one of the most popular results in the area of testing graph properties. For a degree bound d, given an n-vertex graph G (of maximum degree d) and a proximity parameter $\epsilon > 0$, the tester runs in $\text{poly}(1/\epsilon) \cdot \widetilde{O}(n^{1/2})$-time, always accepts if G is bipartite, and rejects with probability at least $2/3$ if the G is ϵ-far from being bipartite (i.e., more than $\epsilon dn/2$ edges must be omitted from G in order to yield a bipartite graph).

The tester itself is very intuitive. Essentially, it selects $O(1/\epsilon)$ vertices and conducts $m = \text{poly}(1/\epsilon) \cdot \widetilde{O}(n^{1/2})$ random walks of length $\ell = \text{poly}(\epsilon^{-1} \cdot \log n)$ from each selected vertex. The tester rejects if and only if its exploration, which uncovers an $O(m \cdot \ell/\epsilon)$-vertex subgraph, contains an odd-length cycle. Actually, the tester only looks for odd-length cycles that are formed by a collision of two walks that start at the same vertex.

As presented in [4] and reviewed in [2, Sec. 9.4.1], the analysis of this tester is very appealing when assuming that the graph is rapid-mixing (hereafter referred to as the rapid-mixing case).[1] Unfortunately, the analysis of the general case is quite imposing (e.g., it takes 26 pages in [4] and is avoided in [2, Sec. 9.4.1]). The key observation is that if the tester accepts the graph G with high probability, then this yields a 2-partition of the vertices of G with few violating edges (i.e., edges with both endpoints on the same side). While in the rapid-mixing case this 2-partition is easy to derive, in the general case it is derived in iterations, while considering auxiliary Markov chains that represent random walks on residual subgraphs.

In contrast, Fei and Rubinfeld [1] propose to abandon this iterative process and combine the n "fractional" 2-partitions derived from the n different start vertices in one shot. The combination is performed by considering n vectors that represent, for each start vertex s, the probability of reaching the various vertices by a random (logarithmically long) walk of odd (resp., even) number of steps, and 2-partitioning these vectors by randomized rounding (via a random hyperplane, as in the GW-algorithm for MaxCUT [6]).

A slightly more detailed description of the analysis. Assuming that the proposed tester rejects with low probability, the analysis of [4] implicitly defines n potential 2-partitions of the n-vertex graph, each based on the distribution of parity-sensitive random walks starting from each of the graph vertices. Specifically, for each vertex $s \in [n]$ and each $\sigma \in \{0, 1\}$, we consider the distribution X_s^σ of the endpoint of a random walk of length $2 \cdot O(\log n) + \sigma$ that starts at s. We then partially 2-partition the vertices of the graph accordingly (i.e., v is place on side σ if $\mathbf{Pr}[X_s^\sigma = v] > \mathbf{Pr}[X_s^{1-\sigma} = v]$).[2] Each of these (partial) 2-partitions violates relatively few edges, but

[1] In this case, we assume that the endpoint of an $O(\log n)$-long random walk (starting at any vertex) is almost uniformly distributed in the vertex set.

[2] Indeed, vertices that are not reached at all are not assigned a side in this 2-partition. Actually, it makes sense not to assign a side to vertices that are reached with very low probability.

unless the graph is sufficiently rapid-mixing none of these 2-partitions covers all the graph vertices.

In the latter case, the original analysis stitches these (partial) 2-partitions in a gradual process in which the vertices that were already assigned are omitted from the graph, which calls for the definition of residual 2-partitions (used instead of the original 2-partitions).[3] Instead, as suggested by [1], one may use the original probabilities to define n (unit) vectors, view them as a solution to an SDP-relaxation of the original problem (i.e., "fractional" 2-partitions), and employ the rounding procedure of [6] in order to obtain a solution to the original problem (i.e., Bipartiteness).

A key issue is the selection of these n vectors. We want these n (unit) vectors to represent the foregoing $2n$ distributions (i.e., the X_s^σ's) such that *if the tester rejects with small probability, then almost all pairs of vectors that correspond to edges are almost antipolar*, which in turn implies that the graph is close to being bipartite. One natural choice is to associate the vertex s with the n-dimensional vector in which the $i^{\rm th}$ coordiante is proportional to $\mathbf{Pr}[X_s^0 = i] - \mathbf{Pr}[X_s^1 = i]$, but working with these vectors runs into technical problems that I was not able to resolve. Specifically, the issue seems to be normalization (i.e., normalizing these vectors so that they are of length 1). Fei and Rubinfeld [1] resolve the problem by using a less natural set of unit-length vectors (see Eq. (3)), and I was forced to adopt their choice.[4]

Let me also point out that, as stressed in [1], the improved analysis yields a query complexity upper bound of $\mathrm{poly}(1/\epsilon) \cdot O(\sqrt{n} \cdot \log n)$ rather than the bound $\mathrm{poly}(1/\epsilon) \cdot O(\sqrt{n} \cdot \log^7 n)$ obtained in [4]. This is due to disposing of the use of auxiliary Markov Chains (which were used in [4] when handling the non-rapid-mixing case; see Footnote 3). Recall that $\Omega\sqrt{n/\epsilon})$ is a lower bound on the query complexity of testing Bipartitness in the bounded-degree model (even for degree-bound 3 and the rapid-mixing case (cf. [3])).[5]

An alternative presentation. The current text offers a digest of [1]. The presentation in [1] is "technically-oriented" (e.g., it is presented in terms of the Markov operator, and offers a measure theoretic framework and an SDP perspective). In contrast, I prefer a conceptually-oriented presentation that utilizes direct arguments and a minimal amount of "technology". I also deviate from [1] in some concrete aspects, which are discuss in Section 5.

An interactive proof of proximity for Bipariteness. Interactive proofs of proximity, introduced by Rothblum, Vadhan, and Wigderson [9], are a hybrid of property testing and interactive proofs. In these systems, the tester is replaced by a verifier, who can communicate with an untrusted prover, who (unlike the verifier) has free access to the input (cf. [9, Sec. 2] and [2, Sec. 12.6]). The new analysis of the Bipartite tester provided by [1] allows us to obtain an extremely efficient interactive proof of proximity for Bipartiteness. The new interactive proof builds on the interactive proof of proximity for the rapid mixing case presented in [9, Sec. 5.1].

[3] Each of these 2-partitions is defined based on an imaginary random walk that takes place on an auxiliary Markov Chain that contains a subgraph of the original graph as well as auxiliary gadgets. Relating random walks on these MCs to random walks on the original graph requires using longer walks on the original graph (i.e., longer than in the case of rapid-mixing).

[4] The $i^{\rm th}$ coordiante of the vector associated with s is $(-1)^{\sigma_s(i)} \cdot \sqrt{(\mathbf{Pr}[X_s^0 = i] + \mathbf{Pr}[X_s^1 = i])/2}$ such that $\sigma_s(i)$ equals the sign of $\mathbf{Pr}[X_s^0 = i] - \mathbf{Pr}[X_s^1 = i]$.

[5] We note that in the case of rapid mixing graphs, Bipartite testing can be done using $O(\sqrt{n/\epsilon})$ queries (see `https://www.wisdom.weizmann.ac.il/~oded/pt-proj1.html`). We also note that, in case of degree-bound 2, all graphs properties can be tested with $\widetilde{O}(1/\epsilon^2)$ queries [5].

Organization. We assume familiarity with property testing in the bounded-degree graph model (cf. [2, Sec. 9.1]). In Section 2 we review the Bipartite tester of [4] and its analysis in the rapid-mixing case. This serves as a good but inessential warm-up towards the core of the current text, presented in Section 3. Additional comment appear in Sections 4 and 5. The interactive proof of proximity for Bipartiteness is presented in Section 6,

2 Background: The tester of [4] in the rapid-mixing case

The Bipartite tester of [4] consists of taking $\widetilde{O}(\sqrt{n})$ random walks of polylogarithmic length from few start vertices, where n denotes the number of vertices in the graph. Specifically, the tester selects $O(1/\epsilon)$ random start vertices, and perform the following vertex test from each of them. Actually, in the rapid mixing case (defined below), using a single start vertex suffices.

> For a start vertex s, the s-vertex-test consists of taking $O(\sqrt{n}/\epsilon)$ lazy random walks of length $O(\log n)$ from s, and rejecting if the explored subgraph is not bipartite.
>
> (The actual analysis refers to a relaxed rejection condition by which the tester rejects only when it sees a pair of walks that form an odd cycle.)

For the sake of analysis, it is convenient to consider lazy random walks. For a degree bound d and length parameter ℓ, an ℓ-step lazy random walk (on a graph G) from a start vertex s is a sequence of vertices $(v_0, v_1, ..., v_\ell)$ such that $v_0 = s$ and v_i equals v_{i-1} with probability $1 - (\deg(v_{i-1})/2d)$ and is a uniformly distributed neighbor of v_{i-1} otherwise. In other words, upon reaching a vertex v of degree $\deg(v)$, the walk stays in v with probability $1 - (\deg(v)/2d)$, and moves to a uniformly selected neighbor otherwise.[6] The steps in which the walk stays in place are called lazy. Here we fix the length of the walk, denoted ℓ, setting it to $O(\log n)$.

Central to the analysis is the parity of a walk, defined as the number of actual (i.e., non-lazy) steps taken in the walk (i.e., the parity of the number of i's such that $v_i \neq v_{i-1}$). We focus on the vertex reached at the end of a random walk, and denote by $p_s^\sigma(v)$ the probability that an ℓ-step lazy random walk (hereafter ℓ-LRW or LRW) starting at s ends at v with parity σ. Noting that the vertex-test rejects whenever two walks of different parity collide (forming an odd cycle of length at most $2\ell - 1$), we focus on analyzing this event.

Using $\ell = \Omega(\log n)$, note that the distribution of the endpoint of an ℓ-LRW (starting at s) is approximated by the distribution of the endpoint of an $(\ell - 1)$-LRW (starting at s). Hence, $p_r^\sigma(v) > (1 - o(1)) \cdot p_s^{1-\sigma}(v)/2d$ if r neighbors s, since a walk from r may start with a move to s. Furthermore, in case the graph is rapid-mixing, the endpoint of an ℓ-LRW (starting at any fixed vertex) is almost uniformly distributed in the vertex set; specifically, we say that the n-vertex graph is rapid mixing if for every s and v it holds that $p_s(v) \in (1 \pm 0.1)/n$, where $p_s(v) = p_s^0(v) + p_s^1(v)$ is the probability that in ℓ-LRW starting at s ends at v (regardless of the walk's parity).

The key observation of [4] is that, in case G is rapid mixing and $\ell = O(\log n)$, the following three claims holds (for any s).

1. The case $\sum_v p_s^0(v)p_s^1(v) = 0$ can be distinguished from the case $\sum_v p_s^0(v)p_s^1(v) > \epsilon/100n$ by taking $O(\sqrt{n/\epsilon})$ random walks from s and looking for collisions between the endpoints of pairs of walks.

[6] Allowing graphs with parallel edges means that we actually pick a random edge incident at v and traversing it.

(The rapid-mixing hypothesis is used to show that the variance of this experiment is small enough.)

2. If $\sum_v p_s^0(v)p_s^1(v) > \epsilon/100n$, then the vertex-test rejects with probability at least 3/4.

   (Recall that the vertex-test rejects if it sees a collision between the endpoints of walks of different parities.)

3. If $\sum_v p_s^0(v)p_s^1(v) \le \epsilon/100n$, then the graph is ϵ-close to being bipartite.

   (We color a vertex by 0 if and only if $p_s^0(v) > p_s^1(v)$, and observe that each monochromatic edge contributes $\Omega(1/dn^2)$ units to the sum. The rapid-mixing hypothesis is used here too.)

An analogous strategy is used in the general case, but this requires additional ideas (as outlined in the introduction). This is exactly where [1] depart from [4].

3 The main presentation

First, unlike [1], we prefer to work in the bounded-degree graph model, with constant degree bound $d \ge 3$, and later transform the result to the general graph model by using a generic reduction (a la [7]). Second, we believe that it is simpler to analyze a slightly modified version of the original algorithm of [4]. Specifically, rather than selecting $O(1/\epsilon)$ random vertices, we select $O(1/\epsilon)$ random edges and take $\text{poly}(1/\epsilon) \cdot \sqrt{n}$ random walks from each endpoint of each edge. That is, for each selected edge $\{r, s\}$, we perform the following sub-test, hereafter called the $\{r, s\}$-edge test.

> We take $m \stackrel{\text{def}}{=} \text{poly}(1/\epsilon) \cdot \sqrt{n}$ lazy random walks of length $\ell \stackrel{\text{def}}{=} \text{poly}(1/\epsilon) \cdot \log n$ from both r and s, and *reject if the subgraph that we explore is not bipartite.*

Actually, in the analysis, for some $\ell \in [L]$ (to be determined later), we only rely on the rejection of the $\{r, s\}$-edge test in case there is a *collision between the endpoints of ℓ-long prefixes of two walks of the same parity* (from the two different start vertices). The first part of the analysis (which refers to this event) is oblivious to ℓ.

(Recall that by a lazy random walk, when using the degree bound d, we mean a walk that upon reaching a vertex v of degree d_v, stays in v with probability $1 - (d_v/2d)$, and moves to a uniformly selected neighbor otherwise. Also recall that the parity of a walk is defined as the number of actual (i.e., non-lazy) steps taken in it, and that $p_s^\sigma(v)$ denotes the probability that an ℓ-step lazy random walk starting at s ends at v with parity σ.)

A road map for the rest of this section. As hinted upfront, the new analysis of the Bipartite tester is based on representing the foregoing $2n$ probability functions by n (unit) vectors. The analysis proceeds in three steps.

1. Showing that if some quantity that depends on the vectors (equiv., on the distributions) is large, then the tester rejects with high probability.

2. Showing that if this quantity is small, then almost all pairs of vectors that correspond to edges are almost antipolar.

3. Showing that if almost all pairs of vectors that correspond to edges are almost antipolar, then the graph is close to being bipartite.

These three steps are captured by the three lemmas that follow. The third step uses the randomized rounding of [6], but the real challenge is preparing for it. The first step uses a strategy akin to the one used in [4], but the challenge is performing it with a quantity that can be linked to the hypothesis of the third step. Hence, the most difficult part is the second step (titled "towards a global (GW-analysis) analysis"), which bridges the gap between what works in the first step and what works in the third step.

(Let me stress that the contribution of [1] is both conceptual and technical. The conceptual contribution lies in realizing that the original $2n$ probability functions (i.e., the p_s^σ's) can be used to define a quantity such that the test rejects if this quantity is large, whereas a 2-coloring with few monochromatic edges is obtained otherwise. The technical contribution lies in materializing this strategy, which amounts to selecting the right quantity and performing the second foregoing step.)

Selecting an adequate quantity and performing the first step. While the analysis in [4] is pivoted at $\sum_v p_s^0(v)p_s^1(v)$ (which reflects the probability of collision from two walks of opposite parity starting at s), one is tempted to consider the *parity-conforming collision probability* of walks starting at the two endpoints of an edge; that is,

$$\mathtt{pccp}(r,s) \stackrel{\rm def}{=} \sum_{v\in[n]} \sum_{\sigma\in\{0,1\}} p_r^\sigma(v)\cdot p_s^\sigma(v). \tag{1}$$

Clearly, if $G=([n],E)$ is bipartite, then, for every edge $\{r,s\}\in E$, the sum in Eq. (1) equals 0. Adapting a key observation of [4], one can show that if the sum in Eq. (1) relatively large (i.e., a constant fraction of the ordinary collision probability), then the corresponding edge-test rejects with high probability.[7] Unfortunately, the pccp does not seem to work well with the global (GW-type) analysis, and so (following [1]) we consider a related quantity, which represents *minimum matching parity probabilities*; that is,

$$\mathtt{mmpp}(r,s) \stackrel{\rm def}{=} \sum_{v\in[n],\sigma\in\{0,1\}} \min(p_r^\sigma(v), p_s^\sigma(v)). \tag{2}$$

Again, if $G=([n],E)$ is bipartite, then for every edge $\{r,s\}\in E$, the sum in Eq. (2) equals 0. And again, as observed in [1], if Eq. (2) is large, then the corresponding edge-test rejects with high probability; that is,

Lemma 1 (large mmpp yields rejection (adapted from [1])): *For any $\tau>0$ and $m\geq\sqrt{n}/\Omega(\tau)$, if for $\{r,s\}\in E$ it holds that $\mathtt{mmpp}(r,s)\geq\tau$, then the $\{r,s\}$-edge-test rejects with probability at least $3/4$.*

Hence, if at least $\epsilon\cdot|E|/2$ of the edges $\{r,s\}$ satisfy $\mathtt{mmpp}(r,s)\geq\sqrt{n}/\Omega(m)$, then the tester rejects with probability at least $2/3$. Thus, we focus on the case that more than $(1-0.5\cdot\epsilon)\cdot|E|$ of the edges have a mmpp value smaller than τ. (We comment that the proof of Lemma 1 is simpler than the analogous analysis of relatively large pccp.)[8]

---

[7] The ordinary collision probability of such two walks equals $\mathtt{cp}(r,s) \stackrel{\rm def}{=} \sum_{v\in[n]} p_r(v)\cdot p_s(v)$, where $p_u(v)=p_u^0(v)+p_u^1(v)$. Specifically, we can show that if $\mathtt{pccp}(r,s)\geq\tau\cdot\mathtt{cp}(r,s)$, then the $\{r,s\}$-edge-test rejects with probability at least $1-\frac{O(\tau^{-2})}{\mathtt{cp}(r,s)^{1/2}\cdot m}$. (Also, for $\ell\geq\Theta(\log n)$, it holds that $\mathtt{cp}(r,s)=\Omega(1/dn)=\Omega(1/n)$.)

[8] This is the case because relating the rejection probability to pccp uses a concentration bound, which refers to all pairs of lazy random walks, whereas the proof of Lemma 1 uses the more crude Birthday paradox.

Proof: For $i,j \in [m]$, let $\zeta_{i,j}$ be a 0-1 random variable such that $\zeta_{i,j} = 1$ if the $i^{\rm th}$ LRW starting at r ends in the same vertex and same parity as the $j^{\rm th}$ LRW starting at s. Then, $\mathbf{Pr}[\zeta_{i,j}=1] \geq \sum_{v,\sigma} q(v,\sigma)^2$, where $q(v,\sigma) \stackrel{\rm def}{=} \min(p_r^\sigma(v), p_s^\sigma(v))$. Using a Birthday-type argument (for the selection of elements with probability q), the claim follows. ■

Towards a global (GW-type) analysis. Assuming that the tester rejects with low probability, we wish to show that the graph is close to being bipartite. The key step is to represent the $2n$ probability functions (i.e., $p_s^\sigma : [n] \to [0,1]$) by n vectors such that vectors that correspond to neighboring vertices are almost antipolar. In addition, in this case, we wish the vectors to be approximately of unit length.

One natural choice is to associate with each vertex $s \in [n]$ an n-dimensional vector, denoted $\vec{\delta_s}$, such that location v in $\vec{\delta_s}$ holds the value $\frac{p_s^0(v)-p_s^1(v)}{\mathtt{cp}(s)^{1/2}}$, where $\mathtt{cp}(s)^{1/2} = \|p_s(\cdot)\|_2$ is a normalizing factor representing the length of the vector p_s such that $p_s(v) = p_s^0(v) + p_s^1(v)$. Indeed, for a bipartite graph, the vector $\vec{\delta_s}$ has length 1, since in that case $|p_s^0(v) - p_s^1(v)| = p_s(v)$ and $\sum_v (p_s(v)/\mathtt{cp}(s)^{1/2})^2 = 1$. Unfortunately, this natural choice does not seem to work (or at least we failed to make it work).[9] It seems that the source of trouble is the normalization by $\mathtt{cp}(s)^{1/2}$, which we are going to avoid next.

In light of the foregoing, we define the vectors in a way that avoids the normalization factor. Specifically, we associate with each vertex $s \in [n]$ an n-dimensional vector, denoted $\vec{\rho_s}$, such that location v in $\vec{\rho_s}$ holds the value

$$\vec{\rho_s}(v) \stackrel{\rm def}{=} \begin{cases} \sqrt{p_s(v)} & \text{if } p_s^0(v) > p_s^1(v) \\ -\sqrt{p_s(v)} & \text{otherwise (i.e., } p_s^0(v) \leq p_s^1(v)) \end{cases} \tag{3}$$

where $p_s(v) = p_s^0(v) + p_s^1(v)$. (This looks a bit weird, since the assignment of vectors (according to a random hyperplane) seems quite arbitrary when $p_s^0(v) \approx p_s^1(v)$, but we can ignore pairs (s,v) of this type (since they make a relatively large contribution towards rejection).) In any case, the good news are that the vector $\vec{\rho_s}$ has unit length. The bad news is that it seems harder to relate the inner product of vectors (associated with neighboring vertices) to the corresponding value of their $\mathtt{mmpp}$. Specifically, the inner product of the vectors $\vec{\rho_r}$ and $\vec{\rho_s}$, denoted $\langle \vec{\rho_r}, \vec{\rho_s} \rangle$, equals

$$\sum_{v\in[n]} (-1)^{\sigma_{r,s}(v)} \cdot \sqrt{p_r(v)\cdot p_s(v)}, \tag{4}$$

where $\sigma_{r,s}(v)$ is the sign of $\vec{\rho_r}(v) \cdot \vec{\rho_s}(v)$. Although $\langle \vec{\rho_r}, \vec{\rho_s} \rangle$ is difficult to relate to $\mathtt{mmpp}(r,s)$, this is possible (as shown in the proof of Lemma 2), provided that we consider specific values of $\ell = O(\log n)$. Recall that so far we used a generic $\ell = O(\log n)$, although we did not even use this condition. This is going to change now, and towards this change we make ℓ explicit in the notation; that is, we use $\mathtt{mmpp}_\ell(r,s)$ instead of $\mathtt{mmpp}(r,s)$, and $\vec{\rho_{\ell,x}}$ instead of $\vec{\rho_x}$.

Lemma 2 (small $\mathtt{mmpp}$ implies inner-product close to -1 (adapted from [1])): *For most $\ell \in [O(\tau^{-1}\cdot \log n)]$, the following holds. If $\sum_{\{r,s\}\in E} \mathtt{mmpp}_\ell(r,s) \leq \tau\cdot|E|$, then $\sum_{\{r,s\}\in E} \langle \vec{\rho_{\ell,r}}, \vec{\rho_{\ell,s}} \rangle \leq -(1-O(\tau))\cdot$*

[9] There exist bipartite graphs in which, for most edges $\{r,s\}$, the inner product of the vectors $\vec{\delta_r}$ and $\vec{\delta_r}$, which equals $\frac{-\mathtt{cp}(r,s)}{\mathtt{cp}(r)^{1/2}\cdot\mathtt{cp}(s)^{1/2}}$, is bounded away from -1. (Recall that $\mathtt{cp}(r,s) = \sum_v (p_r(v)\cdot p_s(v))$.)

$|E|$. *Furthermore, for every $E' \subseteq E$ such that $|E'| = \Omega(|E|)$, the claim holds for E replaced by E'.*[10]

Proof: Recalling that for unit-length vectors $\vec{x}, \vec{y}$ it holds that $\langle \vec{x}, \vec{y} \rangle = -1 + \frac{\|\vec{x}+\vec{y}\|_2^2}{2}$, we focus on upper-bounding $\sum_{\{r,s\}\in E} \|\vec{\rho_{\ell,r}} + \vec{\rho_{\ell,s}}\|_2^2$, which equals $\sum_{\{r,s\}\in E} \sum_{v\in[n]} (\vec{\rho_{\ell,r}}(v) + \vec{\rho_{\ell,s}}(v))^2$. The proof consists of two parts. The first part relates $\|\vec{\rho_{\ell,r}} + \vec{\rho_{\ell,s}}\|_2^2$ and $\texttt{mmpp}_\ell(r,s)$ and holds for any ℓ. Unfortunately, this relation contains a slackness term, which is upper-bounded in the second part of the proof, albeit this holds only for most ℓ.

Relating $\|\vec{\rho_{\ell,r}} + \vec{\rho_{\ell,s}}\|_2^2$ and $\texttt{mmpp}_\ell(r,s)$. Since this part holds for any ℓ, we allow ourself to omit ℓ from the notation so to simplify the expressions. For each $\{r,s\} \in E$ and each $v \in [n]$, we consider the contribution of $(\vec{\rho_r}(v) + \vec{\rho_s}(v))^2$ to the foregoing sum (i.e., to $\sum_{\{r,s\}\in E} \sum_{v\in[n]} (\vec{\rho_r}(v) + \vec{\rho_s}(v))^2$). If $\sigma_{r,s}(v) = -1$, then this contribution equals $(\sqrt{p_r(v)} - \sqrt{p_s(v)})^2$. Otherwise (i.e., $\sigma_{r,s}(v) = 1$), this contribution equals $(\sqrt{p_r(v)} + \sqrt{p_s(v)})^2$, which is upper-bounded as follows, when using $M \stackrel{\rm def}{=} \max(\sqrt{p_r(v)}, \sqrt{p_s(v)})$ and $m \stackrel{\rm def}{=} \min(\sqrt{p_r(v)}, \sqrt{p_s(v)})$:

$$\begin{aligned}
(\sqrt{p_r(v)} + \sqrt{p_s(v)})^2 &= ((M-m) + 2m)^2 \\
&\le 2 \cdot (M-m)^2 + 2 \cdot (2m)^2 \\
&= 2 \cdot \left(\sqrt{p_r(v)} - \sqrt{p_s(v)}\right)^2 + 8 \cdot \min\left(\sqrt{p_r(v)}, \sqrt{p_s(v)}\right)^2 \\
&\le 2 \cdot \left(\sqrt{p_r(v)} - \sqrt{p_s(v)}\right)^2 + 16 \cdot \max_{b\in\{0,1\}} \left\{\min(p_r^b(v), p_s^b(v))\right\}
\end{aligned}$$

where the last inequality uses the fact that $\max_{b\in\{0,1\}}\{\min(p_r^b(v), p_s^b(v))\} \ge \min(p_r(v)/2, p_s(v)/2)$, which in turn holds *since $\sigma_{r,s}(v) = 1$ indicates that $p_r^0(v) > p_r(v)/2$ if and only if $p_s^0(v) > p_s(v)/2$*. Needless to say, this case dominates the prior one. Hence, summing over all $v \in [n]$, we get

$$\|\vec{\rho_r} + \vec{\rho_s}\|_2^2 \;\le\; 2 \cdot \sum_{v\in[n]} \left(\sqrt{p_r(v)} - \sqrt{p_s(v)}\right)^2 + 16 \cdot \sum_{v\in[n]} \max_{b\in\{0,1\}} \left\{\min(p_r^b(v), p_s^b(v))\right\}. \tag{5}$$

The second sum in the r.h.s of Eq. (5) equals $\texttt{mmpp}_\ell(r,s)$, whereas the first sum is a slackness term, which will be denoted $\Delta_\ell(r,s)$. Indeed, towards the second part of the proof, we make ℓ explicit in all notations. In particular, we replace $p_x(v)$ by $p_\ell(x,v)$, and define

$$\Delta_\ell(r,s) \stackrel{\rm def}{=} \sum_{v\in[n]} \left(\sqrt{p_\ell(r,v)} - \sqrt{p_\ell(s,v)}\right)^2. \tag{6}$$

Hence, Eq. (5) is re-written as

$$\|\vec{\rho_{\ell,r}} + \vec{\rho_{\ell,s}}\|_2^2 \;\le\; 2 \cdot \Delta_\ell(r,s) + 16 \cdot \texttt{mmpp}_\ell(r,s). \tag{7}$$

Note that $\Delta_\ell(r,s)$ is negligible if ℓ exceeds the "mixing time" of the graph, but our focus is on the case that ℓ is significantly smaller than the mixing time. We conjecture that $\sum_{\{r,s\}\in E} \Delta_\ell(r,s) \le$

[10] That is, if $\sum_{\{r,s\}\in E'} \texttt{mmpp}_\ell(r,s) \le \tau \cdot |E'|$, then $\sum_{\{r,s\}\in E'} \langle \vec{\rho_{\ell,r}}, \vec{\rho_{\ell,s}} \rangle \le -(1 - O(\tau)) \cdot |E'|$.

$\tau \cdot |E|$ holds for every sufficiently large $\ell = O(\tau^{-1} \cdot \log n)$, but can only prove a weaker statement (which suffices for our application).

Upper-bounding $\Delta_\ell(r,s)$ on the average (for typical ℓ). Recall that $\sum_{\{r,s\}\in E} \mathtt{mmpp}_\ell(r,s) \le \tau \cdot |E|$ (by our hypothesis), and so our current goal is to prove that, for $L = O(\tau^{-1} \cdot \log n)$ and most $\ell \in [L]$, it holds that

$$\sum_{\{r,s\}\in E} \Delta_\ell(r,s) \quad \le \quad \tau \cdot |E|. \tag{8}$$

This is done by proving that

$$\sum_{\ell\in[L]} \sum_{\{r,s\}\in E} \Delta_{\ell-1}(r,s) \;\le\; dn \cdot \log_2 n, \tag{9}$$

where we assume that $|E| = \Omega(dn)$.[11] Let $N(s)$ denote the multiset of neighbors of vertex s (and reacll that $|N(s)| \le d$). Then,

$$\begin{aligned} p_\ell(s,v) \quad &= \quad \frac{2d-|N(s)|}{2d} \cdot p_{\ell-1}(s,v) + \frac{1}{2d} \cdot \sum_{r\in N(s)} p_{\ell-1}(r,v) \\ &= \quad \frac{d-|N(s)|}{d} \cdot p_{\ell-1}(s,v) + \frac{1}{d} \cdot \sum_{r\in N(s)} \frac{p_{\ell-1}(s,v)+p_{\ell-1}(r,v)}{2} \end{aligned}$$

Hence,

$$\mathbf{H}(p_\ell(s,\cdot)) \quad \ge \quad \frac{d-|N(s)|}{d} \cdot \mathbf{H}(p_{\ell-1}(s,\cdot)) + \frac{1}{d} \cdot \sum_{r\in N(s)} \mathbf{H}\left(\frac{p_{\ell-1}(s,\cdot)+p_{\ell-1}(r,\cdot)}{2}\right) \tag{10}$$

At this point we apply the inequality (cf. [1, Lem. A.4])

$$-\frac{x+y}{2} \cdot \log_2 \frac{x+y}{2} \;\ge\; \frac{1}{2} \cdot \left((\sqrt{x}-\sqrt{y})^2 - x\log_2 x - y \log_2 y \right)$$

to each term in the expression for $\mathbf{H}(\frac{p_{\ell-1}(s,\cdot)+p_{\ell-1}(r,\cdot)}{2})$ in Eq. (10), and get

$$\begin{aligned} \mathbf{H}(p_\ell(s,\cdot)) \quad \ge \quad & \frac{d-|N(s)|}{d} \cdot \mathbf{H}(p_{\ell-1}(s,\cdot)) \\ & + \frac{1}{2d} \cdot \sum_{r\in N(s)} \left(\sum_{v\in[n]} \left(\sqrt{p_{\ell-1}(s,v)} - \sqrt{p_{\ell-1}(r,v)} \right)^2 \;+ \mathbf{H}(p_{\ell-1}(s,\cdot)) + \mathbf{H}(p_{\ell-1}(r,\cdot)) \right) \\ = \quad & \mathbf{H}(p_{\ell-1}(s,\cdot)) + \frac{1}{2d} \cdot \sum_{r\in N(s)} \Delta_{\ell-1}(s,r) + \frac{1}{2d} \cdot \sum_{r\in N(s)} \Big(\mathbf{H}(p_{\ell-1}(r,\cdot)) - \mathbf{H}(p_{\ell-1}(s,\cdot)) \Big) \end{aligned}$$

Summing over all $s \in [n]$, while letting $h_t \stackrel{\text{def}}{=} \sum_{s\in[n]} \mathbf{H}(p_t(s,\cdot))$, we get for each $\ell \in [L]$

$$h_\ell \ge h_{\ell-1} + \frac{1}{d} \cdot \sum_{\{s,r\}\in E} \Delta_{\ell-1}(s,r)$$

[11] At this point, the reader may assume that $d = O(1)$. and that G has no isolated vertices.

whereas $h_L \leq n \cdot \log_2 n$ and $h_0 = 0$. Thus, Eq. (9) follows. The furthermore claim is shown by observing that the first part of the proof remains intact, whereas the second part (which establishes Eq. (8)) implies that $\sum_{\{r,s\}\in E'} \Delta_\ell(r,s) \leq \tau \cdot |E|$, whereas $|E| = O(|E'|)$. ■

Using a global (GW-type) analysis. Using Lemmas 1 and 2, we infer that if the tester rejects with small probability (i.e., with probability smaller than 1/3), then almost all pairs of vectors that correspond to edges are almost anti-polar; in other words, the average inner-product of pairs of vectors that correspond to edges is close to -1. We now show that in this case the graph is close to being bipartite.

Lemma 3 (inner products close to -1 implies closeness to Bipartite (adapted from [6])): *Suppose that the average inner-product of vectors that correspond to neighboring vertices in $G = ([n], E)$ is at most $-1+\tau$; that is,*

$$\sum_{\{r,s\}\in E} \langle \vec{\rho_r}, \vec{\rho_s} \rangle \leq -(1-\tau)\cdot |E|. \tag{11}$$

Then, G is $\sqrt{\tau}$-close to bipartite. Furthermore, if Eq. (11) *holds when replacing E by $E' \subseteq E$, then G is $(\sqrt{\tau}+\rho)$-close to bipartite, where $\rho \stackrel{\text{def}}{=} \frac{|E\setminus E'|}{dn/2}$.*

Proof: Letting $\theta(r,s)$ denotes the angle between $\vec{\rho_r}$ and $\vec{\rho_s}$, note that $\langle \vec{\rho_r}, \vec{\rho_s} \rangle = \cos(\theta(r,s))$. Following [6], while letting $\delta(r,s) \stackrel{\text{def}}{=} 1 + \langle \vec{\rho_r}, \vec{\rho_s} \rangle \geq 0$ denote the distance of $\langle \vec{\rho_r}, \vec{\rho_s} \rangle$ from -1, we observe that selecting uniformly a vector $\vec{u}$ in the unit sphere and assigning colors to $\vec{\rho_r}$ and $\vec{\rho_s}$ according to the sign of their inner product with $\vec{u}$, the probability that they are assigned different colors is

$$\begin{aligned}
\frac{\theta(r,s)}{\pi} &= \frac{\arccos(\langle \vec{\rho_r}, \vec{\rho_s} \rangle)}{\pi} \\
&= \frac{\arccos(-1+\delta(r,s))}{\pi} \\
&\geq 1 - \frac{(\pi/2)\cdot\sqrt{2\delta(r,s)}}{\pi}
\end{aligned}$$

where the last inequality uses the fact that, for $x \in [0,1]$, it holds that $\arcsin(x) \leq (\pi/2)\cdot x$ (see [8, Sec. 3.4.1]).[12] Hence, the probability that the edge $\{r,s\}$ violates this random assignment (i.e., both r and s are assigned the to same side) is at most $\sqrt{\delta(r,s)/2}$. Using $\sum_{\{r,s\}\in E'} \delta(r,s) \leq \tau \cdot |E'|$, it follows that the expected number of edges (in E') that violate the random assignment is at most $\sqrt{\tau/2}\cdot|E'|$. The furthermore claim follows. ■

[12] We also use well-known equalities such as $\arccos(-x) = \pi - \arccos(x)$ and $\arccos(x) = \arcsin(\sqrt{1-x^2})$. Hence, we have

$$\begin{aligned}
\arccos(-1+\delta(r,s)) &= \pi - \arcsin((1-(1-\delta(r,s))^2)^{1/2}) \\
&\geq \pi - \arcsin((2\delta(r,s))^{1/2}) \\
&\geq \pi - (\pi/2)\cdot(2\delta(r,s))^{1/2}
\end{aligned}$$

Conclusion. Combining the three claims (while using the furthermore parts of Lemmas 2 and 3 (with $|E'| \geq (1-\epsilon/2) \cdot |E|$)), we conclude that, for sufficiently large $m = O(\sqrt{n}/\epsilon^2)$, if the tester rejects a graph with probability smaller than 1/3, then the graph is ϵ-close to being bipartite. Specifically, we use $\tau = (\epsilon/2)^2$ (and $\rho = \epsilon/2$) in Lemma 3, while using $\tau = \epsilon^2/O(1)$ in Lemmas 1 and 2. Hence, the query complexity of the tester is $O(m/\epsilon) \cdot L = O(\epsilon^{-5} \cdot \sqrt{n} \cdot \log n)$, since $L = O(\epsilon^{-2} \cdot \log n)$.

Digest. As stated in [1], "stability" in the sense of small slackness (as defined in the proof of Lemma 2) is central to the "global" analysis. Achieving this notion of stability requires random walks of length $O(\epsilon^{-2} \cdot \log n)$. This notion of stability replaces a notion of (weak) expansion that is used in [4].

While the notion of parity-conforming collision probability (i.e., pccp) and the difference vectors (i.e., $\vec{\delta}$'s) appear more natural than the mmpp and the vectors $\vec{\rho}$'s, we were not able to carry out the analysis with the former, and were forced to use the latter. Specifically, we failed to prove a result analogous to Lemma 2 for the former (i.e., relate the inner product of the $\vec{\delta}$'s to the pccp). Ironically, even relating mmpp to the rejection probability of the tester (i.e., Lemma 1) is easier than relating pccp to the rejection probability (cf., [2, Sec. 9.4.1]).

We are not really satisfied by the current state of affairs. First, the alternative we work with (i.e., mmpp and $\vec{\rho}$) is less appealing than the one that we cannot analyze well (i.e., pccp and $\vec{\delta}$). But, more importantly, we find the proof of Lemma 2 somewhat uninituive (e.g., involving too many technical steps that have no intuitive justification). Hence, we hope that others will succeed in the place we failed.

4 The dependence on the degree-bound and beyond

We neglected the effect of the degree-bound (i.e., d) on the query complexity of the tester, but this dependence is important when one views d as a parameter, which is crucial when extending the result to the general graph model. Note that we explicitly referred to the value of d only when postulating, right after stating Eq. (9), that we assume $|E| = \Omega(dn)$ (and justifying it (in Footnote 11) by the assumption that $d = O(1)$ and $|E| \geq n/2$). However, we can justify $|E| = \Omega(dn)$, without assuming that $d = O(1)$, by assuming (w.l.o.g.) that all vertices in the graph have degree at least $d/2$. (This can be justified by a reduction that augments the original n-vertex graph with n auxiliary vertices that are each connected to a corresponding original vertex by d parallel edges.)[13] In addition, we implicitly assumed that $d = O(1)$ when presuming that we can easily sample random edges. However, when all vertices have degree at least $d/2$, sampling edges can be done by rejection sampling (i.e., sampling uniformly a vertex v and an index $i \in [d]$ and outputting the i^{th} neighbor of v if such exists).

Alternatively, we can just replace the upper bounds on the sums that appear in Lemmas 2 and 3 such that error terms related to $|E|$ are replaced by error terms related to $dn/2$. Specifically, the main part of Lemma 2 would read *If* $\sum_{\{r,s\}\in E} \texttt{mmpp}_\ell(r,s) \leq \tau \cdot dn/2$, *then* $\sum_{\{r,s\}\in E} \langle \vec{\rho_{\ell,r}}, \vec{\rho_{\ell,s}} \rangle \leq -(1-O(\tau)) \cdot dn/2$, whereas Eq. (11) in Lemma 3 is replaced by $\sum_{\{r,s\}\in E} \langle \vec{\rho_r}, \vec{\rho_s} \rangle \leq -(1-\tau) \cdot dn/2$. (Note that if at most $\tau \cdot dn/2$ edges are "monochromatic", then the graph is τ-close to being 2-colorable.) (Also, in this case, sampling $[n] \times [d]$ without repeats suffices.)

[13] Here we capitalize on the fact that the algorithm and its analysis allows parallel edges.

In any case, having established a complexity bound that is independent of d, we reduce testing the graph $G = ([n], E)$ in the general graph model to testing a related graph in the bounded-degree model (with degree bound $d = O(|E|/n)$). For details see [7] or [2, Sec. 10.2.2].

5 Concrete differences viz a viz [1]

As stated before, the presentation in [1] is "technically-oriented" (e.g., it is presented in terms of the Markov operator, and offers a measure theoretic framework and an SDP perspective). In contrast, we prefer a conceptually-oriented presentation that utilizes direct arguments and a minimal amount of "technology". We also deviate from [1] in some concrete aspects, discussed here. Two of the aspects were already mentioned in our main presentation.

Starting with the bounded-degree graph model and ignoring the dependence on the degree bound. The presentation in [1] is aimed at obtaining the most general result in one step (i.e., they work directly in the "general graph model"). In contrast, we prefer to first focus on the bounded-degree graph model and neglect the dependence on the degree bound (treating it later (as a ramification)).

Taking random walks from the endpoints of a random edge. The analysis in [1] refers to the original description of the tester (as in [4]) in which the start vertices (for $O(\sqrt{n})$ random walks) are selected at random. In contrast, we prefer to analyze a version in which start edges are selected at random and walks are generated from each of their endpoints.

The query complexity is somewhat better. We obtain a query complexity bound of $O(\epsilon^{-5} \cdot \sqrt{n} \cdot \log n)$, whereas the bound in [1] is $O(\epsilon^{-16} \cdot \sqrt{n} \cdot \log n)$. We do not know how much of the gap is due to the difference in presentation and how much is due to the authors of [1] not caring about this issue. (For perspective, we mention that the furthermore parts of Lemmas 2 and 3 allow to reduce the power of $1/\epsilon$ (in the query complexity bound) from 6 to 5.)

6 The interactive proof of proximity

As stated upfront, the analysis in Section 3 yields an interactive proof system for Bipartiteness. Such an interactive proof was previously known only for the rapid mixing case [9]. Essentially, we follow their strategy, but having Lemmas 2 and 3 at our disposal we can use the following.[14]

Corollary 4 (if G is far from bipartite, then $\texttt{mmpp}$ is large): *For $L = O(\epsilon^{-2} \cdot \log n)$, the following holds. If G is ϵ-far from being bipartite, then $\sum_{\{r,s\}\in E} \texttt{mmpp}_\ell(r, s) = \Omega(\epsilon^3 \cdot |E|)$ holds for most $\ell \in [L]$.*

Using Corollary 4, we first present an interactive proof system with perfect completeness and soundness error that is bounded away from 1. On input a graph $G = ([n], E)$ and proximity parameter $\epsilon > 0$, the parties proceed as follows.

[14] Actually, the hypothesis implies that either $\texttt{mmpp}_\ell = \Omega(\epsilon^2)$ holds for at least an $\epsilon/2$ fraction of the edges (and every ℓ) or $\sum_{\{r,s\}\in E} \texttt{mmpp}_\ell(r, s) = \Omega(\epsilon^2 \cdot |E|)$ holds for most $\ell \in [L]$.

1. The verifier selects uniformly at random $\ell \in [L]$ and an edge $\{r, s\} \in E$. It then selects uniformly at random $u \in \{r, s\}$, and takes a ℓ-step lazy random walk from u, denoting the vertex reached at the end of the walk by v and the parity of the walk by σ.

2. The verifier sends ℓ, $\{r, s\}$, v and σ to the prover.

3. The prover responds with $w \in \{r, s\}$ that maximizes $p_\ell^\sigma(w, v)$ (i.e., the probability that an ℓ-LRW starting at w reaches v with parity σ).

4. The verifier accepts if and only if $w = u$.

The verifier's time complexity is $O(L) = O(\epsilon^{-2} \cdot \log n)$. If G is bipartite, then, for every message $(\ell, \{r, s\}, v, \sigma)$ of the verifier, it holds that $p_\ell^\sigma(u, v) > 0$ whereas $p_\ell^\sigma(u', v) = 0$ for every u' that neighbors u. Hence, in case G is bipartite, the verifier accepts with probability 1. On the other hand, if G is ϵ-far from bipartite, then the verifier rejects with probability at least

$$\mathbf{E}_{\ell \in [L]} \mathbf{E}_{\{r,s\} \in E} [\mathtt{mmpp}_\ell(r, s)] \quad = \quad \Omega(\epsilon^3)$$

because the probability that the verifier sent $(\ell, \{r, s\}, v, \sigma)$ and rejected upon receiving the prover's response is at least $\min(p_\ell^\sigma(r, v), p_\ell^\sigma(s, v))$. Repeating the foregoing system for $O(1/\epsilon^3)$ times (in parallel), we get

Theorem 5 (an interactive proof of proximity for Bipartiteness): *There exists a one-round interactive proof of proximity for Bipartiteness in which the verifier time complexity is $O(\epsilon^{-5} \cdot \log n)$.*

Acknowledgements

I am grateful to Yumou Fei for extensive discussions in which he explained to me various aspects of [1]. In particular, the proof of Lemma 2 as well as Footnote 9 were adapted from his emails to me. I am also grateful to Guy Rothblum for numerous discussions that evolved around getting intearctive proofs of proximity for Bipartiteness (in the non-rapid-mixing case).

References

[1] Yumou Fei and Ronitt Rubinfeld. Testing Bipartiteness in Logarithmic Rounds. In `arXiv:2606.13583` [cs.DS].

[2] Oded Goldreich. *Introduction to Property Testing*. Cambridge University Press, 2017.

[3] Oded Goldreich and Dana Ron. Property testing in bounded degree graphs. *Algorithmica*, pages 302–343, 2002. Extended abstract in *29th STOC*, 1997.

[4] Oded Goldreich and Dana Ron. A sublinear bipartite tester for bounded degree graphs. *Combinatorica*, Vol. 19 (3), pages 335–373, 1999. Extended abstract in *30th STOC*, 1998.

[5] Oded Goldreich and Laliv Tauber. Testing in the Bounded-Degree Graph Model with Degree Bound Two. In *Computational Complexity and Local Algorithms*, LNCS15700, Springer, pages 373–382, 2025.

[6] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, Vol. 42 (6), pages 1115–1145, 1995.

[7] Tali Kaufman, Michael Krivelevich, and Dana Ron. Tight Bounds for Testing Bipartiteness in General Graphs. *SIAM Journal on Computing*, Vol. 33 (6), pages 1441–1483, 2004.

[8] Dragoslav S. Mitrinovic. *Analytic Inequalities.* Springer-Verlag, Berlin, 1970.

[9] Guy N. Rothblum, Salil Vadhan, and Avi Wigderson. Interactive Proofs of Proximity: Delegating Computation in Sublinear Time. In *45th ACM Symposium on the Theory of Computing*, pages 793–802, 2013. Full version available from Guy Rothblum's website.[15]

[15] See `https://guyrothblum.wordpress.com/wp-content/uploads/2014/11/rvw13.pdf`

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il