RnD Report: Spring 2024

Sooraj A P, Krishna N Agaram
Guide: Prof. Akash Kumar

Contents

1 Introduction

The main topic of study will be property testing algorithms for an important NP-Complete problem - 3-coloring. We first present results from [BOT02] that prove why testing whether a graph is 3-colorable or far from it is 'HARD', that is, it requires a number of queries linear in the size of the graph. Following that, we also see that the problem is hard even on expanding graphs ([PY23]).

1.1 Notation and Preliminaries

A property testing algorithm $\mathcal{A}$ for a property $\mathcal{P}$ of elements belonging to some set $\mathcal{S}$ is one that given a proximity parameter ε accepts elements of $\mathcal{S}$ that satisfy $\mathcal{P}$ with probability at least $2/3$ and reject those that are ε-far from having the property with probability $2/3$. note that we do not care about the output of the algorithm on inputs that do not satisfy one of the above conditions. Our notion of 'closeness' is as follows: an element $G \in \mathcal{S}$ is said to be ε-close to having the property $\mathcal{P}$ if G can be made to have the property $\mathcal{P}$ by modifying at most an ε fraction of the underlying elements in G.
In our case, we are interesting in the universe of n-vertex graphs with a maximum degree of d, these will form the set $\mathcal{S}$. Our property of interest is that of 3-colorability and we say that $G \in \mathcal{S}$ is ε-close to being 3-colorable if it differs from a 3-colorable graph by at most $\varepsilon nd/2$ edges.

2 The One-Sided Lowerbound

Theorem 1. *For each $\alpha > 0$, there exist constants d and $\delta > 0$ such that if A is a one-sided error tester for degree d graphs that distinguishes between graphs that are 3-colorable from those that are $(1/3 - \alpha)$-far from being 3-colorable, then A requires at least δn queries.*

This is an interesting regime for property testing since no graph is more than $1/3$-far from being 3-colorable, making such a result even more surprising. To prove the result, we try and construct graphs that are indeed far from being 3-colorable but any subgraph consisting of at most δn vertices is 3-colorable. To do so, we go by probabilistic constructions of our distributions for graphs as well as instances of E3-LIN which we use later for reductions. The 'FAR' distribution $\mathcal{G}_{FAR}$ of graphs we shall consider are the "configuration graphs" whose distribution is generated as follows (assume that the number of vertices is even):

- For rounds $1 \leq i \leq d$, do the following: pick a random matching on the vertex set, call it C_i.

- Set the edge set as $E = \cup_{i=1}^{d} C_i$.

(Remark: This is not exactly the configuration model, but with very high probability $(1 - o(1))$ they behave the same. The only difference between the distribution here and the one in the config model is that the one we employ completely avoids self edges.) The distribution thus obtained has multigraphs (we shall call these graphs for convenience) each having a maximum possible degree of d, and these graphs are simple and expand with high probability. The authors first show that a graph G sampled from $\mathcal{G}_{FAR}$ has, with high probability, close to a $1/3$ fraction of violating edges, proving that these graphs indeed form a FAR distribution. This is done using the fact that for each tripartition of the vertex set, the process in which edges of G are revealed one by one forms a Doob Martingale and then using the Azuma-Hoeffding inequality to bound the deviation of the number of internal-edges (ones that do not cross from one partition to another) from the expected value, which is approximately $1/6dn$, which is a $1/3$ fraction of the expected number of edges in G. The conclusion of this analysis is that for any given $\alpha > 0$, with very high probability the fraction of internal edges exceeds $(1/3 - \alpha)$. If X is the random variable denoting the number of internal edges corresonding to some tripartition in some $G \sim \mathcal{G}_{FAR}$, then

$$\mathbb{P}[X < (1/6 - \alpha)dn] \leq e^{-(\alpha - o(1))^2 dn}.$$

Using that each G has 3^n tripartitions, we see that choosing $d > (\ln(3)/\alpha^2)$ ensures that after the union bound, for any 3-coloring, with probability $1 - o(1)$, we have at least $(1/6 - \alpha)dn$ violating(internal) edges. We now obtain $\bar{G}$ from G by dropping all mutiple edges (thus making it a simple graph). By Markov's inequality, the probability that we have more than $d \log n$ distinct repeated edges is $o(1)$ and so the number of edges in $\bar{G}$ is $nd/2 - o(n)$, so the same choice of d makes the same result hold for $\bar{G}$ as well.

Now that we have established that this distribution $\mathcal{G}_{FAR}$ is indeed a FAR one, we shall see that each small subgraph has very few edges and as a consequence is 3-colorable whp.

Lemma 2. *For every $K > 1$, there exists a $\delta > 0$ such that with probability $1 - o(1)$ all the graphs $\bar{G}|_S$ with $S \subseteq V$, $|S| < \delta n$ have at most $K|S|$ edges.*

The proof of the above goes via the probabilistic method. The upshot to this lemma is that by setting $K = 3|S|/2$, we get the following.

Theorem 3. *For every $\alpha > 0$, there exists $\delta > 0$ such that with probability $1-o(1)$, $\bar{G}$ is $(1/3-\alpha)$-far from being 3-colorable but all subgraphs on fewer than δn vertices are 3-colorable.*

As indicated above, this is a consequence of Lemma 2 and the fact that graphs sampled from the distribution $\mathcal{G}_{FAR}$ are $(1/3 - \alpha)$-far from being colorable with probability $1 - o(1)$. More importantly, Theorem 3 states that the instance we constructed is such that any subgraph searched by any δn query algorithm appears 3-colorable, and thus cannot correctly identify the distribution, establishing the desired lower bound.

3 Two-sided Lower Bound

The approach towards proving this bound proceeds via Yao's Lemma, that is, we construct distributions $\mathcal{G}_{3COL}$ and $\mathcal{G}_{FAR}$ such that they are indistinguishable for any tester of sublinear query complexity. Towards this, we first come up with two distributions on E3-LIN instances which we then reduce to 3-SAT and then to instances of 3-coloring. The main result is the following:

Theorem 4. *There exist constants δ, ε such that for any two-sided tester A for distinguishing degree d 3-colorable graphs from those that are $\varepsilon-$far from being 3-colorable have a query complexity of at least δn.*

To start off, we construct distributions for E3-LIN, which on reducing to 3-SAT, and then to 3-coloring instances, are hard to distinguish by any algorithm that is allowed access to only a sublinear number of vertices. For the FAR distribution, we take inspiration from the one for the one-sided tester, at least for constricting the instances for E3-LIN. Just as for the one-sided case, consider a "configuration" distribution which we shall call $\mathcal{H}_{FAR}$ on 3-hypergraphs on n vertices, where we assume that n is a multiple of 3 for convenience. Using identical analysis as for the case where we had graphs instead of hypergraphs, we see that hypergraphs sampled from this distribution also satisfy some properties very similar to Lemma 2.

Lemma 5. *For every $K > 1/2$, there exists $\delta > 0$ such that with probability $1 - o(1)$ all 3-hypergraphs $H|_S$ where $H \sim \mathcal{H}_{FAR}$ and $|S| \leq \delta n$ have at most $K|S|$ edges.*

Using the above, we can prove the following theorem.

Theorem 6. *For every $c > 0$ there exists a $\delta > 0$ such that for every n there exists a matrix $A \in \{0,1\}^{n \times cn}$ with n columns and cn rows, such that each row has exactly three non-zero entries, each column has exactly 3 c non-zero entries, and every collection of δn rows is linearly independent.*

Proof. In Lemma 5, let $|S| < 3\delta n$ and $K = 2/3$, then the sub-hypergraph induced has fewer than $2|S|/3$ hyperedges. Consider the incidence matrix of this hypergraph. The columns of A correspond to vertices of H, the rows of A correspond to hyperedges of H, and $A_{ve} = 1$ if and only if $v \in e$. Suppose that there is a set R of δn rows of A (or hyperedges of H) that are linearly dependent. We may assume that R is a minimal set with this property. Let $S \subseteq V(H)$ denote the set of vertices incident to hyperedges in R, so that $|S| \leq 3\delta n$. By minimality of R, every element of S must appear in at least two rows of R. Therefore, R contains at least $2|S|/3$ hyperedges which is a contradiction to Lemma 5. ∎

Now with this construction of the FAR instances for E3-LIN in hand, we make some definitions that allow the reductions we have in mind to go through seamlessly.

Definition 7. *We say that a CSP (Constraint Satisfaction Problem) on m clauses is $(\delta, 1-\varepsilon)$-satisfiable if any assignment satisfies at least δm clauses but no assignment satisfies more than $(1-\varepsilon)m$ clauses.*

Definition 8. *A map φ from a decision problem A to a decision problem B is said to be a gap-preserving local reduction if there exist constants c_1, c_2 such that the following hold:*

1. *φ maps a YES instance of A to a YES instance of B.*
2. *φ maps an ε-FAR instance of A to a ε/c_1-FAR instance of B.*
3. *The answer to any query to $\varphi(x)$ can be computed by making c_2 queries to the parent instance x [This is the "local" property.].*

Via the standard reduction from a k-CNF to a 3-CNF, we conclude that there exist gap preserving local reductions from any CSP f defined on n variables and m clauses into a 3-CNF with $O(n+m)$ variables and $O(m)$ clauses. It is also easily verified that gap-preserving local reductions are closed under composition. We require the above property as we plan to find reductions φ from E3-LIN to 3-SAT and ψ from 3-SAT to 3-coloring and want $\varphi \circ \psi$ to be a gap-preserving local reduction so that we get the required coloring instance.
We now want to investigate how these reductions behave on the above defined $(\delta, 1-\varepsilon)$ CSPs. After all, our instances of coloring are such CSPs since any graph is 1/3-close to colorable ($\delta = 2/3$) and we require that the graph is ε-far. We state the following result without delving into the proof, though the proof is detailed in [BOT02].

Lemma 9. *If $\varphi : A \to B$ is a gap-preserving local reduction with distortion constants c_1, c_2 and f is a $(\delta, 1-\varepsilon)$-satisfiable CSP, then $\varphi(f)$ is a $\left(\frac{\delta}{c_2}, 1 - \frac{\varepsilon}{c_1}\right)$-satisfiable CSP.*

Note: We were not able to complete the proof for the above lemma when φ is not the standard reduction from E3-Lin to 3-SAT. But since we only need that reduction, there is no break in the result.
Now onto the reduction from 3-SAT to 3-coloring! We require a new construction of the corresponding graph since the standard reduction does not always preserve the property that the graph must have bounded degree if the parent 3-CNF does.

3.1 GPLR for 3-SAT to 3-coloring

Given a 3-SAT instance φ with m clauses and n variables from a family with a bounded number of occurrences of each variable, we construct a *bounded degree* graph $\mathcal{G}_\varphi$ preserving gap. In particular, we establish the following theorem.

Theorem 10. *If φ is an instance of a $(\delta, 1-\varepsilon)$ 3-SAT CSP with $m > n$, then the graph $\mathcal{G}_\varphi$ is an instance of a $(\delta', 1-\varepsilon')$ 3-COL CSP, with*

$$\delta' = \frac{\delta}{bc} \quad \text{and} \quad \varepsilon' = \frac{\varepsilon}{8}.$$

(b and c are constants that we shall define shortly)

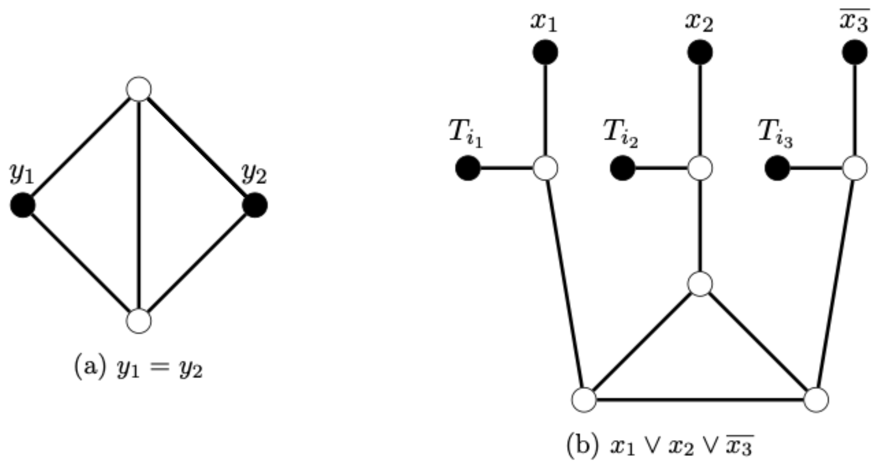

Figure 1: (a) The equality gadget. Notice that the two nodes y_1 and y_2 are forced to have the same color in any proper three coloring of the gadget. (b) The clause gadget. With the mapping $[0 \mapsto F, 1 \mapsto T]$, any satisfying assignment to the clause results in a proper coloring of the gadget. Further, any assignment that does not satisfy the clause cannot properly color the rest of the gadget.

The construction is detailed nicely in [PY23]; we reproduce it here for completeness.
We are given a formula φ on n variables with m clauses. Each variable occurs at most k times in φ. Please keep in mind that we would like to color the graph $\mathcal{G}_\varphi$ in three colors. Our three colors will be called true, false and dummy, abbreviated T, F and D.
We make heavy use of the two gadgets in Figure 1.
The vertex set of $\mathcal{G}_\varphi$:

1. Construct three sets of nodes D_i, T_i, F_i where $1 \le i \le 2kn$. These are the *color class* vertices. A coloring of $\mathcal{G}_\varphi$ will have the D_is colored D, the T_is T and the F_is F.

2. Construct for each $1 \le i \le n$ the nodes ${x_i}^1, \dots, {x_i}^k$ and ${\bar{x}_i}^1, \dots, {\bar{x}_i}^k$. There are $2kn$ total *literal* nodes. A coloring of $\mathcal{G}_\varphi$ will have these nodes colored either T or F according to some 0/1-assignment to the literals. ${x_i}^j$ and ${\bar{x}_i}^j$ will always have opposite colors.

3. We also construct more nodes that arise as internal nodes in some gadget (these comprise the white nodes in Figure 1). We call these A(dditional)-nodes. A coloring of $\mathcal{G}_\varphi$ will have these nodes' colors forced so as to make each gadget colorable.

For each literal node ${x_i}^j$, we call the nodes D_ℓ, T_ℓ, F_ℓ where $\ell := k(i-1)+j$ its *corresponding* color class nodes. Similarly, for ${\bar{x}_i}^j$, the corresponding color nodes have $\ell := kn + k(i-1) + j$. This correspondence is simply a bijection between the set of $2kn$ literal nodes and each color class. We also say ${x_i}^j$ represents x_i and ${\bar{x}_i}^j$ represents $\bar{x}_i$.
The edge set of $\mathcal{G}_\varphi$:

1. Pick an d-regular *expander* on $2kn$ nodes (d is some other constant > 0). For each edge $\{i, j\}$ of the expander, construct an equality gadget between D_i and D_j ($D_i = y_1, D_j = y_2$), similarly between T_i, T_j and F_i, F_j. This adds some more vertices. Since expanders are connected, all D_is (sim. T_i, F_i) will be assigned the same color in any proper coloring.

2. Further for each $1 \le i \le 2kn$, add edges $\{D_i, T_i\}$, $\{T_i, F_i\}$, $\{F_i, D_i\}$. This ensures they all have different colors, WLOG we can let the D_is have color D (sim. T_is get T and F_is get F).

3. Connect each literal node to its corresponding D-color class node to ensure a literal is only colored T/F. For each $1 \le i \le n, 1 \le j \le k$, add edge $\{{x_i}^j, {\bar{x}_i}^j\}$.

4. Finally, for each clause $\ell_1 \vee \ell_2 \ell_3$ (each ℓ_s is x_i or $\bar{x}_i$ for some i), add a clause gadget using fresh (not used in any clause gadget so far) nodes representing literals ℓ_s and each chosen fresh node's corresponding T-color class node. This also adds some more vertices.

That completes the construction of the graph $\mathcal{G}_\varphi$. It is clear that the maximum degree of a vertex in this graph (the graph is very not-regular, note) is some function of k and d - a constant, call it b. The number of vertices n' is also at most cn for some constant c depending on k and d. The number of edges m' in $\mathcal{G}_\varphi$ is thus at most $bn' \le bcn$.
We now prove that it is a GPLR.

Proof. The easy direction is to show how many edges we can satisfy via a coloring. Pick an assignment that satisfies δm clauses of φ, and color the graph using the natural coloring. At least δm clause gadgets are correctly colored. Finish by noting that $\delta' m' \le \delta n \le \delta m$, so surely at least $\delta' m'$ edges are satisfied.
Now, consider a coloring for $\mathcal{G}_\varphi$ that maximizes the number of satisfied edges. The idea is as follows: call the monochromatic edges bad, we shall be pessimistic and throw away all gadgets with at least one bad edge, and continue to throw away edges and vertices till we get back a (fully colorable) graph $\mathcal{G}_{\varphi'}$ that is the transformation of another formula φ'. φ' will be the conjunction of a *huge* subset of the clauses of φ; we leverage that the size of this subset cannot of course exceed $(1-\varepsilon)m$.
Let $\gamma 2kn$ be an upper bound on the number of bad edges of each particular type (described below).

1. Suppose that $\gamma 2kn$ equality gadgets were thrown away from the D-color class subgraph. By expansion, we can recover a connected component after throwing away the edges that has size at least $(1-\gamma)2kn$. Throw away the rest of the D-nodes. So the same for the T and F subgraphs.

2. For each i, if any one of the three edges $\{D_i, T_i\}$, $\{T_i, F_i\}$, $\{F_i, D_i\}$ is now removed, delete all three of D_i, T_i, F_i. The deletion happens iff one of D_i, T_i, F_i was deleted, and each such deletion additionally removes at most the other two nodes. In total, at most $9\gamma 2kn$ nodes are deleted ($3\gamma 2kn$ from each of the three classes).

3. At this point the color class nodes left have their old structure and their subgraph is fully colorable. We next remove literal nodes that are adjacent to bad edges.

4. Each literal node whose corresponding D-node has been removed is removed. At most $3\gamma 2kn$ literal nodes removed. All literals adjacent to the at most $\gamma 2kn$ bad edges going between a $D/T/F$ node to the literal are removed. At most $\gamma 2kn$ of these.

5. Any literal node ${x_i}^j$ $({\bar{x}_i}^j)$ whose complement literal ${\bar{x}_i}^j$ $({x_i}^j)$ is removed. So far at most $4\gamma 2kn$ literals are removed, so this step removes at most an equal number more.

6. At this point all remaining literal nodes have their old structure. It remains to remove some clauses.

7. Any clause gadget with a bad edge is removed. At most $\gamma 2kn$ of these. Clause gadgets with a T node or literal node that is non-existent are removed - at most $3\gamma 2kn + 8\gamma 2kn$ more (notice here we are using the fact that a particular T/literal node participates in at most one clause) clauses are removed. Note that only the A-vertices and edges of the clause gadgets are removed; the T and literal nodes involved remain in the graph.

8. At this point the graph is the transformation $\mathcal{G}_{\varphi'}$ (possibly with some extra vertices) of φ' that is φ minus the clauses corresponding to the removed clause gadgets. The extra vertices do not affect the fact that all clauses in φ' are satisfiable.

$11\gamma 2kn$ clauses have been removed, and the resulting graph is fully 3-colorable, implying that all clauses left over are satisfiable. This means, by the promise on φ, that

$$22\gamma kn \ge \varepsilon m.$$

However, we had $\gamma 2kn \times 3 + \gamma 2kn + \gamma 2kn = 10\gamma kn$ bad edges (skirting over the fact here that $\gamma 2kn$ was not exactly the number of bad edges of each type; it seems to be simply a technical issue, one can use $\gamma_r 2kn$ for bad edges of type r and sum those up, etc). Since this was the minimum possible bad edges, at least

$$10\gamma kn \ge \frac{10}{22}\varepsilon m \ge \frac{1}{3}\varepsilon n \ge \frac{\varepsilon}{3bc} m' = \varepsilon' m'$$

edges are bad in any coloring of $\mathcal{G}_\varphi$. The argument in [BOT02] yields a much larger $\varepsilon' = \varepsilon/8$, but we are not sure we understand why the argument works; it seems like the conclusions drawn from each inequality in their argument is necessary but not sufficient. ■

We are now done with the lower bound proof for two-sided testers. We restate the theorem here for convenience.

Theorem 11. *There exist constants δ, ε such that for any two-sided tester A for distinguishing degree d 3-colorable graphs from those that are $\varepsilon-$far from being 3-colorable have a query complexity of at least δn.*

Proof. We set up first two distributions over E3-Lin instances. Pick a matrix A satisfying the conditions in Theorem 6, and consider the systems $Ax = b$ (with b uniformly drawn from $\{0,1\}^{cn}$) and $Ax = Ab'$ (with b' drawn uniformly from $\{0,1\}^n$). An instance drawn from the first distribution is w.h.p a $(1/2, 1/2 + \alpha)$-CSP, while from the second is fully satisfiable. The important point about these distributions is that restricting the systems to any subset of δn rows makes the distributions statistically indistinguishable.
Applying the GPLR from E3-Lin to (E)3-SAT gives us two distributions, the first being $(7/8, 7/8 + \alpha/4)$-CSP and of course, the second one fully satisfiable. Applying the GPLR from E3-SAT to 3COL yields one distribution which is $(\star, 1-\varepsilon)$-CSP for some constant $\varepsilon > 0$ and the second again, fully satisfiable.
Suppose now that there is a sublinear two-sided tester for satisfiable vs ε-far from 3 colorable. The first distribution, call it FAR, is w.h.p ε-far (since w.h.p the E3-Lin is $(1/2-\alpha)$-far). Call the second distribution 3COL. Let $\mathcal{A}(I)$ denote the event that the algorithm's output on instance I is YES. We have

$$\Pr_{I \sim \text{FAR}}[\mathcal{A}(I)] \leq \Pr[\mathcal{A}(I) | I \text{ is far}]\left(1 - \frac{1}{n}\right) + \frac{1}{n} \leq \frac{1}{3} + \frac{2}{3n}.$$

Since $\mathcal{A}$ is a sublinear time tester, $I \sim$ FAR and $I \sim$ 3COL look the same (a sublinear number of queries on the graph, by GPLR properties, depends exclusively on a sublinear number of rows in the underlying E3-Lin instances, and so the distributions of subgraphs received is identical). In particular, this means that

$$\Pr_{I \sim \text{FAR}}[\mathcal{A}(I)] = \Pr_{I \sim \text{3COL}}[\mathcal{A}(I)].$$

The latter quantity, because each $I \in$ 3COL is 3-colorable, is at least $2/3$. This yields

$$1/3 + 2/3n \geq \Pr_{I \sim \text{FAR}}[\mathcal{A}(I)] = \Pr_{I \sim \text{3COL}}[\mathcal{A}(I)] \geq 2/3,$$

contradiction. Poof. ∎

4 Extensions

It is known ([Bol88]) that the random graphs generated by the union of random matchings are w.h.p expanding, which means that there is an expander whose δn-vertex subgraphs are all colorable; this yields a one-sided lower bound for 1 vs $1/3 - \alpha$-far from 3 colorable on expanders. The graph $\mathcal{G}_\varphi$ was also proved to be an expander in [PY23], yielding a two-sided lower bound for 1 vs ε-far from 3-colorable even on expanders.
It is of interest to find an analogous two-sided tester lower bound for the 1 vs $1/3 - \alpha$ case. To establish such a bound, we set out to find distributions $\mathcal{D}_{FAR}, \mathcal{D}_{3COL}$ such that for any algorithm ALG:

$$\sum_{T \text{ - a } q\text{-query transcript}} [\mathbb{P}_{G \sim \mathcal{D}_{FAR}}(ALG \text{ gets } T \text{ on } G) - \mathbb{P}_{G \sim \mathcal{D}_{3COL}}(ALG \text{ gets } T \text{ on } G)] < \frac{1}{3}.$$

4.1 Our attempts at proving the lower bound

The first thing to take note of while addressing the questions that arise relating to a two-sided lower bound is the oracle which gives information about the graph to the algorithm that tries to test for 3-colorability.
The oracle: Given access to the graph $G = (V, E)$, when queried

1. Sample a vertex v uniformly at random.
2. Pick a vertex u among the neighbours of v uniformly at random and return the edge $e = (u, v)$.

Before trying other constructions, we first emphasise why the distributions considered by [BOT02], i.e the configuration model, fails. The problem with the configuration model when trying to prove
An approach we tried was the following: Let $\mathcal{D}_{FAR}$ be the instances from Section 2 and $\mathcal{D}_{3COL} = \{G \setminus CE : G \sim \mathcal{D}_{FAR}, CE \text{ is the set of contradicting edges in } G\}$. By this construction, we have that eny graph in $\mathcal{D}_{3COL}$ has atleast 2/3rds the edges of one particular corresponding graph in $\mathcal{D}_{FAR}$. This of course, is not completely sufficient as a construction, so we employ the "blow-up" methodology. The blow-up method is a standard approach to prove lower bounds for property testing problems. It involves taking a base graph H (which we shall refer to as the *host graph*) and 'blowing

up' each vertex of H into a copy of some other graph, and adding a suitable matching. The host graph, the graphs we place at the old nodes, as well as the matching would depend on the problem we are trying to approach, and the goal is to use these to come up with hard instances and thus establish lower bounds.

A failed attempt: Let the host graph be a union of d random matchings, in other words, a config graph. In the YES instance, replace each node with a graph drawn from $\mathcal{D}_{3COL}$ and in the FAR instance, replace each node with a graph sampled from $\mathcal{D}_{FAR}$. Between each blown-up node (we shall informally refer to these graphs as 'blobs'), add a perfect matching that preserves the colorability. The goal of this approach was the following: try and hide the contradictions to 3-colorability inside the blobs for the FAR instance, and thus to probe and find a contradiction, an algorithm must find a witness for non 3-colorbaility within a blob. However, this approach is fundamentally wrong, as the same proof may then be employed (by changing the blobs to 'bipartite graphs' and 'graphs that are FAR from being bipartite' in the YES and FAR cases respectively) to prove a $\Omega(n)$ query lower bound for testing bipartiteness on bounded degree graphs (of course, the graphs generated above are indeed of bounded degree), which is known to be false as [GR98] shows a $\Omega(n^{1/2})$ lower bound for the same.

Attempt 2: We use the same base graph as above, but in place of the above blobs, we replace each node of the hsot graph using 3-colorable graphs. Where the constructions for the YES and FAR instance differ, are the matchings we use to wrap up the construction.

- In the YES case, add a matching between the adjacent blobs that preserves 3-colorability.
- In the FAR case, add a random matching between adjacent blobs, this will create many contradicitons to colorablility.

(**Note:** By 'adjacent blobs', we mean the blobs added to both ends of an edge of the configuration host graph.)
This approach does not lead to a false proof for a lower bound for bipartiteness as all our contradicitons to 3-colorability are hidden within the matchings, so merely changing the blobs to the suitable ones outlined above for bipartiteness will not necessarily make the graph in the YES instance bipartite. However, we have not yet managed to complete the proof of the lower bound for an algorithm on this instance.

References

[Bol88] Béla Bollobás. The isoperimetric number of random regular graphs. *European Journal of combinatorics*, 9(3):241–244, 1988.

[BOT02] Andrej Bogdanov, Kenji Obata, and Luca Trevisan. A lower bound for testing 3-colorability in bounded-degree graphs. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 93–102. IEEE, 2002.

[GR98] Oded Goldreich and Dana Ron. A sublinear bipartiteness tester for bounded degree graphs. In *Proceedings of the thirtieth annual ACM Symposium on Theory of Computing*, pages 289–298, 1998.

[PY23] Pan Peng and Yuichi Yoshida. Sublinear-time algorithms for max cut, max e2lin (q), and unique label cover on expanders. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 4936–4965. SIAM, 2023.