# Algorithmic Aspects of Property Testing in the Dense Graphs Model

Oded Goldreich*
Department of Computer Science
Weizmann Institute of Science
Rehovot, Israel.
oded.goldreich@weizmann.ac.il

Dana Ron†
Department of EE–Systems
Tel-Aviv University
Ramat-Aviv, Israel.
danar@eng.tau.ac.il

October 21, 2010

**Abstract**

In this paper we consider two basic questions regarding the query complexity of testing graph properties in the adjacency matrix model. The first question refers to the relation between adaptive and non-adaptive testers, whereas the second question refers to testability within complexity that is inversely proportional to the proximity parameter, denoted $\epsilon$. The study of these questions reveals the importance of algorithmic design in this model. The highlights of our study are:

- A gap between the complexity of adaptive and non-adaptive testers. Specifically, there exists a natural graph property that can be tested using $\widetilde{O}(\epsilon^{-1})$ adaptive queries, but cannot be tested using $o(\epsilon^{-3/2})$ non-adaptive queries.
- In contrast, there exist natural graph properties that can be tested using $\widetilde{O}(\epsilon^{-1})$ non-adaptive queries, whereas $\Omega(\epsilon^{-1})$ queries are required even in the adaptive case.

We mention that the properties used in the foregoing conflicting results have a similar flavor, although they are of course different.

**Keywords:** Property Testing, Adaptivity vs. Non-adaptivity, Graph Properties,

*Partially supported by the Israel Science Foundation (grants No. 460/05 and 1041/08).
†Partially supported by the Israel Science Foundation (grants No. 89/05 and 246/08).

# Contents

# 1 Introduction

In the last couple of decades, the area of property testing has attracted much attention (see, e.g., a couple of recent surveys [R1, R2]). Loosely speaking, property testing typically refers to sub-linear time probabilistic algorithms for deciding whether a given object has a predetermined property or is far from any object having this property. Such algorithms, called testers, obtain bits of the object by performing queries, which means that the object is seen as a function and the testers get oracle access to this function. Thus, a tester may be expected to work in time that is sub-linear in the length of the description of this object.

Much of the aforementioned work (see, e.g., [GGR, AFKS, AFNS]) was devoted to the study of testing graph properties in the adjacency matrix model, which is also the setting of the current work. In this model, introduced in [GGR], graphs are viewed as symmetric Boolean functions over a domain consisting of all possible vertex-pairs. Namely, an $N$-vertex graph $G = ([N], E)$ is represented by the function $g : [N] \times [N] \to \{0, 1\}$ such that $\{u, v\} \in E$ if and only if $g(u, v) = 1$. Consequently, an $N$-vertex graph represented by the function $g : [N] \times [N] \to \{0, 1\}$ is said to be $\epsilon$-far from some predetermined graph property if more than $\epsilon \cdot N^2$ entries of $g$ must be modified in order to yield a representation of a graph that has this property. We refer to $\epsilon$ as the proximity parameter. Given this representation, the algorithm may query whether there is an edge between any pair of vertices of its choice, and the query and time complexity of testing are stated in terms of $\epsilon$ and possibly the number, $N$, of vertices in the graph. We note that this model is most suitable for testing *dense* graphs, that is, graphs in which the number of edges is $\Omega(N^2)$. This is true both because the adjacency matrix is an appropriate representation of dense graphs, and because distance between graphs is related to their size, $\Theta(N^2)$. Discussion of other models, more suitable for sparse graphs, can be found, in [GR02, PR, KKR].

Interestingly, many natural graph properties can be tested in the adjacency matrix model with query complexity that depends only on the proximity parameter; see [GGR], which presents testers with query complexity poly$(1/\epsilon)$, and [AFNS], which characterizes the class of properties that are testable within query complexity that depends only on the proximity parameter (where this dependence may be an arbitrary function of $\epsilon$). However, a common phenomenon in all the aforementioned works is that they utilize quite naive algorithms and their focus is on the analysis of these algorithms, which is often quite sophisticated. This phenomenon is no coincidence: As shown in [AFKS, GT], when ignoring a quadratic blow-up in the query complexity, property testing in this model reduces to sheer combinatorics. Specifically, without loss of generality, the tester may just inspect a random induced subgraph (of an appropriate size) of the input graph.

In this paper we demonstrate that a more refined study of property testing in this model reveals the importance of algorithmic design also in this model. This is demonstrated both by studying the advantage of adaptive testers over non-adaptive ones as well as by studying the class of properties that can be tested within complexity that is inversely proportional to the proximity parameter.

## 1.1 Two Related Studies

We start by reviewing the two related studies conducted in the current work.

### 1.1.1 Adaptivity vs. Non-adaptivity

A tester is called non-adaptive if it determines all its queries independently of the answers obtained for previous queries, and otherwise it is called adaptive. Indeed, by [AFKS, GT], the benefit of adaptivity (or, equivalently, the cost of non-adaptivity) is polynomially bounded: Specifically, any (possibly adaptive) tester, for any graph property, of query complexity $q(N,\epsilon)$ can be transformed into a non-adaptive tester of query complexity $O(q(N,\epsilon)^2)$. But is this quadratic gap an artifact of the known proofs (of [AFKS, GT]) or does it reflect something inherent?

A recent work by [GR07] suggests that the latter case may hold: For every $\epsilon > 0$, they showed that the set of $N$-vertex bipartite graphs of maximum degree $O(\epsilon N)$ is $\epsilon$-testable (i.e., testable with respect to proximity parameter $\epsilon$) by $\widetilde{O}(\epsilon^{-3/2})$ queries, while by [BT] a non-adaptive tester for this set must use $\Omega(\epsilon^{-2})$ queries. Thus, there exists a case where non-adaptivity has the cost of increasing the query complexity; specifically, for any $c < 4/3$, the query complexity of the non-adaptive tester is greater than a $c$-power of the query complexity of the adaptive tester (i.e., $\widetilde{O}(\epsilon^{-3/2})^c = o(\epsilon^{-2})$). We stress that the result of [GR07] does not refer to property testing in the "proper" sense; that is, the complexity is not analyzed with respect to a varying value of the proximity parameter for a fixed property. It is rather the case that, for every value of the proximity parameter, a different property, which depends on this parameter, is considered. The upper bounds and lower bounds refer to this combination of a property tailored for a fixed value of the proximity parameter. Thus, *the work of [GR07] leaves open the question of whether there exists a single graph property such that adaptivity is beneficial for any value of the proximity parameter* (as long as $\epsilon > N^{-\Omega(1)}$). That is, the question is whether adaptivity is beneficial for the standard asymptotic-complexity formulation of property testing.

#### 1.1.2 Complexity linearly related to the proximity parameter

As shown in [GGR], many natural graph properties can be tested within query complexity that is polynomial in the reciprocal of the proximity parameter and independent of the size of the graph. We ask whether a linear complexity is possible at all, and if so which properties can be tested with query complexity that is linear (or almost linear) in the reciprocal of the proximity parameter, that is, with query complexity $\tilde{O}(1/\epsilon)$.[1]

The first question is easy to answer even when avoiding *trivial* properties. We say that a graph property $\Pi$ is trivial for testing if for every $\epsilon > 0$ there exists $N_0 > 0$ such that for every $N \geq N_0$ either all $N$-vertex graphs belong to $\Pi$ or all of them are $\epsilon$-far from $\Pi$. Note that the property of being a clique (equiv., an independent set) can be tested by $O(1/\epsilon)$ queries, even when these queries are non-adaptive (e.g., make $O(1/\epsilon)$ random queries and accept if and only if all return 1). Still, we ask whether "more interesting" graph theoretical properties can also be tested within similar complexity, either only adaptively or also non-adaptively. In particular, the property of being a clique (or an independent set) is viewed as "non-interesting" since it contains a single $N$-vertex graph (per each $N$) and is represented by a constant function.

### 1.2 Our Results

We address the foregoing questions by studying a sequence of natural graph properties, which are defined formally in Section 2.2. The first property in the sequence, called clique collection and denoted $\mathcal{CC}$, is the set of graphs such that each graph consists of a collection of isolated cliques. Testing this property corresponds to the following natural clustering problem: can a set of possibly related elements be partitioned into "perfect clusters" (i.e., two elements are in the same cluster if and only if they are related)? For this property, $\mathcal{CC}$, we prove a gap between adaptive and non-adaptive query complexity, where the adaptive query complexity is almost linear in the reciprocal of the proximity parameter. That is:

[1]Note that $\Omega(1/\epsilon)$ queries are required for testing any of the graph properties considered in the current work; for a more general statement see the beginning of Section 6.

**Theorem 1.1** (the query complexity of clique collection):

1. *There exists an adaptive tester for $\mathcal{CC}$ whose query complexity is $\widetilde{O}(\epsilon^{-1})$. Furthermore, this tester has one-sided error and runs in time $\widetilde{O}(\epsilon^{-1})$.*[2]

2. *Any non-adaptive tester for $\mathcal{CC}$ must have query complexity $\Omega(\epsilon^{-4/3})$.*

3. *There exists a non-adaptive tester for $\mathcal{CC}$ whose query complexity is $O(\epsilon^{-4/3})$. Furthermore, this tester has one-sided error and runs in time $O(\epsilon^{-4/3})$.*

Note that the complexity gap between Parts 1 and 2 of Theorem 1.1 matches the gap established by [GR07] for "non-proper" testing. A larger gap is established for a property of graphs, called bi-clique collection and denoted $\mathcal{BCC}$, where a graph is in $\mathcal{BCC}$ if it consists of a collection of isolated bi-cliques (i.e., complete bipartite graphs). We note that bi-cliques may be viewed as the bipartite analogues of cliques (w.r.t. general graphs), and indeed they arise naturally in clustering applications that are modeled by bipartite graphs over two types of elements.

**Theorem 1.2** (the query complexity of bi-clique collection):

1. *There exists an adaptive tester for $\mathcal{BCC}$ whose query complexity is $\widetilde{O}(\epsilon^{-1})$. Furthermore, this tester has one-sided error and runs in time $\widetilde{O}(\epsilon^{-1})$.*

2. *Any non-adaptive tester for $\mathcal{BCC}$ must have query complexity $\Omega(\epsilon^{-3/2})$. Furthermore, this holds even if the input graph is promised to be bipartite.*

The furthermore clause in Part 2 of Theorem 1.2 holds also for the model studied in [AFN], where the bi-partition of the graph is given.

Theorem 1.2 asserts that the gap between the query complexity of adaptive and non-adaptive testers may be a power of $1.5 - o(1)$. Recall that the results of [AFKS, GT] assert that the gap may not be larger than quadratic. We conjecture that this upper bound can be matched.

**Conjecture 1.3** (an almost-quadratic complexity gap): *For every positive integer $t \geq 5$, there exists a graph property $\Pi$ for which the following holds:*

1. *There exists an adaptive tester for $\Pi$ whose query complexity is $\widetilde{O}(\epsilon^{-1})$.*

2. *Any non-adaptive tester for $\Pi$ must have query complexity $\Omega(\epsilon^{-2+(2/t)})$.*

3. *There exists an efficient non-adaptive tester for $\Pi$ whose query complexity is $\widetilde{O}(\epsilon^{-2+2t^{-1}})$.*

*Furthermore, $\Pi$ consists of graphs that are each a collection of "super-cycles" of length $t$, where a super-cycle is a set of $t$ independent sets arranged on a cycle such that each pair of adjacent independent sets is connected by a complete bipartite graph.*

We were able to prove Part 2 of Conjecture 1.3, but failed to provide a full analysis of an algorithm that we designed for Part 1. However, we were able to prove a *promise problem version of Conjecture 1.3*; specifically, this promise problem (stated in Theorem 5.7) refers to inputs promised to reside in a set $\Pi' \supset \Pi$ and the tester is required to distinguish graphs in $\Pi$ from graphs that are $\epsilon$-far from $\Pi$.

[2] We refer to a model in which elementary operations regarding pairs of vertices are charged at unit cost.

In contrast to the foregoing results that aim at identifying properties with a substantial gap between the query complexity of adaptive versus non-adaptive testing, we also study cases in which no such gap exists. Since query complexity that is linear in the reciprocal of the proximity parameter is minimal for many natural properties, and, in fact, for any property that is "non-trivial for testing" (as defined at the end of Subsection 1.1), we focus on non-adaptive testers that approximately meet this bound. Among the results obtained in this direction, we highlight the following one.

**Theorem 1.4** (the query complexity of collections of $O(1)$ cliques): *For every positive integer c, there exists a non-adaptive tester of query complexity $\widetilde{O}(\epsilon^{-1})$ for the set of graphs such that each graph consists of a collection of up to c cliques. Furthermore, this tester has one-sided error and runs in time $\widetilde{O}(\epsilon^{-1})$.*

Theorem 1.4 should be viewed as a first step in the study of graph properties that are the simplest to test; that is, the class of graph properties that have a non-adaptive of query complexity $\widetilde{O}(\epsilon^{-1})$. We mention that a second step, which significantly generalizes Theorem 1.4, has been subsequently taken in [A09, AG].

**Discussion.** Our results demonstrate that a finer look at property testing of graphs in the adjacency matrix model reveals the role of algorithm design in this model. In particular, in some cases (see, e.g., Theorems 1.1 and 1.2), *carefully designed adaptive algorithms outperform any non-adaptive algorithm.* Indeed, this conclusion stands in contrast to [GT, Thm. 2], which suggests that a less fine view, which ignores polynomial blow-ups,[3] deems algorithm design irrelevant to this model. We also note that, in some cases (see, e.g., Theorem 1.4 and Part 3 of Theorem 1.1), *carefully designed non-adaptive algorithms outperform canonical ones.*

As discussed previously, one of the goals of this work was to study the relation between adaptive and non-adaptive testers in the adjacency matrix model. Our results demonstrate that, in this model, the relation between the adaptive and non-adaptive query-complexities is not fixed, but rather varies with the computational problem at hand. In some cases (e.g., Theorem 1.4) the complexities are essentially equal, indeed, as in the case of sampling [CEG]. In other cases (e.g., Theorem 1.1), these complexities are related by a fixed power (e.g., $4/3$) that is strictly between 1 and 2. And, yet, in other cases (e.g., Theorem 5.7) the non-adaptive complexity is quadratic in the adaptive complexity, which is the maximum gap possible (by [AFKS, GT]). Furthermore, by Theorem 5.7, for any $t \geq 4$, there exists a promise problem for which the aforementioned complexities are related by a power of $2 - (2/t)$.

Needless to say, the fundamental relation between adaptive and non-adaptive algorithms was studied in a variety of models, and the current work studies it in a specific natural model (i.e., of property testing in the adjacency matrix representation). In particular, this relation has been studied in the context of property testing in other domains. Specifically, in the setting of testing the satisfiability of linear constraints, it was shown that adaptivity offers absolutely no gain [BHR]. A similar result holds for testing monotonicity of sequences of positive integers [F04]. In contrast, an exponential gap between the adaptive and non-adaptive complexities may exist in the context of testing other properties of functions [F04]. Lastly, we mention that an even more dramatic gap exists in the setting of testing graph properties in the bounded-degree model (of [GR02]); see [RS06].

[3] Recall that [GT, Thm. 2] asserts that canonical testers, which merely select a random subset of vertices and rule according to the induced subgraph, have query-complexity that is at most quadratic in the query-complexity of the best tester. We note that [GT, Thm. 2] also ignores the time-complexity of the testers.

## 1.3 A Complexity Theoretic Perspective

Let us start by rephrasing Conjecture 1.3, while recalling that it refers to properties for which testing requires (adaptive) query complexity that is at least linear in the reciprocal of the proximity parameter (see Proposition 6.1).

**Conjecture 1.3 (rephrased).** *For every integer $t \geq 2$, there exists a* (natural) *graph property $\Pi_t$ such that* non-adaptively *testing $\Pi_t$ has query complexity $\widetilde{\Theta}(q^{2-(2/t)})$, where $q = q(N, \epsilon)$ denotes the query complexity of* (adaptively) *testing $\Pi_t$.*

Recall that it is known that the non-adaptive query complexity of testing any graph property is at most quadratic in the adaptive query complexity. We stress that Conjecture 1.3 not only asserts that this upper bound is essentially tight, but rather asserts an infinite hierarchy of possible functional relations between the non-adaptive and adaptive query complexity.

The results in this work refer to "two and a half" elements in the conjectured hierarchy as well as to a corresponding hierarchy of promise problems. Specifically, denoting the (adaptive) query complexity by $q = q(N, \epsilon)$, we have:

- Theorem 1.4 establishes the conjecture for $t = 2$. Specifically, Theorem 1.4 presents natural graph properties that have non-adaptive query complexity $\widetilde{\Theta}(q)$.
- Theorem 1.1 establishes the conjecture for $t = 3$. Specifically, Theorem 1.1 presents a natural graph property that has non-adaptive query complexity $\widetilde{\Theta}(q^{4/3})$.
- Theorem 1.2 establishes half of the conjecture for $t = 4$. Specifically, Theorem 1.2 presents a natural graph property that has non-adaptive query complexity $\widetilde{\Omega}(q^{3/2})$.
- Theorem 5.7 fully establishes the conjecture in the setting of promise problems. We stress that these promise problems are fixed (independently of the proximity parameter).

Indeed, in all our results $q = q(N, e) = \widetilde{\Omega}(1/\epsilon)$. We also mention that in all our results the upper bounds are established by one-sided error testers, whereas the lower bounds hold also for general (i.e., two-sided error) testers.

**Open problems.** In addition to the resolution of Conjecture 1.3, our study raises many other open problems; the most evident ones are listed next.

1. What is the non-adaptive query complexity of $\mathcal{BCC}$? Note that Theorem 1.2 only establishes a lower bound of $\Omega(\epsilon^{-3/2})$. We conjecture that an efficient non-adaptive algorithm of query complexity $\widetilde{O}(\epsilon^{-3/2})$ can be devised.
2. For which constants $c \in [1, 2]$ does there exist a property that has adaptive query complexity of $q(\epsilon)$ and non-adaptive query complexity of $\widetilde{\Theta}(q(\epsilon)^c)$? Note that Theorem 1.1 shows that $4/3$ is such a constant, and the same holds for the constant 1 (see, e.g., Theorem 1.4). We conjecture (see Conjecture 1.3) that, for any $t \geq 2$, it holds that the constant $2 - (2/t)$ also satisfies the foregoing requirement. It may be the case that these constants are the only ones that satisfy this requirement.
3. Characterize the class of graph properties for which the query complexity of non-adaptive testers is almost linear in the query complexity of adaptive testers. Note that this class may not contain the property of bipartiteness [GR07].

4. Characterize the class of graph properties for which the query complexity of non-adaptive testers is almost quadratic in the query complexity of adaptive testers.

5. Characterize the class of graph properties for which the query complexity of adaptive (resp., non-adaptive) testers is almost linear in the reciprocal of the proximity parameter.

The last characterization project may be the most feasible among the three foregoing characterization projects. We mention that this is partially addressed in [A09, AG], which significantly extends and builds upon Theorem 1.4. Finally, we recall the well-known open problem, partially addressed in [AS], of providing a characterization of the class of graph properties that are testable within query complexity that is polynomial in the reciprocal of the proximity parameter.

### 1.4 Organization

Section 2 contains a review of the basic notions underlying this work as well as formal definitions of the graph properties that we study. In Section 3 we present an adaptive tester for Clique Collection that has almost-linear query complexity. This result stands in contrast to the tight lower bound on the query complexity of non-adaptive testers for Clique Collection, presented in Section 4. Theorem 1.1 follows by combining the results in these sections. Larger gaps between the query complexity of adaptive versus non-adaptive testers (i.e., Theorems 1.2 and 5.7) are presented in Section 5. On the other hand, in Section 6, we present non-adaptive testers of query complexity that is almost linear in the reciprocal of the proximity parameter. We conclude this paper, in Section 7, by explicitly presenting three perspectives on our results.

## 2 Preliminaries

In this section we review the definition of property testing, when specialized to graph properties in the adjacency matrix model. We also define several natural graph properties, which will serve as the pivot of our study.

### 2.1 Basic Notions

For an integer $n$, we let $[n] = \{1, \ldots, n\}$. A generic $N$-vertex graph is denoted by $G = ([N], E)$, where $E \subseteq \{\{u, v\} : u, v \in [N]\}$ is a set of unordered pairs of vertices. Any set of such graphs that is closed under isomorphism is called a graph property. By oracle access to such a graph $G = ([N], E)$ we mean oracle access to the Boolean function that answers the query $\{u, v\}$ (or rather $(u, v) \in [N] \times [N]$) with the bit 1 if and only if $\{u, v\} \in E$.

**Definition 2.1** (property testing for graphs in the adjacency matrix model): *A* tester *for a graph property $\Pi$ is a probabilistic oracle machine that, on input parameters $N$ and $\epsilon$ and access to an $N$-vertex graph $G = ([N], E)$, outputs a binary verdict that satisfies the following two conditions.*

1. *If $G \in \Pi$ then the tester accepts with probability at least $2/3$.*

2. *If $G$ is $\epsilon$-far from $\Pi$ then the tester accepts with probability at most $1/3$, where $G$ is* $\epsilon$-far *from $\Pi$ if for every $N$-vertex graph $G' = ([N], E') \in \Pi$ it holds that the symmetric difference between $E$ and $E'$ has cardinality that is greater than $\epsilon N^2$.*[4]

[4]Indeed, it is more natural to require that this symmetric difference should have cardinality that is greater than $\epsilon \cdot \binom{N}{2}$. The current convention is adopted for the sake of convenience.

*If the tester accepts every graph in $\Pi$ with probability 1, then we say that it has* one-sided error. *A tester is called* non-adaptive *if it determines all its queries based solely on its internal coin tosses* (and the parameters $N$ and $\epsilon$)*; otherwise it is called* adaptive.

The query complexity of a tester is the number of queries it makes to any $N$-vertex graph oracle, as a function of the parameters $N$ and $\epsilon$. We say that a tester is efficient if it runs in time that is polynomial in its query complexity, where basic operations on elements of $[N]$ (and in particular, uniformly selecting an element in $[N]$) are counted at unit cost. We note that all testers presented in this paper are efficient, whereas the lower bounds hold also for non-efficient testers.

We shall focus on properties that can be tested with query complexity that only depends on the proximity parameter, $\epsilon$. Thus, the query complexity upper bounds that we state hold for any values of $\epsilon$ and $N$, but will be meaningful only for $\epsilon > 1/N^2$ or so. In contrast, the lower bounds (e.g., of $\Omega(1/\epsilon)$) cannot possibly hold for $\epsilon < 1/N^2$, but they will indeed hold for any $\epsilon > N^{-\Omega(1)}$. Alternatively, one may consider the query-complexity as a function of $\epsilon$, where for each fixed value of $\epsilon > 0$ the value of $N$ tends to infinity.

**Notation and a convention.** For a fixed graph $G = ([N], E)$, we denote by $\Gamma(v) = \{u : \{u, v\} \in E\}$ the set of neighbors of vertex $v$. At times, we look at $E$ as a subset of $[N] \times [N]$; that is, we often identify $E$ with $\{(u, v) : \{u, v\} \in E\}$. For two sets $V_1, V_2 \subseteq [N]$, we denote by $E(V_1, V_2)$ the set of pairs $(u, v) \in E \cap (V_1 \times V_2)$.

If a graph $G = ([N], E)$ is not $\epsilon$-far from a property $\Pi$ then we say that $G$ is $\epsilon$-close to $\Pi$; this means that at most $\epsilon N^2$ edges should be added and/or removed from $G$ such to yield a graph in $\Pi$.

## 2.2 The Graph Properties to be Studied

The set of graphs that consists of a collection of isolated cliques is called clique collection and is denoted $\mathcal{CC}$; that is, a graph $G = ([N], E)$ is in $\mathcal{CC}$ if and only if the vertex set $[N]$ can be partitioned into $(C_1, \ldots, C_t)$ such that the subgraph induced by each $C_i$ is a clique and there are no edges with endpoints in different $C_i$'s (i.e., for every $u < v \in [N]$ it holds that $\{u, v\} \in E$ if and only if there exists an $i$ such that $u, v \in C_i$). In other words, the relation defined by the graph edges is *transitive*. If $t \leq c$ then we say that $G$ is in $\mathcal{CC}^{\leq c}$; that is, $\mathcal{CC}^{\leq c}$ is the subset of $\mathcal{CC}$ that contains graphs that are each a collection of up-to $c$ isolated cliques.

A bi-clique is a complete bipartite graph (i.e., a graph $G = (V, E)$ such that $V$ is partitioned into $(S, V \setminus S)$ such that $\{u, v\} \in E$ if and only if $u \in S$ and $v \in V \setminus S$). Note that a graph is a bi-clique if and only if its complement is in $\mathcal{CC}^{\leq 2}$. The set of graphs that consists of a collection of isolated bi-cliques is called bi-clique collection and denoted $\mathcal{BCC}$; that is, a graph $G = ([N], E)$ is in $\mathcal{BCC}$ if and only if the vertex set $[N]$ can be partitioned into $(V_1, \ldots, V_t)$ such that the subgraph induced by each $V_i$ is a bi-clique and there are no edges with endpoints in different $V_i$'s (i.e., each $V_i$ is partitioned into $(S_i, V_i \setminus S_i)$ such that for every $u < v \in [N]$ it holds that $\{u, v\} \in E$ if and only if there exists an $i$ such that $(u, v) \in S_i \times (V_i \setminus S_i)$).

Generalizations of $\mathcal{BCC}$ are obtained by considering collections of "super-paths" and "super-cycles" respectively. A super-path (of length $t$) is a sequence of disjoint sets of vertices, $S_1, \ldots, S_t$, such that vertices $u, v \in \bigcup_{i \in [t]} S_i$ are connected by an edge if and only if for some $i \in [t-1]$ it holds that $u \in S_i$ and $v \in S_{i+1}$. Note that a bi-clique can be viewed as a super-path of length two. We denote the set of graphs that consists of a collection of isolated super-paths of length $t$ by $\mathcal{S}\mathsf{P}_t\mathcal{C}$ (e.g., $\mathcal{S}\mathsf{P}_2\mathcal{C} = \mathcal{BCC}$). Similarly, a super-cycle (of length $t$) is a sequence of disjoint sets of vertices, $S_1, \ldots, S_t$, such that vertices $u, v \in \bigcup_{i\in[t]} S_i$ are connected by an edge if and only if for some $i \in [t]$ it holds that $u \in S_i$ and $v \in S_{(i \bmod t)+1}$. Note that a bi-clique that has at least two vertices on each side can be viewed as a super-cycle of length four (by partitioning each of its sides into two parts). We denote the set of graphs that consists of a collection of isolated super-cycles of length $t$ by $\mathcal{SC}_t\mathcal{C}$ (e.g., $\mathcal{SC}_4\mathcal{C} \subset \mathcal{BCC}$, where the strict containment is due to the pathological case of bi-cliques having at most one node on one side).

## 2.3 On Proving Lower Bound for Property Testing

All our lower bounds employ the following method, which is commonly attributed to Yao [Y77]. To prove that a certain class, $\mathcal{C}$, of algorithms cannot decide a certain (promise) problem, we present two distributions, one concentrated on YES-instances and the other concentrated on NO-instances and prove that any algorithm in $\mathcal{C}$ cannot distinguish these two distributions. In the context of property testing, the first distribution, $D_1$, is over objects that have the predetermined property $\Pi$, whereas the second distribution, $D_2$, is over objects that are $\epsilon$-far from $\Pi$, where $\epsilon$ is the value of the proximity parameter for which we seek to prove the hardness of testing. Now, if $T$ is a tester for $\Pi$, then on input proximity parameter $\epsilon$, it should hold that:

1. With probability at least 2/3 (taken over both $D_1$ and $T$'s internal coin tosses), when given access to an object selected according to $D_1$ the tester $T$ accepts.

2. With probability at most 1/3 (taken over both $D_2$ and $T$'s internal coin tosses), when given access to an object selected according to $D_2$ the tester $T$ accepts.

Let us define the distinguishing gap of $M$ between $D_1$ and $D_2$ as $|p_1 - p_2|$ where $p_i$ denotes the probability that $M$ outputs 1 ("accept") when given access to an object drawn according to $D_i$. Thus, $T$ must be able to distinguish, with a gap of at least 1/3 between objects distributed according to $D_1$ and objects distributed according to $D_2$. Therefore, in order to prove a query complexity lower bound $q$, we show that oracle machines $M$ making fewer than $q$ queries cannot distinguish such distributions with gap at least 1/3. In other words, it suffices to establish an upper bound on the distinguishing gap of any oracle machine that makes a number of queries that is smaller than the claimed lower bound. Using an averaging argument (and relying on the lack of a uniformity condition), it suffices to establish this upper bound for deterministic machines.

Finally, when considering non-adaptive testers, it suffices to consider a fixed sequence of queries, and the distribution of answers provided by objects selected according to the two distributions. Thus, for non-adaptive oracle machines, the distinguishing gap is upper bounded by the statistical difference between these two distributions of answers. Recall that the statistical difference between two distributions $A$ and $B$ is

$$\max_S \{\Pr_{e\sim A}[e \in S] - \Pr_{e\sim B}[e \in S]\} \;=\; \frac{1}{2} \cdot \sum_v |\Pr_{e\sim A}[e = v] - \Pr_{e\sim B}[e = v]| \tag{1}$$

where $\Pr_{e\sim D}[P(e)]$ denotes the probability that an element drawn according to distribution $D$ satisfies the predicate $P$.

## 2.4 Annoying Technicalities

We allowed ourselves various immaterial inaccuracies. For example, various quantities (e.g., $\log_2(1/\epsilon)$) are treated as if they are integers, whereas one should actually use some rounding and compensate for the rounding error. At times, we ignore events that occur with probability that is inversely proportional to the number of vertices; for example, when we select a random sample of $s = O(1)$ (or $s = \widetilde{O}(1/\epsilon)$) vertices, we often analyze it as if sampling was done with repetitions. In some places, we do not specify the "high" (constant) probability with which some events occur; but such missing details are easy to fill-up. In other places, we specify high constants that are not the best ones possible.

# 3 The Adaptive Query Complexity of Clique Collection

In this section we study the (adaptive) query complexity of clique collection, presenting an almost optimal (adaptive) tester for this property. Loosely speaking, the tester starts by finding a few random neighbors of a few randomly selected start vertices, and then examines the existence of edges among the neighbors of each start vertex as well as among these neighbors and the non-neighbors of each start vertex. Note that if for some vertex $v$ the algorithm either finds two neighbors of $v$ that do not have an edge between them, or the algorithm finds a neighbor of $v$ and a non-neighbor of $v$ that have an edge between them, then it has evidence that the graph is not a clique collection.

We highlight the fact that adaptivity is used in order to perform queries that refer only to pairs of neighbors of the same start vertex. To demonstrate the importance of this fact, consider the case that the $N$-vertex graph is partitioned into $O(1/\epsilon)$ connected components each having $O(\epsilon N)$ vertices. Suppose that we wish to tell whether the connected component that contains the vertex $v$ is indeed a clique, or there is a constant fraction of missing edges between the neighbors of $v$. Using adaptive queries we may first find a constant number of neighbors of $v$, by selecting $t \stackrel{\text{def}}{=} O(1/\epsilon)$ random vertices and checking whether each selected vertex is adjacent to $v$. We can then check whether these *constant number* of neighbors are adjacent to each other. In contrast, intuitively, a non-adaptive procedure cannot avoid making all $\binom{t}{2}$ possible queries, since it "does not know" which of the $t$ vertices are neighbors of $v$.

The foregoing adaptive procedure is tailored to the case that the $N$-vertex graph is partitioned into $O(1/\epsilon)$ ("strongly connected") components, each having $O(\epsilon N)$ vertices. In such a case, it suffices to check that a constant fraction of these components are in fact cliques (or rather close to being so), as described in the foregoing adaptive procedure, and that there are no edges (or rather relatively few edges) between the cliques. However, if the components (and potential cliques) are larger, then we should check more of them. Fortunately, due to their larger size, finding neighbors requires less queries, and the total number of queries remains invariant.

Thus, the algorithm, described next, works in iterations, where the iterations differ in the number of start vertices selected and in the size of the sample used to get uniformly selected neighbors (and non-neighbors) of these start vertices. Each iteration is designed to detect the existence of vertices with a certain, iteration dependent, lower bound on their degree, for which the following holds. Either there are relatively many missing edges between their neighbors, or there are relatively many "superfluous" edges between their neighbors and their non-neighbors. The quantification of "relatively many" is also dependent on the iteration. As we show in our correctness proof, if the graph is $\epsilon$-far from $\mathcal{CC}$, then there must be relatively many such vertices for at least one of the iterations (where again, the quantification of "relatively many" depends on the iteration and is related to the number of start vertices that are selected in the iteration).

**Algorithm 3.1** (adaptive tester for $\mathcal{CC}$): *On input $N$ and $\epsilon$ and oracle access to a graph $G = ([N], E)$, set $t = \Theta(\log^3(1/\epsilon))$, and proceed in $\ell \stackrel{\text{def}}{=} \log_2(1/\epsilon) + 2$ iterations as follows: For $i =$*

$1, \ldots, \ell$, *select uniformly* $10 \cdot 2^i$ *start vertices and for each selected vertex* $v \in [N]$ *perform the following sub-test, denoted* sub-test$_i(v)$*:*

1. *Select at random a sample,* $S$*, of* $t/(2^i \epsilon)$ *vertices.*

2. *Determine* $\Gamma_S(v) = S \cap \Gamma(v)$*, by making the queries* $(v, w)$ *for each* $w \in S$*.*

3. *If* $|\Gamma_S(v)| \leq \sqrt{t/2^i \epsilon}$ *then check that for every* $u, w \in \Gamma_S(v)$ *it holds that* $(u, w) \in E$*. Otherwise* (i.e., $|\Gamma_S(v)| > \sqrt{t/(2^i \epsilon)}$), *select a sample of* $t/(2^i \epsilon)$ *pairs in* $\Gamma_S(v) \times \Gamma_S(v)$ *and check that each selected pair is in* $E$*.*

4. *Select a sample of* $t/(2^i \epsilon)$ *pairs in* $\Gamma_S(v) \times (S \setminus \Gamma_S(v))$ *and check that each selected pair is* not *in* $E$*.*

*The sub-test* (i.e., sub-test$_i(v)$) *accepts if and only if all checks were positive* (i.e., no edges were missed in Step 3 and no edges were detected in Step 4)*. The tester itself accepts if and only if all* $\sum_{i=1}^{\ell} 10 \cdot 2^i$ *invocations of the sub-test accepted.*

The query complexity of this algorithm is $\sum_{i=1}^{\ell} 10 \cdot 2^i \cdot O(t/(2^i \epsilon)) = O(\ell \cdot t/\epsilon) = \widetilde{O}(1/\epsilon)$, and the running time is of the same order as the query complexity. Clearly, this algorithm accepts with probability 1 any graph that is in $\mathcal{CC}$. It remains to analyze its behavior on graphs that are $\epsilon$-far from $\mathcal{CC}$, and thus establish Part 1 of Theorem 1.1, which states that there exists an adaptive (one-sided error) tester for $\mathcal{CC}$ whose complexity is $\widetilde{O}(\epsilon^{-1})$.

**Lemma 3.2** *If* $G = ([N], E)$ *is* $\epsilon$*-far from* $\mathcal{CC}$*, then on input* $N, \epsilon$ *and oracle access to* $G$*, Algorithm 3.1 rejects with probability at least* $2/3$*.*

**Proof:** We shall prove the contrapositive statement; that is, that if Algorithm 3.1 accepts a graph $G$ with probability at least $1/3$, then $G$ is $\epsilon$-close to $\mathcal{CC}$. The proof makes use of the following notion of $i$-good start vertices (for $i \in [\ell]$). We first show that if Algorithm 3.1 accepts with probability at least $1/3$, then the number of vertices with a relatively high degree that are not $i$-good is relatively small, and next show how to use $i$-good vertices (with a relatively high degree) in order to construct a partition of the vertices that demonstrates that the graph is $\epsilon$-close to $\mathcal{CC}$.

The following central definition of $i$-good vertices refers to a parameter $\gamma$, which is set to $c/t$, where $t$ is as determined in Algorithm 3.1 and $c$ is a constant (which will be chosen to be sufficiently large for the purposes of the analysis). In fact, it is useful to think of first setting $\gamma$ to be $1/(c' \log^3(1/\epsilon))$ for some sufficiently large constant $c'$ (which ensures that we get a good partition based on $i$-good vertices), and then setting $t$ (which determines the sample sizes selected by the algorithm) to be $c/\gamma$. (In fact, at the heart of the analysis is a parameter $\beta$ which is set to be $1/(c''\ell)$ for a constant $c''$, and $\gamma$ is set to be $\beta/(c'''\ell^2)$ for a constant $c'''$.)

**Definition 3.2.1** *A vertex* $v$ *is* $i$-good *if the following two conditions hold.*

1. *The number of missing edges in the subgraph induced by* $\Gamma(v)$ *is at most* $\gamma \cdot 2^i \epsilon \cdot |\Gamma(v)| \cdot N$*.*

2. *For every positive integer* $j \leq j_0 \stackrel{\rm def}{=} \log_2(|\Gamma(v)|/(\gamma \cdot 2^i \epsilon N))$*, the number of vertices in* $\Gamma(v)$ *that have at least* $\gamma \cdot 2^{i+j} \epsilon \cdot N$ *neighbors that do not belong to* $\Gamma(v)$ *is at most* $2^{-j} \cdot |\Gamma(v)|$*.*

Note that Condition 1 holds vacuously whenever $|\Gamma(v)| < \gamma \cdot 2^i\epsilon \cdot N$ (since in such a case, $|\Gamma(v)|^2 < \gamma \cdot 2^i\epsilon \cdot |\Gamma(v)| \cdot N$). However, when $|\Gamma(v)|$ is sufficiently larger than $\gamma \cdot 2^i\epsilon \cdot N$, then Condition 1 implies that a large fraction of the vertices in $\Gamma(v)$ neighbor almost all vertices in $\Gamma(v)$, so that $\Gamma(v)$ is close to being a clique. Condition 2 implies that almost all vertices in $\Gamma(v)$ have relatively few neighbors outside of $\Gamma(v)$, where "almost all" and "relatively few" are quantified and related. On the other hand, as the next claim establishes, if a vertex $v$ is not $i$-good (and has a sufficiently high degree), then sub-test$_i(v)$ will detect it with high constant probability.

**Claim 3.2.2** *If $v$ has degree at least $\gamma \cdot 2^i\epsilon \cdot N$ and is not $i$-good, then the probability that* sub-test$_i(v)$ *accepts is less than* 0.05.

Proof: Intuitively, the lower bound on $|\Gamma(v)|$ implies that the violation of any of the two conditions of Definition 3.2.1 is detected with high probability by sub-test$_i(v)$. For example, if a 0.01 fraction of the vertices in $\Gamma(v)$ have less than $0.99 \cdot |\Gamma(v)|$ neighbors in $\Gamma(v)$, then the residual sample $\Gamma_S(v)$ (created by sub-test$_i(v)$) is likely to contain a constant fraction of vertices that miss a constant fraction of neighbors in $\Gamma_S(v)$. The actual proof, which refers to the two conditions of $i$-goodness, follows. In this proof, whenever we say: "with high constant probability", we mean with probability at least $1 - \delta$, where $\delta$ is a constant that is sufficiently small, so that when we sum all failure probabilities, we get at most 0.05.

Assume that Condition 1 of $i$-goodness does not hold for $v$, and let

$$\rho \stackrel{\text{def}}{=} \frac{\gamma \cdot 2^i\epsilon \cdot |\Gamma(v)| \cdot N}{|\Gamma(v)|^2} = \frac{\gamma \cdot 2^i\epsilon \cdot N}{|\Gamma(v)|} \tag{2}$$

denote the lower bound on the fraction of missing edges in $\Gamma(v)$. As noted in the discussion following Definition 3.2.1, Condition 1 of $i$-goodness may be violated only if $|\Gamma(v)| \geq \gamma \cdot 2^i\epsilon \cdot N$. Recall that sub-test$_i(v)$ selects a sample, $S$ of $t/(2^i\epsilon)$ vertices, and that $t = c/\gamma$ (for a constant $c$). By a multiplicative Chernoff bound, for a sufficiently large $c$, with high constant probability, it holds that $|\Gamma_S(v)| \geq m/2$, where

$$m \stackrel{\text{def}}{=} \frac{t}{\epsilon 2^i} \cdot \frac{|\Gamma(v)|}{N} \tag{3}$$

is the expected size of $\Gamma_S(v)$, and so $m \geq t \cdot \gamma = c$.

Assume from this point on that indeed $|\Gamma_S(v)| \geq n = m/2$, and note that the members of $\Gamma_S(v)$ are distributed uniformly in $\Gamma(v)$. Therefore, we may consider $n = m/2$ uniformly distributed vertices in $\Gamma(v)$, and define the following 0/1-valued random variables $\zeta_{j,k}$ for every $1 \leq j < k \leq n$. We let $\zeta_{j,k} = 1$ if there is no edge between the $j^{\text{th}}$ and $k^{\text{th}}$ vertices in the sample (of vertices in $\Gamma(v)$). Hence, $\text{Exp}[\zeta_{j,k}] \geq \rho$. We next give an upper bound (in terms of $c$) on $\text{Var}\left[\sum_{j<k} \zeta_{j,k}\right] / \text{Exp}^2\left[\sum_{j<k} \zeta_{j,k}\right]$, so that by applying Chebyshev's Inequality, it will follow that, with high constant probability, the fraction of edges that are missing in the subgraph induced by the said sample is at least $\rho/2$.

By the definition of $\zeta_{j,k}$, we have $\binom{n}{2}$ random variables, which are partially pairwise independent (i.e., $\zeta_{j,k}$ is independent of $\zeta_{j',k'}$ if $|\{j,k,j',k'\}| = 4$). Furthermore, these random variables assume values in $\{0,1\}$ (and so $\zeta_{j,k}^2 = \zeta_{j,k}$) and it holds (by the definitions of $n$ and $\rho$) that $n \cdot \rho = t\gamma/2 = c/2$. Assume, for simplicity that $\text{Exp}[\zeta_{j,k}]$ equals $\rho$ (and is not only lower bounded by $\rho$). It follows that $\text{Exp}\left[\sum_{j<k} \zeta_{j,k}\right] = \binom{n}{2} \cdot \rho > n^2\rho/3$ and $\text{Var}\left[\sum_{j<k} \zeta_{j,k}\right] < 4 \cdot \text{Exp}\left[\sum_{j<k,k'} \zeta_{j,k}\zeta_{j,k'}\right] \leq 4n \cdot \text{Exp}\left[\sum_{j<k} \zeta_{j,k}\right] < 2n^3\rho$. Thus, $\frac{\text{Var}\left[\sum_{j<k} \zeta_{j,k}\right]}{\text{Exp}^2\left[\sum_{j<k} \zeta_{j,k}\right]} < \frac{18}{n\rho} = \frac{36}{t\gamma} = 36/c$, which can be made an arbitrary small constant by choosing $c$ to be sufficiently large.

We thus obtain that if Condition 1 of $i$-goodness does not hold for $v$, then with high constant probability, the fraction of pairs of vertices in $\Gamma_S(v)$ that do not have an edge between them is at least $\rho/2$. Conditioned on this event, if $|\Gamma_S(v)| \le \sqrt{t/(2^i\epsilon)}$, so that Step 3 of sub-test$_i(v)$ checks whether $(u,w) \in E$ for every $u, w \in \Gamma_S(v)$, then we are done. Otherwise, the sub-test selects a random sample of $\frac{t}{2^i\epsilon} \ge \frac{t\gamma}{\rho} = \frac{c}{\rho}$ pairs of vertices in $\Gamma_S(v)$, and with probability at least $1-(1-\rho/2)^{c/\rho}$, which is close to 1 for a sufficiently large $c$, it will detect a missing edge.

Next assume that Condition 2 of $i$-goodness does not hold for $v$; that is, there exists a $j \le j_0$ such that more than $2^{-j} \cdot |\Gamma(v)|$ vertices in $\Gamma(v)$ have each a "high out-degree", that is, have each at least $\gamma \cdot 2^{i+j}\epsilon \cdot N$ neighbors that do not belong to $\Gamma(v)$. Using the same setting of $m$ and $n$ as in the previous paragraph (as well as the premise of the claim: $|\Gamma(v)| \ge \gamma \cdot 2^i\epsilon \cdot N$), we note (again) that $|\Gamma_S(v)| \ge n = m/2$ with high constant probability. Similarly, since there must be at least $\gamma \cdot 2^{i+j}\epsilon \cdot N$ vertices in $[N] \setminus \Gamma(v)$ (the neighbors of the high out-degree vertices that are not in $\Gamma(v)$), the number of vertices in $S \setminus \Gamma(v)$ is also at least half its expected value with high constant probability. Assume that these events in fact hold.

Since $v$ has at least $2^{-j} \cdot |\Gamma(v)|$ high out-degree neighbors, once again by a multiplicative Chernoff bound, with high constant probability we have that $\Gamma_S(v)$ contains at least $|\Gamma(v)| \cdot 2^{-j}/2 \ge t\gamma \cdot 2^{j_0-j}/4 \ge t\gamma/4 = c/4$ such vertices (where these vertices are uniformly distributed among the high out-degree neighbors of $v$). Consider a fixed choice of such a high out-degree vertex $u$ in $\Gamma_S(v)$. Since the vertices in $S \setminus \Gamma_S(v)$ are uniformly distributed in $[N] \setminus \Gamma(v)$, with high constant probability (by a multiplicative Chernoff bound), the number of neighbors that $u$ has in $S \setminus \Gamma_S(v)$ is at least half its expected value (i.e., at least $\gamma \cdot 2^{i+j}\epsilon \cdot |S|/2$). It follows by Markov's inequality that with high constant probability, the edge density in $\Gamma_S(v) \times (S \setminus \Gamma_S(v))$ is at least $\rho' \stackrel{\text{def}}{=} 2^{-j} \cdot \gamma \cdot 2^{i+j}\epsilon/4 = \gamma \cdot 2^i\epsilon/4$. Thus, a sample of $\frac{t}{2^i\epsilon} = \frac{c}{\gamma \cdot 2^i\epsilon}$ random pairs in $\Gamma_S(v) \times (S \setminus \Gamma_S(v))$ will hit an edge with high constant probability and cause Step 4 (of sub-test$_i(v)$) to reject. The claim follows. ■

**Claim 3.2.3** *If Algorithm 3.1 accepts with probability at least* $1/3$*, then, for every* $i \in [\ell]$ *the number of vertices of degree at least* $\gamma \cdot 2^i\epsilon \cdot N$ *that are not i-good is at most* $2^{-i} \cdot N/4$.

Claim 3.2.3 follows by combining Claim 3.2.2 with the fact that Algorithm 3.1 invokes sub-test$_i$ on $10 \cdot 2^i$ random vertices (and using $(1-2^{-i}/4)^{10 \cdot 2^i} + 0.05 < \exp(-10/4) + 0.05 < 1/3$). Next, using the conclusion of Claim 3.2.3, we turn to construct a partition $(C_1, \ldots, C_t)$ of $[N]$ such that the following holds: the total number of missing edges (in $G$) within the $C_i$'s is at most $\epsilon \cdot N^2/2$ and the total number of (superfluous) edges between the $C_i$'s is at most $\epsilon \cdot N^2/2$. The partition is constructed in iterations. *We start with a motivating discussion.*

Note that any $i$-good vertex, $v$, yields a set of vertices (i.e., $\Gamma(v)$) that is "close" to being a clique, where "closeness" has a stricter meaning when $i$ is smaller. Specifically, by Condition 1, the number of missing edges between pairs of vertices in this set is at most $\gamma \cdot 2^i\epsilon \cdot |\Gamma(v)| \cdot N$. But we should also care about how this set "interacts" with the rest of the graph, which is where Condition 2 comes into play. Letting $C_v$ contain only the vertices in $\Gamma(v)$ that have less than $|\Gamma(v)|$ neighbors outside of $\Gamma(v)$, we upper-bound the number of edges going out of $C_v$ as follows: We first note that these edges are either edges between $C_v$ and $\Gamma(v) \setminus C_v$ or edges between $C_v$ and $[N] \setminus \Gamma(v)$. The number of edges of the first type is upper-bounded by $|C_v| \cdot |\Gamma(v) \setminus C_v|$, whereas $|\Gamma(v) \setminus C_v| \le 2^{-j_0}|\Gamma(v)|$ (by using Condition 2 with $j = j_0$, while noting that $\gamma \cdot 2^{i+j_0}\epsilon N = |\Gamma(v)|$ (since $j_0 = \log_2(|\Gamma(v)|/(\gamma \cdot 2^i\epsilon N))$)). Thus, the number of edges of the first type is upper-bounded by $|C_v| \cdot 2^{-j_0}|\Gamma(v)| = |C_v| \cdot \gamma 2^i\epsilon N \le \gamma 2^i\epsilon \cdot |\Gamma(v)| \cdot N$. The number of edges of the second type is upper-bounded by assigning each vertex $u \in C_v$ to the smallest $j \in [j_0]$ such that $|\Gamma(u) \setminus \Gamma(v)| < \gamma \cdot 2^{i+j}\epsilon \cdot N$. (This means that $u$ violates Condition 2 w.r.t $j-1$.) Thus, the number of edges of the second type is upper-bounded by

$$\sum_{j=1}^{j_0} 2^{-(j-1)}|\Gamma(v)| \cdot \gamma \cdot 2^{i+j}\epsilon \cdot N \;=\; 2j_0 \cdot \gamma 2^i \epsilon \cdot |\Gamma(v)| \cdot N, \tag{4}$$

where the equality follows from the definition of $j_0$. Thus, the total number of the edges of both types is upper-bounded by $(2j_0+1) \cdot \gamma 2^i \epsilon \cdot |\Gamma(v)| \cdot N$, which is upper-bounded by $3\ell \cdot \gamma 2^i \epsilon \cdot |\Gamma(v)| \cdot N$ (since $j_0 \leq \log_2(1/(\gamma \cdot 2^i \epsilon)) \leq \log_2(1/\gamma\epsilon) = (1+o(1)) \cdot \ell$).

The foregoing paragraph identifies a single (good) clique, while we wish to identify all cliques. Starting with $i = 1$, the basic idea is to identify new cliques by using $i$-good vertices that are not covered by previously identified cliques. If we are lucky and the entire graph is covered this way, then we halt. But it may indeed be the case that some vertices are left uncovered and that they are not $i$-good. At this point we invoke Claim 3.2.3 and conclude that these vertices either have low degree (i.e., have degree at most $\gamma \cdot 2^i \epsilon \cdot N$) or are relatively few in number (i.e., their number is at most $2^{-i} \cdot N/4$). Ignoring (for a moment) the vertices of low degree, we deal with the remaining vertices by invoking the same reasoning with respect to an incremented value of $i$ (i.e., $i \leftarrow i+1$). The key observation is that the number of violations, caused by cliques identified in each iteration $i$, is upper-bounded by the product of the number of vertices covered in that iteration (which is linearly related to $2^{-i}$) and the "density" of violations caused by each identified clique (which is linearly related to $2^i\epsilon$). Thus, intuitively, each iteration contributes $O(\ell\gamma\epsilon \cdot N^2)$ violations, and after the last iteration (i.e., $i = \ell$) we are left with at most $2^{-i} \cdot N/4 < (\epsilon/4)N$ vertices, which we can afford to identify as a single clique (or alternatively as isolated vertices).

Two problems, which were ignored by the foregoing description, arise from the fact that vertices that are identified as belonging to the clique $C_v$ (of some $i$-good vertex $v$) may belong either to previously identified cliques or to the set of vertices cast aside as having low degree. Our solution is to use only $i$-good vertices for which the majority of their neighbors do not belong to these two categories (i.e., vertices $v$ such that most of $\Gamma(v)$ belongs neither to previously identified cliques nor have low degree). This leads to the following description.

The partition reconstruction procedure. The iterative procedure is initiated with $C = L_0 = \emptyset$, $R_0 = [N]$ and $i = 1$, where $C$ denotes the set of vertices "covered" (by cliques) so far, $R_{i-1}$ denotes the set of "remaining" vertices after iteration $i-1$ and $L_{i-1}$ denotes the set of vertices cast aside (as having "low degree") in iteration $i-1$. In each iteration, a set $F_i$ is constructed, where each vertex $v \in F_i$ is used to determine a clique (or, more precisely, a subset that is close to being a clique). The procedure refers to a parameter $\beta = 1/(c_3\ell)$, where $c_3 > 1$ is a sufficiently large constant, which determines the "low degree" threshold (for each iteration). Recall that $\gamma = \Theta(\log^{-3}(1/\epsilon)) = \Theta(1/\ell^3)$, so that $\gamma = o(\beta)$. For $i = 1, \ldots, \ell$, the $i^{\text{th}}$ iteration proceeds as follows, where $F_i$ is initialized to $\emptyset$.

1. Pick an arbitrary vertex $v \in R_{i-1} \setminus C$ that satisfies the following three conditions

   (a) $v$ is $i$-good.

   (b) $v$ has sufficiently high degree; that is, $|\Gamma(v)| \geq \beta \cdot 2^i \epsilon \cdot N$.

   (c) $v$ has relatively few neighbors in $C$; that is, $|\Gamma(v) \cap C| \leq |\Gamma(v)|/4$.

   If no such vertex exists, define $L_i = \{v \in R_{i-1} \backslash C : |\Gamma(v)| < \beta \cdot 2^i \epsilon \cdot N\}$ and $R_i = R_{i-1} \backslash (L_i \cup C)$. If $i < \ell$ then proceed to the next iteration, and otherwise terminate.

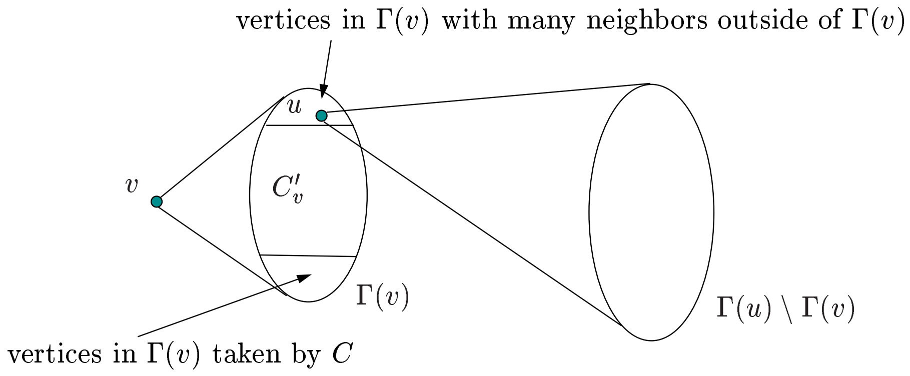


Figure 1: An Illustration for the clique collection partition reconstruction procedure.

2. For a vertex $v$ as selected in Step 1, let $C_v = \{u \in \Gamma(v) : |\Gamma(u) \setminus \Gamma(v)| < |\Gamma(v)|\}$. Form a new clique with the vertex set $C'_v \leftarrow C_v \setminus C$, and update $F_i \leftarrow F_i \cup \{v\}$ and $C \leftarrow C \cup C'_v$.

For an illustration, see Figure 1. Note that by Condition 1c, for every $v \in F_i$, it holds that $|C'_v| \geq |C_v| - (|\Gamma(v)|/4)$, whereas by $i$-goodness[5] (and $j_0 = \log_2(|\Gamma(v)|/(\gamma \cdot 2^i \epsilon N)) \geq \log_2(\beta/\gamma) = \omega(1)$) we have $|C_v| > (1 - o(1)) \cdot |\Gamma(v)|$. Thus, quality guarantees that are quantified in terms of $|\Gamma(v)|$ translate well to similar guarantees in terms of $|C'_v|$. This fact, combined with the fact that $C_v$ cannot contain many low degree vertices (i.e., vertices cast aside in prior iterations as having low degree), plays an important role in the following analysis.

**Claim 3.2.4** *Assume that $\gamma \leq \beta/(48\ell^2)$. Referring to the partition reconstruction procedure, for every $i \in [\ell]$, the following holds.*

1. *The number of missing edges inside the cliques formed in iteration $i$ is at most $8\gamma\epsilon \cdot N^2$; that is,*

$$\left| \bigcup_{v \in F_i} \{(u,w) \in C'_v \times C'_v : (u,w) \notin E\} \right| \leq 8\gamma\epsilon \cdot N^2. \tag{5}$$

2. *The number of* ("superfluous") *edges between cliques formed in iteration $i$ and either $R_i$ or other cliques formed in the same iteration is at most $24\ell \cdot \gamma\epsilon \cdot N^2$; actually,*

$$\left| \bigcup_{v \in F_i} \{(u,w) \in C'_v \times (R_{i-1} \setminus C'_v) : (u,w) \in E\} \right| \leq 24\ell \cdot \gamma\epsilon \cdot N^2. \tag{6}$$

3. *$|R_i| \leq 2^{-i} \cdot N$ and $|L_i| \leq 2^{-(i-1)} \cdot N$.*

Thus, the total number of violations caused by the cliques that are formed by the foregoing procedure is upper-bounded by $(24 + o(1))\ell^2 \cdot \gamma\epsilon \cdot N^2 = o(\epsilon N^2)$, assuming $\gamma = o(\ell^{-2})$. (Recall that $\ell \stackrel{\text{def}}{=} \log_2(1/\epsilon) + 2$, and that we shall set $\gamma = \Theta(\log^{-3}(1/\epsilon))$ and $\beta = \Theta(\log^{-1}(1/\epsilon))$.)

Proof: We prove all items simultaneously, by induction from $i = 0$ to $i = \ell$. Needless to say, all items hold vacuously for $i = 0$, and thus we focus on the induction step.

[5] Every $v \in F_i$ is $i$-good and thus satisfies $|C_v| > (1 - 2^{-j_0}) \cdot |\Gamma(v)|$.

Starting with Item 1, we note that every $v \in F_i$ is $i$-good and thus the number of edges missing in $C'_v \times C'_v \subseteq \Gamma(v) \times \Gamma(v)$ is at most $\gamma 2^i \epsilon \cdot |\Gamma(v)| \cdot N < 2\gamma 2^i \epsilon \cdot |C'_v| \cdot N$, where the inequality follows from $|C'_v| > |\Gamma(v)|/2$ (which follows by combining $|C'_v| \geq |C_v| - (\Gamma(v)|/4)$ and $|C_v| \geq (1 - 2^{-j_0}) \cdot |\Gamma(v)|$, where $j_0 = \log_2(|\Gamma(v)|/(\gamma \cdot 2^i \epsilon N)) > 2$). Observe that the $i$-goodness of $v$, combined with $|\Gamma(v)| \geq \beta \cdot 2^i \epsilon \cdot N$ and the relation between $\gamma$ and $\beta$ (i.e., $\gamma = o(\beta)$), implies that $\Gamma(v)$ contains at least $0.99 \cdot |\Gamma(v)|$ vertices of degree exceeding $0.99 \cdot |\Gamma(v)|$. This implies that $|\Gamma(v) \cap (\bigcup_{j \in [i-1]} L_j)| < |C_v|/4$, because $|\Gamma(v)| \geq \beta 2^i \epsilon \cdot N$ whereas every vertex in $\bigcup_{j \in [i-1]} L_j$ has degree at most $\beta 2^{i-1} \epsilon \cdot N$. Observing that $C'_v = (C'_v \cap R_{i-1}) \cup (C'_v \cap \bigcup_{j \in [i-1]} L_j)$, it follows that $|\bigcup_{v \in F_i} C'_v \cap R_{i-1}| > |\bigcup_{v \in F_i} C'_v|/2$, and thus $\sum_{v \in F_i} |C'_v| \leq 2|R_{i-1}|$. Combining all these bounds, we obtain

$$\begin{aligned}
\left| \bigcup_{v \in F_i} \{(u,w) \in C'_v \times C'_v : (u,w) \notin E\} \right| &= \sum_{v \in F_i} |\{(u,w) \in C'_v \times C'_v : (u,w) \notin E\}| && (7) \\
&\leq 2\gamma 2^i \epsilon \cdot \sum_{v \in F_i} |C'_v| \cdot N && (8) \\
&\leq 2\gamma 2^i \epsilon \cdot 2|R_{i-1}| \cdot N. && (9)
\end{aligned}$$

Using the induction hypothesis regarding $R_{i-1}$ (i.e., $|R_{i-1}| < 2^{-(i-1)} \cdot N$), Item 1 follows.

Item 2 is proved in a similar fashion. Recall that in the motivating discussion (i.e., the text preceding and following Eq. (4)) we showed that the $i$-goodness of $v$ (which follows from $v \in F_i$) implies that the number of edges in $C'_v \times (R_{i-1} \setminus C'_v) \subseteq C_v \times ([N] \setminus C_v)$ is at most $3\ell \cdot \gamma 2^i \epsilon \cdot |\Gamma(v)| \cdot N$. Since we have shown that $|C'_v| \geq |\Gamma(v)|/2$, this expression is upper-bounded by $6\ell \cdot \gamma 2^i \epsilon \cdot |C'_v| \cdot N$. Using again $\sum_{v \in F_i} |C'_v| < 2|R_{i-1}|$ and $|R_{i-1}| < 2^{-(i-1)} \cdot N$, we establish Item 2.

Turning to Item 3, we first note that $L_i \subseteq R_{i-1}$ and thus $|L_i| \leq |R_{i-1}| \leq 2^{-(i-1)} \cdot N$. As for $R_i$, it may contain only vertices that are neither in $L_i$ nor in $\bigcup_{v \in F_i} C'_v$. It follows that for every $v \in R_i$ either $v$ is not $i$-good (although it has degree at least $\beta \cdot 2^i \epsilon \cdot N$) or it has at least $|\Gamma(v)|/4$ neighbors in previously identified cliques (which implies $|\Gamma(v) \cap (\bigcup_{w \in \bigcup_{j \in [i]} F_j} C'_w)| \geq |\Gamma(v)|/4$). By Claim 3.2.3, the number of vertices of the first type is at most $2^{-i} \cdot N/4$. As for vertices of the second type, each such vertex $v$ (in $R_i$) requires at least $|\Gamma(v)|/4 \geq \beta \cdot 2^i \epsilon \cdot N/4$ edges from $C' \stackrel{\text{def}}{=} \bigcup_{w \in \bigcup_{j \in [i]} F_j} C'_w$ to it (because $C'$ is the set of vertices covered by previously identified cliques at the time iteration $i$ is completed). By Item 2, the total number of edges going out from $C'$ to $R_i$ is at most $i \cdot 24\ell \cdot \gamma\epsilon \cdot N^2 \leq 24\ell^2 \cdot \gamma\epsilon \cdot N^2$. On the other hand, as noted above, each vertex of the second type has at least $\beta \cdot 2^i \epsilon \cdot N/4$ edges incident to vertices in $C'$. Hence, the number of vertices of the second type is upper-bounded by

$$\frac{24\ell^2 \cdot \gamma\epsilon \cdot N^2}{\beta \cdot 2^i \epsilon \cdot N} = \frac{24\ell^2 \cdot \gamma}{\beta} \cdot 2^{-i} N, \tag{10}$$

Thus, $|R_i| \leq ((1/4) + 24\ell^2 \gamma \beta^{-1}) \cdot 2^{-i} \cdot N$, and, for $\gamma \leq \beta/(48\ell^2)$, we get that $|R_i| \leq 2^{-i} \cdot N$. ■

**Completing the reconstruction and its analysis.** The foregoing construction leaves "unassigned" the vertices in $R_\ell$ as well as some of the vertices in $L_1, \ldots, L_\ell$. (Note that some vertices in $\bigcup_{i=1}^{\ell-1} L_i$ may be placed in cliques constructed in later iterations, but there is no guarantee that this actually happens.) We now assign each of these remaining vertices to a singleton clique (i.e., an isolated vertex). The number of violations caused by this assignment equals the number of edges with both endpoints in $R' \stackrel{\text{def}}{=} R_\ell \cup \bigcup_{i=1}^{\ell} L_i$, because edges with a single endpoint in $R'$ were already accounted for in Item 2 of Claim 3.2.4. Nevertheless, we upper-bound the number of violations by the total number of edges adjacent at $R'$, which in turn is upper-bounded by

$$\begin{aligned} \sum_{v \in R_\ell \cup \bigcup_{i\in[\ell]} L_i} |\Gamma(v)| &\leq |R_\ell| \cdot N + \sum_{i=1}^{\ell} \sum_{v\in L_i} |\Gamma(v)| && (11) \\ &\leq \frac{\epsilon N}{4} \cdot N + \sum_{i=1}^{\ell} 2^{-(i-1)} N \cdot \beta 2^i \epsilon N && (12) \\ &= \frac{\epsilon}{4} \cdot N^2 + 2\ell \cdot \beta \cdot \epsilon N^2. && (13) \end{aligned}$$

For $\beta \leq 1/(8\ell)$, it follows that the number of these edges is smaller than $\epsilon N^2/2$. Combining this with the bounds on the number of violating edges (or non-edges) as provided by Claim 3.2.4, the lemma follows. Note that the foregoing uses $\beta \leq 1/(8\ell)$ and well as $\gamma \leq \beta/(48\ell^2) = o(\ell^2)$, which can be satisfied by setting $\beta = \Theta(\log^{-1}(1/\epsilon))$ and $\gamma = \Theta(\log^{-3}(1/\epsilon))$, since $\ell = \log_2(1/\epsilon) + 2$. ■

# 4 The Non-Adaptive Query Complexity of Clique Collection

In this section we study the non-adaptive query complexity of clique collection. We first establish the lower bound claimed in Part 2 of Theorem 1.1, and next show that this lower bound is tight.

## 4.1 The Lower Bound

In this section we establish Part 2 of Theorem 1.1. Specifically, for every value of $\epsilon > 0$, we consider two different sets of graphs, one consisting of graphs in $\mathcal{CC}$ and the other consisting of graphs that are $\epsilon$-far from $\mathcal{CC}$, and show that a non-adaptive algorithm of query complexity $o(\epsilon^{-4/3})$ cannot distinguish between graphs selected at random in these sets. Each set is actually determined by a single graph and all possible permutations of the vertex names.

### 4.1.1 The two sets

The first set, denoted $\mathcal{CC}_\epsilon$, contains all $N$-vertex graphs such that each graph consists of $(3\epsilon)^{-1}$ cliques, and each clique has size $3\epsilon \cdot N$. It will be instructive to partition these $(3\epsilon)^{-1}$ cliques into $(6\epsilon)^{-1}$ pairs (each consisting of two cliques). The second set, denoted $\mathcal{BCC}_\epsilon$, contains all $N$-vertex graphs such that each graph consists of $(6\epsilon)^{-1}$ bi-cliques, and each bi-clique has $3\epsilon \cdot N$ vertices on each side. For an illustration, see Figure 2. Indeed, $\mathcal{CC}_\epsilon \subseteq \mathcal{CC}$, whereas, as we show next, the graphs in $\mathcal{BCC}_\epsilon$ are all $\epsilon$-far from $\mathcal{CC}$.

**Claim 4.1** *Every graph in $\mathcal{BCC}_\epsilon$ is $\epsilon$-far from $\mathcal{CC}$.*

**Proof:** Let $G = ([N], E)$ be a graph in $\mathcal{BCC}_\epsilon$, let $(V_j^1, V_j^2)$ be the pair of sets of vertices in its $j^{\text{th}}$ biclique, and let $V_j = V_j^1 \cup V_j^2$. For any partition $\mathcal{P} = (X_1, \ldots, X_\ell)$ of $[N]$, let $\Delta_G(\mathcal{P})$ denote the number of edge modifications that are required in order to make the sets $X_1, \ldots, X_\ell$ into cliques with no edges between them. Then,

$$\Delta_G(\mathcal{P}) = \sum_{i=1}^{\ell} |\overline{E}(X_i)| + \sum_{i<i'} |E(X_i, X_{i'})| \,, \tag{14}$$

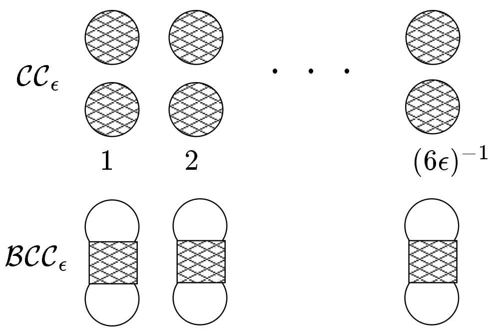


Figure 2: An Illustration for the lower bound construction that establishes Part 2 of Theorem 1.1.

where $\overline{E}(X_i)$ denotes the set of (unordered) pairs of (different) vertices in $X_i$ that do not have an edge between them. Thus, the distance between $G$ and $\mathcal{CC}$ is $N^{-2}$ times the minimum, taken over all partitions $\mathcal{P}$, of $\Delta_G(\mathcal{P})$. We need to show that $\Delta_G(\mathcal{P}) > \epsilon N^2$, for every partition $\mathcal{P}$.

We first observe that, without loss of generality, we may assume that each set $X_i$ intersects at most one $V_j$. This is true since otherwise, by refining the partition (i.e., replacing each $X_i$ with the collection of all nonempty $X_i \cap V_j$), the value of $\Delta_G(\cdot)$ can only decrease (because there are no edges between the different $V_j$'s, hence this refinement causes no new violations). Next, we show that, without loss of generality, we may also assume that each $V_j$ intersects at most one $X_i$. To see why this is true, consider the case that a set $V_j$ has a non-empty intersection with more than one $X_i \subseteq V_j$, and let $\alpha_i = |V_j^1 \cap X_i|/|V_j^1|$, and $\beta_i = |V_j^2 \cap X_i|/|V_j^2|$ (so that $\sum_i \alpha_i = 1$ and $\sum_i \beta_i = 1$). Let $\mathcal{P}'$ be the partition that replaces all $X_i$'s that intersect $V_j$ with a single set. Then, if we denote $|V_j^1| = |V_j^2|$ by $K$, we have

$$
\begin{aligned}
\Delta_G(\mathcal{P}) - \Delta_G(\mathcal{P}') &= \sum_{i=1}^{\ell} |E(V_j^1 \cap X_i, V_j^2 \setminus X_i)| && (15)\\
&\quad -\frac{1}{2} \cdot \sum_{i=1}^{\ell} |\overline{E}(V_j^1 \cap X_i, V_j^1 \setminus X_i)| - \frac{1}{2} \cdot \sum_{i=1}^{\ell} |\overline{E}(V_j^2 \cap X_i, V_j^2 \setminus X_i)| && (16)\\
&= \sum_{i=1}^{\ell} \left( \alpha_i K \cdot (1-\beta_i)K - \left( \frac{1}{2}\alpha_i K(1-\alpha_i)K + \frac{1}{2}\beta_i K(1-\beta_i)K \right) \right) && (17)\\
&= \frac{K^2}{2} \cdot \sum_{i=1}^{\ell} (\alpha_i - \beta_i)^2 \;\; \geq \;\; 0 \, , && (18)
\end{aligned}
$$

where $\overline{E}(Y, Z)$ denotes the set of pairs of vertices in $Y \times Z$ that do not have an edge between them. Hence, $\Delta_G(\mathcal{P}) \geq \Delta_G(\mathcal{P}')$, meaning that the distance can only decrease by taking the union of all sets $X_i$ that intersect $V_j$. It follows that it suffices to compute $\Delta_G(\mathcal{P})$ for the partition $\mathcal{P} = \{V_j\}_{j=1}^{1/6\epsilon}$. For this partition we get

$$
\begin{aligned}
\Delta(\mathcal{P}) &= \sum_{j=1}^{1/6\epsilon} \left( |E(V_j^1, V_j^1)| + |E(V_j^2, V_j^2)| \right) && (19)\\
&= \frac{1}{6\epsilon} \cdot \left( 9\epsilon^2 N^2 - 3\epsilon N \right) \; > \; \epsilon N^2 \, , && (20)
\end{aligned}
$$

using $\epsilon > 1/N$. The claim follows. ■

### 4.1.2 The indistinguishability result

In order to motivate the claim that a non-adaptive algorithm of query complexity $o(\epsilon^{-4/3})$ cannot distinguish between graphs selected at random in these sets, consider the (seemingly best such) algorithm that selects $o(\epsilon^{-2/3})$ vertices and inspects the induced subgraph. Consider the partition of a graph in $\mathcal{CC}_\epsilon$ into $(6\epsilon)^{-1}$ pairs of cliques, and correspondingly the partition of a graph in $\mathcal{BCC}_\epsilon$ into $(6\epsilon)^{-1}$ bi-cliques. Then, the probability that a sample of $o(\epsilon^{-2/3})$ vertices contains at least three vertices that reside in the same part (of $6\epsilon \cdot N$ vertices) is $o(\epsilon^{-2/3})^3 \cdot (6\epsilon)^2 = o(1)$. On the other hand, if this event does not occur, then the answers obtained from both graphs are indistinguishable (because in each case a random pair of vertices residing in the same part is connected by an edge with probability very close to 1/2). As is outlined next, this intuition extends to an arbitrary non-adaptive algorithm.

Specifically, by an averaging argument, it suffices to consider deterministic algorithms, which are fully specified by the sequence of queries that they make and their decision on each corresponding sequence of answers. Recall that these (fixed) queries are elements of $[N] \times [N]$. We shall show that, for every sequence of $o(\epsilon^{-4/3})$ queries, the answers provided by a randomly selected element of $\mathcal{CC}_\epsilon$ are statistically close to the answers provided by a randomly selected element of $\mathcal{BCC}_\epsilon$. We shall use the following notation: For an $N$-vertex graph $G$ and a query $(u, v)$, we denote the corresponding answer by $\mathrm{ans}_G(u, v)$; that is, $\mathrm{ans}_G(u, v) = 1$ if $\{u, v\}$ is an edge in $G$ and $\mathrm{ans}_G(u, v) = 0$ otherwise.

**Lemma 4.2** *Let $G_1$ and $G_2$ be random $N$-vertex graphs uniformly distributed in $\mathcal{CC}_\epsilon$ and $\mathcal{BCC}_\epsilon$, respectively. Then, for every sequence $(v_1, v_2), \ldots, (v_{2q-1}, v_{2q}) \in [N] \times [N]$, where the $v_i$'s are not necessarily distinct, it holds that the statistical difference between $\mathrm{ans}_{G_1}(v_1, v_2), \ldots, \mathrm{ans}_{G_1}(v_{2q-1}, v_{2q})$ and $\mathrm{ans}_{G_2}(v_1, v_2), \ldots, \mathrm{ans}_{G_2}(v_{2q-1}, v_{2q})$ is $O(q^{3/2}\epsilon^2)$.*

Part 2 of Theorem 1.1 follows (cf., also, Section 2.3).

**Proof:** We consider a 1-1 correspondence, denoted $\phi$, between the vertices of an $N$-vertex graph in $\mathcal{CC}_\epsilon \cup \mathcal{BCC}_\epsilon$ and triples in $[(6\epsilon)^{-1}] \times \{1, 2\} \times [3\epsilon \cdot N]$. Specifically, $\phi(v) = (i, j, w)$ indicates that $v$ resides in the $j^{\text{th}}$ "side" of the $i^{\text{th}}$ part of the graph, and it is vertex number $w$ in this set. That is, for a graph in $\mathcal{CC}_\epsilon$ the pair $(i, j)$ indicates the $j^{\text{th}}$ clique in the $i^{\text{th}}$ pair of cliques, whereas for a graph in $\mathcal{BCC}_\epsilon$ the pair $(i, j)$ indicates the $j^{\text{th}}$ side in the $i^{\text{th}}$ bi-clique. Consequently, the answers provided by uniformly distributed $G_1 \in \mathcal{CC}_\epsilon$ and $G_2 \in \mathcal{BCC}_\epsilon$ can be emulated by the following two corresponding random processes.

1. The process $A_1$ selects uniformly a bijection $\phi : [N] \to [(6\epsilon)^{-1}] \times \{1, 2\} \times [3\epsilon \cdot N]$ and answers each query $(u, v) \in [N] \times [N]$ by 1 if and only if $\phi(u)$ and $\phi(v)$ agree on their first two coordinates (and differ on the third). That is, for $\phi(u) = (i_1, j_1, w_1)$ and $\phi(v) = (i_2, j_2, w_2)$, it holds that $A_1(u, v) = 1$ if and only if both $i_1 = i_2$ and $j_1 = j_2$ (and $w_1 \neq w_2$).

2. The process $A_2$ selects uniformly a bijection $\phi : [N] \to [(6\epsilon)^{-1}] \times \{1, 2\} \times [3\epsilon \cdot N]$ and answers each query $(u, v) \in [N] \times [N]$ by 1 if and only if $\phi(u) = (i, j, w_1)$ and $\phi(v) = (i, 3 - j, w_2)$. That is, for $\phi(u) = (i_1, j_1, w_1)$ and $\phi(v) = (i_2, j_2, w_2)$, it holds that $A_2(u, v) = 1$ if and only if $i_1 = i_2$ but $j_1 \neq j_2$.

Let us denote by $\phi'(v)$ (resp., $\phi''(v)$ and $\phi'''(v)$) the first (resp., second and third) coordinates of $\phi(v)$; that is, $\phi(v) = (\phi'(v), \phi''(v), \phi'''(v))$. Then, both processes answer the query $(u, v)$ with 0 if $\phi'(u) \neq \phi'(v)$, and the difference between the processes is confined to the case that $\phi'(u) = \phi'(v)$. Specifically, conditioned on $\phi'(u) = \phi'(v)$ (and $\phi'''(u) \neq \phi'''(v)$), it holds that $A_1(u, v) = 1$ if and only if $\phi''(u) = \phi''(v)$, whereas $A_2(u, v) = 1$ if and only if $\phi''(u) \neq \phi''(v)$. However, since the (random) value of $\phi''$ is not present at the answer, the forgoing difference may go unnoticed. The foregoing considerations apply to a single query, but things may change in case of several queries. For example, if $\phi'(u) = \phi'(v) = \phi'(w)$ then the answers to $(u, v), (v, w)$ and $(w, v)$ will indicate whether we are getting answers from $A_1$ or from $A_2$ (since $A_1$ will answer positively on an odd number of these queries whereas $A_2$ will answer positively on an even number). In general, the event that allows distinguishing the two processes is an odd cycle of vertices that have the same $\phi'$ value. Minor differences may also be due to equal $\phi'''$ values, and so we also consider these in our "bad" event. For sake of simplicity, the bad event is defined more rigidly as follows, where the first condition represents the essential aspect and the second is a technicality.

**Definition 4.2.1** *We say that $\phi$ is* bad (w.r.t. the sequence $(v_1, v_2), \ldots, (v_{2q-1}, v_{2q}) \in [N] \times [N]$), *if any of the following two conditions hold:*

1. *For some $i \in [(6\epsilon)^{-1}]$, the subgraph $Q_i = (V_i, E_i)$, where $V_i = \{v_k : k \in [2q] \wedge \phi'(v) = i\}$ and $E_i = \{\{v_{2k-1}, v_{2k}\} : v_{2k-1}, v_{2k} \in V_i\}$, contains a simple cycle.*
2. *There exists $i \neq j \in [2q]$ such that $\phi'''(v_i) = \phi'''(v_j)$.*

Indeed, the query sequence $(v_1, v_2), \ldots, (v_{2q-1}, v_{2q})$ will be fixed throughout the rest of the proof, and so we shall omit it from our terminology.

**Claim 4.2.2** *The probability that a uniformly distributed bijection $\phi$ is bad is at most*

$$6000 \cdot q^{3/2}\epsilon^2 + \frac{2q^2}{3\epsilon N} \tag{21}$$

Proof: We start by upper-bounding the probability that the second event in Definition 4.2.1 holds. This event is the union of $\binom{2q}{2}$ sub-events, and each sub-event holds with probability $1/(3\epsilon \cdot N)$. Thus, we obtain a probability (upper) bound of $2q^2/3\epsilon N$. As for the first event, for every $t \geq 3$, we upper-bound the probability that some $Q_i$ contains a simple cycle of length $t$. We observe that the query graph $Q = (V_Q, E_Q)$, where $V_Q = \{v_k : k \in [2q]\}$ and $E_Q = \{\{v_{2k-1}, v_{2k}\} : k \in [q]\}$), contains at most $(2q)^{t/2}$ cycles of length $t$ (cf. [A81, Thm. 3]), whereas the probability that a specific simple $t$-cycle is contained in some $Q_i$ is $(6\epsilon)^{t-1}$. Thus, the probability of the first event is upper-bounded by

$$\begin{aligned} \sum_{t \geq 3} (2q)^{t/2} \cdot (6\epsilon)^{t-1} &< \sum_{t \geq 3} \left(\sqrt{2q} \cdot 6 \cdot \epsilon^{(t-1)/t}\right)^t && (22) \\ &< \sum_{t \geq 3} \left(9\sqrt{q} \cdot \epsilon^{2/3}\right)^t, && (23) \end{aligned}$$

which is upper-bounded by $2 \cdot (9\sqrt{q} \cdot \epsilon^{2/3})^3 < 1500q^{3/2}\epsilon^2$, provided $9\sqrt{q} \cdot \epsilon^{2/3} < 1/2$ (and the claim holds trivially otherwise). ■

**Claim 4.2.3** *Conditioned on the bijection $\phi$ not being bad, the sequences $(A_1(v_1, v_2), \ldots, A_1(v_{2q-1}, v_{2q}))$ and $(A_2(v_1, v_2), \ldots, A_2(v_{2q-1}, v_{2q}))$ are identically distributed.*

Proof: Noting that Definition 4.2.1 only refers to $\phi'$ and $\phi'''$, we fix any choice of $\phi'$ and $\phi'''$ that yields a good $\phi$ and consider the residual random choice of $\phi''$. Referring to the foregoing subgraphs $Q_i$'s, recall that pairs with endpoints in different $Q_i$'s are answered by 0 in both processes. Note that (by the second condition in Definition 4.2.1) the hypothesis implies that $\phi'''$ assigns different values to the different vertices in $\{v_k : k \in [2q]\}$, and it follows that $\phi''$ assigns these vertices values that are uniformly and independently distributed in $\{1,2\}$. Now, using the first condition in Definition 4.2.1, the hypothesis implies that each $Q_i$ is a forest. This implies that, for each of the two processes, the answer assigned to each edge in $Q_i$ is independent of the answer given to other edges of $Q_i$. That is, we assert that (in each of the two processes) the edges of each forest $Q_i = (V_i, E_i)$ are assigned a sequence of answers that is uniformly distributed in $\{0,1\}^{|E_i|}$. To formally prove this assertion, consider the constraints on the $\phi''$-values (of $V_i$) that arise from any possible sequence of answers. These constraints form a system of $|E_i|$ linear equations over $GF(2)$ with variables corresponding to the possible $\phi''$-values and constant terms encoding possible equality and inequality constraints.[6] Note that the (coefficients of the) linear systems are not affected by the identity of the process, which does effect the free terms. Furthermore, this linear system is of full rank; and thus, for each of the two processes and each sequence of answers, the corresponding system has $2^{|V_i|-|E_i|} = 2$ solutions (i.e., possible assignments to $\phi''$ restricted to $V_i$). Thus, in each of the two processes, each query is answered by the value 1 with probability exactly 1/2, independently of the answers to all other queries. The claim follows. ■

Combining Claims 4.2.2 and 4.2.3, it follows that the statistical distance between the sequences $(A_1(v_1, v_2), \ldots, A_1(v_{2q-1}, v_{2q}))$ and $(A_2(v_1, v_2), \ldots, A_2(v_{2q-1}, v_{2q}))$ is at most $O(q^{3/2}\epsilon^2 + q^2(\epsilon N)^{-1})$, and the lemma follows for sufficiently large $N$. ■

## 4.2 A Matching Upper-Bound

In this section we establish Part 3 of Theorem 1.1. We mention that this improves over the $\widetilde{O}(\epsilon^{-2})$ bound of [AS, Thm. 2] (which is based on inspecting the subgraph induced by a random set of $O(\epsilon^{-1}\log(1/\epsilon))$ vertices).

**Algorithm 4.3** (non-adaptive test for $\mathcal{CC}$): *On input $N$ and $\epsilon$ and oracle access to a graph $G = ([N], E)$, set $\ell = \log_2(1/\epsilon)$ and proceeds as follows.*

1. *Select a random sample of $s \stackrel{\rm def}{=} \Theta(\epsilon^{-2/3})$ vertices, denoted $S$, and examine all vertex pairs* (in $S \times S$).
2. *For $i = 1, \ldots, (2\ell/3) + \Theta(1)$, uniformly select a subset $S_i \subseteq S$ of cardinality $s_i \stackrel{\rm def}{=} \Theta(2^i)$ and a sample of $\widetilde{\Theta}(\epsilon^{-1})/s_i$ vertices (in $[N]$), denoted $R_i$, and examine all the vertex pairs in $S_i \times R_i$.*
3. *The tester accepts if and only if its view of the graph as obtained in Steps 1-2 is consistent with some graph in $\mathcal{CC}$. Namely, let $g' : ((S \times S) \cup \bigcup_{i=1}^{\ell'} (S_i \times R_i)) \to \{0,1\}$ be the function determined by the answers obtained in Steps 1–2. Then, the tester accepts if and only if for $S' = S \cup \bigcup_{i=1}^{\ell'} R_i$, the function $g'$ can be extended to a function over $S' \times S'$ that represents a graph in $\mathcal{CC}$.*

The query complexity of Algorithm 4.3 is dominated by Step 1, which uses $O(\epsilon^{-2/3})^2 = O(\epsilon^{-4/3})$ queries. Step 3 can be implemented efficiently by first constructing the connected components of the graph defined by the positive answers obtained in Steps 1–2, and then checking whether or not all the negative answers (obtained in Steps 1–2) refer to pairs that reside in different components.

Clearly, Algorithm 4.3 accepts (with probability 1) any graph that is in $\mathcal{CC}$. It remains to analyze its behavior on graphs that are $\epsilon$-far from $\mathcal{CC}$.

[6] The condition $A_1(u,w) = 1$ iff $\phi''(u) = \phi''(v)$ is encoded by $\phi''(u) + \phi''(v) = A_1(u,w) + 1$, whereas the condition $A_2(u,w) = 1$ iff $\phi''(u) \neq \phi''(v)$ is encoded by $\phi''(u) + \phi''(v) = A_2(u,w)$.

**Lemma 4.4** *If $G = ([N], E)$ is $\epsilon$-far from $\mathcal{CC}$, then on input $N, \epsilon$ and oracle access to $G$, Algorithm 4.3 rejects with probability at least $2/3$.*

Part 3 of Theorem 1.1 follows.

**Overview of the proof of Lemma 4.4.** We say that a triple $(v, u, w)$ of (different) vertices (resp., a 3-set $\{v, u, w\} \subset [N]$) is a witness (for rejection) *if the subgraph* of $G$ *induced by* $\{v, u, w\}$ *contains exactly two edges.* Indeed, Algorithm 4.3 rejects if (and only if), for some witness $(v, u, w)$, the algorithm has made all three relevant queries (i.e., the queries $(v, u)$, $(u, w)$, and $(w, v)$).[7] A sufficient condition for this to happen is that either $\{v, u, w\} \subset S$ or, for some $i$, two of the vertices in $\{v, u, w\}$ belong to $S_i$, and the third belongs to $R_i$. Thus, we say that a witness is effective with respect to the said samples (i.e., $S$ and the $R_i$'s) if the foregoing sufficient condition holds. We shall show that, with probability at least $2/3$, the samples contain an effective witness.

Let $G' = (V, E')$ be a graph in $\mathcal{CC}$ that is closest to $G = (V, E)$, and let $(V_1, \ldots, V_t)$ be its partition into cliques. For the sake of simplicity, we shall refer to the $V_i$'s as cliques, even though they are not (necessarily) cliques in $G$, and we shall refer to the partition $(V_1, \ldots, V_t)$ as the *best possible partition* for $G$. Two main observations regarding this partition follow.

**Observation 1:** For every $i \in [t]$ and every $S \subseteq V_i$, it holds that $|E \cap (S \times (V_i \setminus S))| \geq |S \times (V_i \setminus S)|/2$, because otherwise replacing the clique $V_i$ by two cliques, $S$ and $V_i \setminus S$, yields a better partition for $G$.

**Observation 2:** For every $i \neq j \in [t]$, it holds that $|E \cap (V_i \times V_j)| \leq |V_i \times V_j|/2$, because otherwise replacing the two cliques $V_i$ and $V_j$ by a single clique $V_i \cup V_j$, yields a better partition for $G$.

Now, since $G$ is $\epsilon$-far from $\mathcal{CC}$, either there are at least $\frac{\epsilon}{2} \cdot N^2$ missing edges (in $G$) within these $V_i$'s or there are at least $\frac{\epsilon}{2} \cdot N^2$ superfluous edges between distinct $V_i$'s. We show that in either case, with high constant probability, the samples produced by Algorithm 4.3 contain an effective witness.

The pivot of the analysis is relating the fraction of bad vertex pairs (i.e., either missing "internal" edges or superfluous "external" edges) to the fraction of witnesses. Specifically, we shall show that the existence of $\frac{\epsilon}{2} \cdot N^2$ missing internal edges (resp., $\frac{\epsilon}{2} \cdot N^2$ superfluous external edges) implies the existence of $\Omega(\epsilon^2 N^3)$ witnesses. Intuitively, missing internal edges yield many witnesses, because $(v, u) \in (V_i \times V_i) \setminus E$ form a witness with any $w \in V_i \cap \Gamma(v) \cap \Gamma(u)$, whereas Observation 1 implies that $|V_i \cap \Gamma(v)| \geq |V_i|/2$ and most pairs in $(V_i \setminus \Gamma(v)) \times (V_i \cap \Gamma(v))$ are edges. Similar considerations, which rely on Observation 2, can be shown to imply that superfluous external edges yield many witnesses, intuitively because $(v, u) \in (V_i \times V_j) \cap E$ form a witness with any $w$ such that $|\Gamma(w) \cap \{u, v\}| = 1$, whereas many $w \in V_i \cup V_j$ satisfy this condition. These combinatorial considerations are detailed in Section 4.2.1.

It is tempting to think that we are done as soon as we establish the existence of $\Omega(\epsilon^2 N^3)$ witnesses. Unfortunately, this is not quite true. Indeed, if we were to select independently at random $O(\epsilon^{-2})$ triples and examine their internal edge relation, then we would have hit a witness with high probability. However, while Algorithm 4.3 does inspect the internal edge relations of $\Omega(\epsilon^{-2})$ triples (and each triple is uniformly distributed), these triples are not independently distributed. Thus, we shall establish additional features of the structure of the set of witnesses, and use these features to show that with high probability the random sample (as produced by Algorithm 4.3) contains an

[7] We note that only the (easy to establish) sufficiency of the foregoing rejection condition is used in the analysis.

effective witness. That is, these additional features, which are established in the elaborate parts of Claims 4.4.1 and 4.4.2, are instrumental to the detection of a witness (as analyzed in Claim 4.4.3).

Unfortunately, the implementation of the foregoing strategy is quite lengthy and complicated. Some readers may prefer to skip it and proceed directly to Section 5.

### 4.2.1 The structure of the set of witnesses

To facilitate the exposition, for every two sets $A, B \subset [N]$, we let $E(A, B)$ denote the set of edges with one endpoint in $A$ and another endpoint in $B$ (i.e., $E(A, B) \stackrel{\rm def}{=} E \cap (A \times B)$). For each vertex $v$ and $j \in [t]$, let

$$\Gamma_j(v) \stackrel{\rm def}{=} V_j \cap \Gamma(v) \;=\; \{u \in V_j : \{u, v\} \in E\} \tag{24}$$

and

$$\overline{\Gamma}_j(v) \stackrel{\rm def}{=} V_j \setminus (\Gamma(v) \cup \{v\}) \;=\; \{u \in (V_j \setminus \{v\}) : \{u, v\} \notin E\}. \tag{25}$$

If $v \in V_i$, then we use the shorthand: $\overline{\Gamma}(v) = \overline{\Gamma}_i(v)$. Indeed, $\overline{\Gamma}(v)$ corresponds to the set of *internal edges that are missed by vertex $v$*.

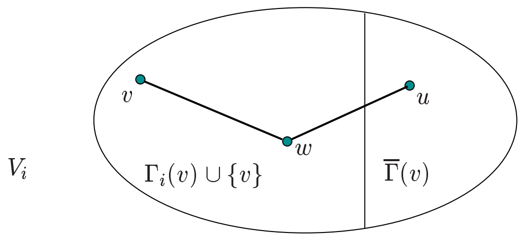


Figure 3: An Illustration for the proof of Claim 4.4.1.

Introduction to Claim 4.4.1. For every vertex $v$, the set $\overline{\Gamma}(v) \times \Gamma(v)$ contains pairs of vertices that may form witnesses together with $v$; that is, $(u, w) \in \overline{\Gamma}(v) \times \Gamma(v)$ forms a witness with $v$ if and only if $(u, w) \in E$. The basic claim asserts that the number of such pairs is at least $\Omega(|\overline{\Gamma}(v)|^2)$ (even when restricting $w$ to the same $V_i$ as $v$; see illustration in Figure 3). Moving to the elaborate claim, we encourage the reader to first consider the case that $F = \emptyset$. (In fact, this case is one of the two cases that will be actually used.) The point of Part 1 (in the elaborate claim) is to set the stage for Part 2, which upper-bounds the number of designated witnesses in which each $w$ appears, where this upper-bound is instrumental for the probabilistic analysis provided by Claim 4.4.3.

**Claim 4.4.1** (using missing internal edges):

Basic claim: *For every vertex $v$, the number of witnesses that contain $v$ is $\Omega(|\overline{\Gamma}(v)|^2)$.*

Elaborate claim: *For every* (possibly empty) *set $F$ of* ("forbidden") *vertex-pairs, where $F \subseteq \binom{[N]}{2} \setminus E$, the following holds:*

1. *For every $v \in [N]$ there exists a set $W_v \subseteq \overline{\Gamma}(v) \setminus \{u : \{v, u\} \in F\}$ such that*

$$\sum_{v \in [N]} |W_v| \;>\; \left( \sum_{v \in [N]} \frac{|\overline{\Gamma}(v)|}{4} \right) \;-\; 2 \cdot |F| \tag{26}$$

*and for every $u \in W_v$ there exists a set $W_{v,u} \subseteq (\Gamma(v) \cap \Gamma(u))$ of size at most $|W_v|$ such that*

$$\sum_{u \in W_v} |W_{v,u}| \ \geq \ |W_v|^2/4. \tag{27}$$

*Moreover, if $F = \emptyset$ then for every $v$ it holds that $|W_v| \geq |\overline{\Gamma}(v)|/4$.*

(Indeed, each $(u, v, w)$ such that $u \in W_v$ and $w \in W_{v,u}$ constitutes a witness.)

2. *For the sets $W_v$ and $W_{v,u}$ as in Part 1 of the claim, letting $U_w^{(2)} \stackrel{\rm def}{=} \{\{v, u\} : w \in W_{v,u}\}$ it holds that if each set $W_v$ has cardinality at most $\epsilon^{2/3}N/2$ then each $U_w^{(2)}$ has cardinality at most $\epsilon^{4/3}N^2$.*

It follows that the total number of witnesses is $\Omega(\sum_{v \in [N]} |\overline{\Gamma}(v)|^2)$. In particular, if the number of missing internal edges is at least $\frac{\epsilon}{2} \cdot N^2$ (i.e., $\sum_{v \in [N]} |\overline{\Gamma}(v)| \geq \epsilon \cdot N^2$), then the total number of witnesses is at least $N \cdot \Omega((\epsilon N)^2) = \Omega(\epsilon^2 \cdot N^3)$.

Proof: Using Observation 1, we note that for any choice of $i \in [t]$ and for every $v \in V_i$ it holds that

$$|\overline{\Gamma}(v)| \ = \ |V_i \setminus \{v\}| - |E(\{v\}, V_i \setminus \{v\})| \ \leq \ \frac{|V_i| - 1}{2} \ \leq \ |\Gamma_i(v)| \tag{28}$$

and

$$|E(\overline{\Gamma}(v), \Gamma_i(v)| \ = \ |E(\overline{\Gamma}(v), \Gamma_i(v) \cup \{v\})| \ > \ \frac{1}{2} \cdot |\overline{\Gamma}(v)| \cdot |\Gamma_i(v)| \, . \tag{29}$$

Letting $T_v = \{(v, u, w) : (u, w) \in \overline{\Gamma}(v) \times \Gamma_i(v)\}$, it follows that at least half of the triples $(v, u, w)$ in $T_v$ are witnesses (i.e., $(u, w) \in E$, $(u, v) \notin E$, and $(w, v) \in E$), whereas $|T_v| \geq |\overline{\Gamma}(v)|^2$. This establishes the basic claim.

Let us first establish the elaborate claim for the special case of $F = \emptyset$. In this case, for every $v \in V_i$, we consider the set

$$W_v \stackrel{\rm def}{=} \left\{ u \in \overline{\Gamma}(v) : |E(\{u\}, \Gamma_i(v))| \geq \frac{|\Gamma_i(v)|}{4} \right\} \, . \tag{30}$$

By Eq. (29), $\sum_{u \in \overline{\Gamma}(v)} |E(\{u\}, \Gamma_i(v))| \geq |\overline{\Gamma}(v)| \cdot |\Gamma_i(v)|/2$. It follows that $|W_v| \geq |\overline{\Gamma}(v)|/4$. We note that (by Eq. (30)), for every $u \in W_v$, it holds that $|\Gamma_i(v) \cap \Gamma(u)| \geq |\Gamma_i(v)|/4 \geq |W_v|/4$. Next, for every $u \in W_v$, let $W_{v,u}$ be an arbitrary subset of $|W_v|/4$ elements in $\Gamma_i(v) \cap \Gamma(u)$. Note that, indeed $W_v \subseteq \overline{\Gamma}(v)$ and for every $u \in W_v$ it holds that $W_{v,u} \subseteq \Gamma(v) \cap \Gamma(u)$. Recalling that $|W_v| \geq |\overline{\Gamma}(v)|/4$ and $|W_{v,u}| = |W_v|/4$, Part 1 follows.

To establish Part 2, we first note that if we select $W_{v,u}$ uniformly among all $(|W_v|/4)$-subsets of $\Gamma_i(v) \cap \Gamma(u)$, then, for any $w \in V_i$, the expected size of $U_w^{(2)}$ is upper-bounded by

$$\sum_{v \in V_i} \sum_{u \in W_v} \frac{|W_v|/4}{|\Gamma_i(v) \cap \Gamma(u)|} \ \leq \ \sum_{v \in V_i} \sum_{u \in W_v} \frac{|W_v|/4}{|V_i|/8} \ = \ \frac{2}{|V_i|} \cdot \sum_{v \in V_i} |W_v|^2 \tag{31}$$

where the inequality uses $|\Gamma_i(v) \cap \Gamma(u)| \geq |\Gamma_i(v)|/4 \geq |V_i|/8$. Thus, if $\frac{2}{|V_i|} \cdot \sum_{v \in V_i} |W_v|^2 \leq \epsilon^{4/3}N^2/2$ then, with overwhelmingly high probability, it holds that $|U_w^{(2)}| \leq \epsilon^{4/3}N^2$. Picking the sets (i.e., the $W_{v,u}$'s) so that none of the negligible probability events (associated with the different $w \in V_i$) occurs, we infer that $|U_w^{(2)}| > \epsilon^{4/3}N^2$ implies that $\sum_{v \in V_i} |W_v|^2 > \epsilon^{4/3}N^2|V_i|/4$ (which implies the existence of $v$ such that $|W_v| > \epsilon^{2/3}N/2$). Part 2 follows.

Note that so far we have established the (elaborate) claim for the special case of $F = \emptyset$. We now establish the general case by reduction to the former special case. We first modify the sets $W_v$, by omitting from each $W_v$ each vertex $u$ such that $\{v,u\} \in F$. This modification decreases $\sum_v |W_v|$ by at most $2|F|$. Next, we modify the sets $W_{v,u}$ by omitting from each $W_{v,u}$ a few elements, selected at random, so that $|W_{v,u}| = |W_v|/4$ holds (for the modified sets $W_v$). Clearly, Part 1 holds for the modified sets. To see that Part 2 holds too, we note that the foregoing argument only relies on the fact that $W_{v,u}$ is a random $(|W_v|/4)$-size subset of $\Gamma_i(v) \cap \Gamma(u)$, which is unaffected by $F$. The claim follows. ■

Another piece of notation. For every $i \in [t]$ and every $v \in V_i$, let

$$\Gamma'(v) \stackrel{\rm def}{=} \Gamma(v) \setminus V_i \tag{32}$$

denote the set of *vertices outside of* $V_i$ *that have a superfluous edge to* $v$. That is, $\Gamma'(v) = \bigcup_{j \neq i} \Gamma_j(v)$.

Introduction to Claim 4.4.2. For every vertex $v$, the set $\Gamma'(v)$ contains vertices $u$ such that $v$ and $u$ are part of a witness; specifically, $(v,u,w)$ is a witness if $u \in \Gamma'(v)$ and $|\Gamma(w) \cap \{v,u\}| = 1$. The basic claim asserts that the number of such pairs is at least $\Omega(|\Gamma'(v)|^2)$. Moving to the elaborate claim, we note that the greater complexity of Claim 4.4.2 (when compared to Claim 4.4.1) is reflected in the fact that even in the "simple" case of $F = \emptyset$ (which is treated in Part 1) we do not obtain a uniform bound on all $W_v$, but rather allow some exceptional vertices (which are shown to have small contribution to the sum of $\Gamma'(\cdot)$s). Furthermore, in this case, the basic claim does not follow from Part 1. In Part 2 we deal with a general forbidden set $F$, and get results analogous to (but quantitatively weaker than) the general case of Claim 4.4.1. Analogously to Claim 4.4.1, Part 2a sets the stage for Part 2b, which upper-bounds the number of designated witnesses in which each $w$ appears, where this upper-bound is instrumental for the probabilistic analysis provided by Claim 4.4.3.

**Claim 4.4.2** (using superfluous external edges):

Basic claim: *For every vertex $v$, the number of witnesses that contain $v$ is $\Omega(|\Gamma'(v)|^2)$.*

Elaborate claim: *There exist positive constants $c_1, \ldots, c_4$ such that the following holds:*

1. *For every $\alpha > 0$, if*

$$\sum_{v \in [N]} |\Gamma'(v)| \ > \ \frac{125}{\alpha} \cdot \sum_{v \in [N]} |\overline{\Gamma}(v)| \, , \tag{33}$$

*then for every $v \in [N]$ there exists a set $W_v \subseteq \Gamma'(v)$ such that letting $V' = \{v : |W_v| \geq |\Gamma'(v)|/c_1\}$ it holds that*

$$\sum_{v \in V'} |\Gamma'(v)| \ \geq \ (1-\alpha) \cdot \sum_{v \in [N]} |\Gamma'(v)| \, . \tag{34}$$

*In addition, for every $u \in W_v$ there exists a set $W_{v,u}$, which is either a subset of $\Gamma(v) \setminus \Gamma(u)$ or a subset of $\Gamma(u) \setminus \Gamma(v)$, such that $|W_{v,u}| \geq |W_v|/c_2$.*

(Indeed, each $(v,u,w)$ such that $u \in W_v$ and $w \in W_{v,u}$ constitutes a witness.)

2. *Let $F$ be any set of* "forbidden" *vertex-pairs, where $F \subseteq \bigcup_{i \neq j} E(V_i, V_j)$, and let $F(v) \stackrel{\rm def}{=} \{u : \{v,u\} \in F\} \subseteq \Gamma'(v)$, for every $v \in [N]$. Then:*

(a) *For each vertex $v$, there exists a subset $W_v \subseteq \Gamma'(v) \setminus F(v)$ such that*

$$\sum_{v\in[N]} |W_v| \; > \; \frac{1}{c_3} \cdot \left( \sum_{v\in[N]} |\Gamma'(v)| \right) \; - \; c_4 \cdot |F| \; . \tag{35}$$

*In addition, as in Part 1, for every $u \in W_v$, there exists a set $W_{v,u}$, which is either a subset of $\Gamma(v) \setminus \Gamma(u)$ or a subset of $\Gamma(u) \setminus \Gamma(v)$, such that $|W_{v,u}| \geq |W_v|/c_2$.*

(b) *For the sets $W_{v,u}$ as in Part 2a, let $U_w^{(2)} \stackrel{\rm def}{=} \{(v,u) : w \in W_{v,u}\}$. Then, if for every $v$ it holds that $|\Gamma'(v) \setminus F(v)| \leq \epsilon^{2/3}N/2$, then $U_w^{(2)}$ has cardinality at most $10\epsilon^{4/3}N^2$.*

*In all cases, it holds that $|W_{v,u}| \leq |W_v|$.*

It follows that the total number of witnesses is $\Omega(\sum_{v\in[N]} |\Gamma'(v)|^2)$. In particular, if the number of superfluous external edges is at least $\frac{\epsilon}{2} \cdot N^2$ (i.e., $\sum_{v\in[N]} |\Gamma'(v)| \geq \epsilon \cdot N^2$), then the total number of witnesses is at least $N \cdot \Omega((\epsilon N)^2) = \Omega(\epsilon^2 \cdot N^3)$.

Proof: The claim is proved by a (rather tedious) case analysis, which refers to a generic vertex $v$. In each of the cases, it is relatively easy to prove the basic claim, and things get more complicated when moving to Part 1 of the elaborate claim, and more so when moving to Part 2. Indeed, in our presentation we first establish Part 1, and only then move to Part 2 (which refers to a general set of forbidden pairs $F$).

Each case deals with a different subset of vertices. With the exception of one case, Part 1 is proved by presenting, for every relevant vertex $v$ (i.e., $v$ that satisfies the case hypothesis), a subset $W_v \subseteq \Gamma'(v)$ of size at least $|\Gamma'(v)|/c_1$ and adequate sets $W_{v,u}$ for each $u \in W_v$. Furthermore, it will be shown that the vertices covered by these (non-exceptional) cases account for at least a $1-\alpha$ fraction of the sum $\sum_v |\Gamma'(v)|$.

As in the proof of Claim 4.4.1, when we prove Part 2 we use $W_v \subseteq \Gamma'(v) \setminus F(v)$ and select the sets $W_{v,u}$ as random $\Theta(|W_v|)$-subsets of the sets of admissible elements. We note that when establishing Part 2, for each of the foregoing cases, we consider the restriction of $U_w^{(2)}$ to pairs $(v,u)$ such that $v$ obeys the case hypothesis. We show that if $|\Gamma'(v)| \leq \epsilon^{2/3}N/2$ for every such $v$, then the total contribution to $U_w^{(2)}$ of the corresponding pairs $(v,u)$ is at most $\epsilon^{4/3}N^2$. Since there are less than ten cases, Part 2 follows.

We stress that, while the following analysis refers to possible cases that may apply to a generic vertex $v$, we actually consider the set of all vertices that satisfy the hypothesis of each of these cases. (Hence, when we say that Part 1 (resp., Part 2) is established for the vertices that satisfy a particular case hypothesis, we mean that the contribution of these vertices is as claimed in the corresponding part.) We now turn to the actual case analysis.

Case 1: Much of $\boldsymbol{\Gamma'(v)}$ is contained in a single $\boldsymbol{V_j}$; that is, there exists an index $j$ such that $|\Gamma_j(v)| > |\Gamma'(v)|/10$. Fixing such an index $j$, we distinguish two subcases regarding the fraction of $V_j$ that is not covered by $\Gamma'(v)$ (i.e., the relative density of $\overline{\Gamma}_j(v)$ in $V_j$). For $v \notin V_j$, the natural case is that $|\overline{\Gamma}_j(v)| \geq |V_j|/10$ (see Case 1.1), and in this case we seek witnesses of the form $(v,u,w)$ such that $(u,w) \in (\Gamma_j(v) \times \overline{\Gamma}_j(v)) \cap E$ (i.e., $W_v \subseteq \Gamma_j(v)$ and $W_{v,u} \subseteq \overline{\Gamma}_j(v) \cap \Gamma_j(u)$). The other case (i.e., Case 1.2) is that $|\overline{\Gamma}_j(v)| < |V_j|/10$, where we seek witnesses of the form $(v,u,w)$ such that $v$ and $w$ resides in the same $V_i$, while $u$ resides in $V_j$, and $(v,u),(v,w) \in E$, while $(u,w) \notin E$. Details follow.

***Case 1.1:*** $|\overline{\Gamma}_j(v)| \geq |V_j|/10$. In this case, we let $W_v$ be a subset of the neighbors that $v$ has in $V_j$, that is, a subset of $\Gamma_j(v)$. For each $u \in W_v$ we let $W_{v,u}$ be a subset of the non-neighbors of $v$ in $V_j$ that are neighbors of $w$, that is, a subset of $\overline{\Gamma}_j(v) \cap \Gamma_j(u)$. Thus, for every $u \in W_v$ and $w \in W_{v,u}$, the triple $(v,u,w)$ is a witness. For an illustration, see Figure 4. Combining this case hypothesis (which asserts that $v$ has many non-neighbors in $V_j$) with Observation 1 (which guarantees many edges between neighbors and non-neighbors of $v$ in $V_j$), we obtain many (i.e., $\Omega(|\Gamma'(v)|^2)$) such witnesses, and the basic claim follows.

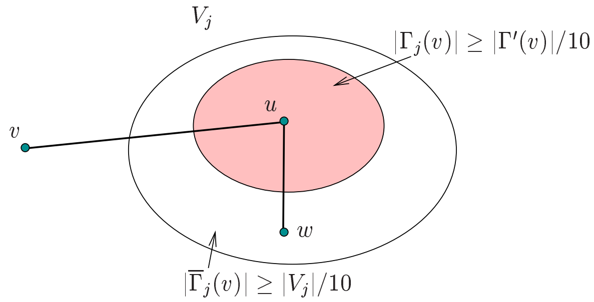


Figure 4: An Illustration for Case 1.1 in the proof of Claim 4.4.2.

In order to actually prove Parts 1 and 2, we now provide a more detailed description of the choice of $W_v$ and $W_{v,u}$. Let the subset of vertices for which the case (1.1) hypothesis holds be denoted by $V^{1.1}$. For each vertex $v \in V^{1.1}$, let $\xi(v) \stackrel{\text{def}}{=} j$ if $j$ is the smallest integer such that $|\Gamma_j(v)| > |\Gamma'(v)|/10$. Next, we define the set

$$W_v \stackrel{\text{def}}{=} \{u \in \Gamma_{\xi(v)}(v) : |\Gamma(u) \cap (\overline{\Gamma}_{\xi(v)}(v))| \geq |\overline{\Gamma}_{\xi(v)}(v)|/4\}, \tag{36}$$

and note that (by the case hypothesis) for every $u \in W_v$ it holds that $|\Gamma(u) \cap (\overline{\Gamma}_{\xi(v)}(v))| \geq |V_{\xi(v)}|/40$. By Observation 1, $|E(\Gamma_{\xi(v)}(v), \overline{\Gamma}_{\xi(v)}(v))| \geq |\Gamma_{\xi(v)}(v)| \cdot |\overline{\Gamma}_{\xi(v)}(v)|/2$. Noting that $|E(\Gamma_{\xi(v)}(v), \overline{\Gamma}_{\xi(v)}(v))| = \sum_{u \in \Gamma_{\xi(v)}(v)} |\Gamma(u) \cap (\overline{\Gamma}_{\xi(v)}(v))|$ and referring to the definition of $W_v$, it follows that $|W_v| \geq |\Gamma_{\xi(v)}(v)|/4 \geq |\Gamma'(v)|/40$. We complete the proof of Part 1 by noting that, for every $u \in W_v$, the set $\overline{\Gamma}_{\xi(v)}(v) \cap \Gamma(u)$ contains at least $|\overline{\Gamma}_{\xi(v)}(v)|/4 \geq |V_{\xi(v)}|/40$ elements, whereas each such element $w$ yields a witness $(v,u,w)$ (since $(u,v) \in E$ and $w \in \overline{\Gamma}_{\xi(v)}(v) \cap \Gamma(u)$).

Towards proving Part 2, we first omit from the foregoing $W_v$ all elements of $F(v)$; that is, we redefine $W_v$ as the set of all $u \in \Gamma_{\xi(v)}(v) \setminus F(v)$ such that $|\Gamma(u) \cap (\overline{\Gamma}_{\xi(v)}(v))| \geq |\overline{\Gamma}_{\xi(v)}(v)|/4$. Surely, this decreases $\sum_v |W_v|$ by at most $\sum_v |F(v)| = 2|F|$. Now, for every $u \in W_v$, let $W_{v,u}$ be a random subset of $|W_v|/40$ elements in $\overline{\Gamma}_{\xi(v)}(v) \cap \Gamma(u)$, while recalling that the latter set has size at least $|\overline{\Gamma}_{\xi(v)}(v)|/4 \geq |V_{\xi(v)}|/40$. Thus, Part 2a follows.

In order to establish Part 2b, we fix an arbitrary $j$, and let $V^{1.1_j} \stackrel{\text{def}}{=} \{v \in V^{1.1} : \xi(v) = j\}$. We first note that, for any $w \in V_j$, the expected size of $U_w^{(2)}$ is upper-bounded by

$$\sum_{v \in V^{1.1_j}} \sum_{u \in W_v} \frac{|W_v|/40}{|\overline{\Gamma}_j(v) \cap \Gamma(u)|} \leq \frac{1}{|V_j|} \cdot \sum_{v \in V^{1.1_j}} |W_v|^2 \tag{37}$$

where the inequality uses $|\overline{\Gamma}_j(v) \cap \Gamma(u)| \geq |V_j|/40$. As in the proof of Claim 4.4.1, it is possible to choose the subsets $W_{v,u}$ so that the sizes of the sets $U_w^{(2)}$ are not much larger than (the upper bounds on the value of) their expected sizes. It follows that if some $w \in V_j$ satisfies $|U_w^{(2)}| > \epsilon^{4/3}N^2$, then $\sum_{v\in V^{1.1_j}} |W_v|^2 > \epsilon^{4/3}N^2|V_j|/2$. Assume, contrary to the claim, that $|\Gamma'(v) \setminus F(v)| \leq \epsilon^{2/3}N/2$ (so that $|W_v| \leq \epsilon^{2/3}N/2$) for every $v$, but $|U_w^{(2)}| > \epsilon^{4/3}N^2$ for some $w \in V_j$ (so that $\sum_{v\in V^{1.1_j}} |W_v|^2 > \epsilon^{4/3}N^2|V_j|/2$). In such a case we have:

$$
\begin{aligned}
|E(V^{1.1_j}, V_j) \setminus F| &\geq \frac{1}{2}\sum_{v\in V^{1.1_j}} |W_v| &(38)\\
&\geq \frac{1}{2}\sum_{v\in V^{1.1_j}} \frac{|W_v|^2}{\epsilon^{2/3}N/2} &(39)\\
&> \frac{1}{2}|V_j|\cdot\epsilon^{2/3}N\,. &(40)
\end{aligned}
$$

It follows that there exists a vertex $u \in V_j$ such that $|\Gamma'(u) \setminus F(u)| \geq |(\Gamma(u) \setminus F(u)) \cap V^{1.1_j}| > \epsilon^{2/3}N/2$, and we have reached a contradiction. Thus, Part 2 follows in this case.

***Case 1.2:*** $|\overline{\Gamma}_j(v)| \leq |V_j|/10$ (i.e., $|\Gamma_j(v)| \geq 0.9|V_j|$)**.** Let $i$ be such that $v \in V_i$. We first note that $|\Gamma_i(v)| \geq 0.8|\Gamma_j(v)|$, because otherwise we would obtain a better partition by moving the vertex $v$ from $V_i$ to $V_j$ (since the gain from such a move is at least $(|\Gamma_j(v)| - |\overline{\Gamma}_j(v)|) - |\Gamma_i(v)|$, whereas $|\Gamma_j(v)| - |\overline{\Gamma}_j(v)| \geq 0.8|V_j| \geq 0.8|\Gamma_j(v)|$). It follows that $|\Gamma_i(v)| \geq 0.8 \cdot |\Gamma'(v)|/10 > |\Gamma'(v)|/13$. We consider two subcases regarding the cardinality of the set $\Gamma_i(v)$:

1. If $|\Gamma_i(v)| \geq 0.9 \cdot |V_i|$, then we let $W_v$ be a subset of $\Gamma_j(v)$, and for each $u \in W_v$, we let $W_{v,u}$ be a subset of $\Gamma_i(v) \setminus \Gamma(u)$. Thus each triple $(v, u, w)$ where $u \in W_v$ and $w \in W_{v,u}$ is a witness. For an illustration, see Figure 5. Combining the case hypotheses (which asserts that $V_j \times V_i$ is essentially covered by $\Gamma_j(v) \times \Gamma_i(v)$) with Observation 2 (which guarantees many non-edges in $V_j \times V_i$), we obtain $\Omega(|\Gamma'(v)|^2)$ such witnesses. Details follow.

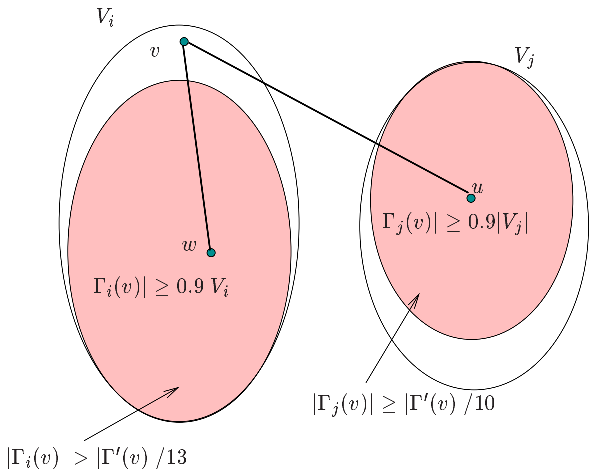


Figure 5: An Illustration for the 1st subcase of Case 1.2 in the proof of Claim 4.4.2.

Let us denote the subset of vertices (in $V_i$) for which the case hypothesis holds by $V_i^{1.2}$, and for each $v \in V_i^{1.2}$ define $\xi(v)$ as in Case 1.1. Fixing any $i$ and $v \in V_i^{1.2}$, let

$$W_v \stackrel{\text{def}}{=} \{u \in \Gamma_{\xi(v)}(v) : |\Gamma_i(v) \setminus \Gamma(u)| \geq |\Gamma_i(v)|/10\}\,. \tag{41}$$

Note that for any $u \in W_v$ it holds that $|\Gamma_i(v) \setminus \Gamma(u)| \geq 0.1|\Gamma_i(v)| \geq 0.08|\Gamma_j(v)|$, where $j \stackrel{\text{def}}{=} \xi(v)$. Using Observation 2 we have that

$$\begin{aligned}
|E(\Gamma_j(v), \Gamma_i(v))| &\leq |E(V_j, V_i)| && (42)\\
&\leq \frac{1}{2} \cdot |V_j| \cdot |V_i| && (43)\\
&\leq \frac{1}{2} \cdot \frac{|\Gamma_j(v)|}{0.9} \cdot \frac{|\Gamma_i(v)|}{0.9} && (44)\\
&< 0.7 \cdot |\Gamma_j(v)| \cdot |\Gamma_i(v)| \,. && (45)
\end{aligned}$$

Hence there are at least $0.3 \cdot |\Gamma_j(v)| \cdot |\Gamma_i(v)|$ pairs $(u, w) \in \Gamma_j(v) \times \Gamma_i(v)$ such that $(u, w) \notin E$; that is, $\sum_{u \in \Gamma_j(v)} |\Gamma_i(v) \setminus \Gamma(u)| > 0.3 \cdot |\Gamma_j(v)| \cdot |\Gamma_i(v)|$. It follows that $|W_v| > |\Gamma_j(v)|/5$, where by the hypothesis of Case 1 this value is greater than $|\Gamma'(v)|/50$. Next, recalling that for any $u \in W_v$ it holds that $|\Gamma_i(v) \setminus \Gamma(u)| \geq 0.08|\Gamma_j(v)|$, we let $W_{v,u}$ be an arbitrary $0.08|W_v|$-size subset of $\Gamma_i(v) \setminus \Gamma(u) \subseteq \Gamma(v) \setminus \Gamma(u)$, and note that indeed for every $u \in W_v$ and $w \in W_{v,u}$ it holds that $u, w \in \Gamma(v)$ and $(u, w) \notin E$. Thus, Part 1 follows in this case.

As for Part 2, we first omit from $W_v$ all vertices in $F(v)$ (i.e., we redefine $W_v$ as the set of all $u \in \Gamma_{\xi(v)}(v) \setminus F(v)$ satisfying $|\Gamma_i(v) \setminus \Gamma(u)| \geq |\Gamma_i(v)|/10$), and let each $W_{v,u}$ be a random $0.08|W_v|$-size subset of $\Gamma_i(v) \setminus \Gamma(u) \subseteq \Gamma(v) \setminus \Gamma(u)$. This establishes Part 2a, and so we turn to Part 2b. We then note that, for every $w \in V_i$, the expected size of $U_w^{(2)}$ (when restricted to pairs $(v, u)$ with $v \in V_i^{1.2}$) is upper-bounded by

$$\sum_{v \in V_i} \sum_{u \in W_v} \frac{0.08|W_v|}{|\Gamma_i(v) \setminus \Gamma(u)|} \leq \frac{0.08}{0.09|V_i|} \cdot \sum_{v \in V_i} |W_v|^2 \qquad (46)$$

where the inequality uses $|\Gamma_i(v) \setminus \Gamma(u)| \geq 0.1|\Gamma_i(v)| \geq 0.09|V_i|$. Again, we may select the sets $W_{v,u}$ such that for each $w \in V_i$ it holds that $|U_w^{(2)}| < \sum_{v \in V_i} |W_v|^2/|V_i|$. Thus, if some $w \in V_i$ satisfies $|U_w^{(2)}| > \epsilon^{4/3} N^2$, then $\sum_{v \in V_i} |W_v|^2 > \epsilon^{4/3} N^2 |V_i|$. It follows that there exists a vertex $v \in V_i$ such that $|W_v| > \epsilon^{2/3} N$, and Part 2 follows.

2. If $|\Gamma_i(v)| \leq 0.9 \cdot |V_i|$, then we proceed somewhat differently than in the other cases (this is the exceptional case mentioned at the preamble of the proof). Recall that $\overline{\Gamma}(v) = \overline{\Gamma}_i(v) = V_i \setminus \Gamma(v)$, and so $|\overline{\Gamma}(v)| \geq 0.1 \cdot |V_i| \geq 0.008 \cdot |\Gamma'(v)|$ (because $|V_i| \geq |\Gamma_i(v)| \geq 0.8|\Gamma_j(v)|$ and $|\Gamma_j(v)| \geq |\Gamma'(v)|/10$). For the basic claim, we invoke Claim 4.4.1, translating the lower bound in terms of $|\overline{\Gamma}(v)|$ (provided by Claim 4.4.1) into a lower bound in terms of $|\Gamma'(v)|$. For the elaborate claim, we set $W_v = \emptyset$ for every $v$ as in the case hypothesis (i.e., the current Case 1.2.2). Thus, we trivially have that $|W_{v,u}| \geq |W_v|/c_2$ for every $u \in W_v$, and Part 2 of the claim holds trivially as well. Finally, we use the hypothesis of Eq. (33) (i.e., $\sum_{v \in [N]} |\Gamma'(v)| > (125/\alpha) \sum_{v \in [N]} |\overline{\Gamma}(v)|$) to infer that the current subcase (in which $|\Gamma'(v)| \leq 125|\overline{\Gamma}(v)|$) may account for less than an $\alpha$ fraction of the sum $\sum_{v \in [N]} |\Gamma'(v)|$. All other vertices $v$ will be placed in $V'$, and hence Eq. (34) holds.

This completes the treatment of the current case (i.e., Case 1.2), which in turn completes the treatment of Case 1. (We thus proceed to the following complementary Case 2.)

Case 2: No single $\boldsymbol{V_j}$ contains much of $\boldsymbol{\Gamma'(v)}$; that is, for every $j$ it holds that $|\Gamma_j(v)| \leq |\Gamma'(v)|/10$. As in Case 1, we consider two subcases regarding the relative part of each $V_j$

covered by $\Gamma'(v)$, but in the current case we consider a partition of the set $J \stackrel{\rm def}{=} \{j : |\Gamma_j(v)| \geq 1\}$ and distinguish cases regarding the intersection of $\Gamma'(v)$ with the sets $V_j$ in each part.[8] Specifically, we let $J' \stackrel{\rm def}{=} \{j : |\Gamma_j(v)| > 0.9|V_j|\}$, where each $V_j$ with $j \in J'$ is analogous to Case 1.2, except that having several such $V_j$ calls for seeking witnesses of the form $(v,u,w)$ such that $(u,w) \in (\Gamma(v) \times \Gamma(v)) \setminus E$. The case that $\sum_{j\in J'} |\Gamma_j(v)|$ accounts for much of $|\Gamma'(v)|$ is treated first (in Case 2.1), and the complementary case is postponed to Case 2.2.

***Case 2.1:*** $\boldsymbol{\sum_{j\in J'} |\Gamma_j(v)| \geq 0.5 \cdot |\Gamma'(v)|}$**.** In this case $J'$ has cardinality at least five (since $\sum_{j\in J'} |\Gamma_j(v)| \geq 0.5 \cdot |\Gamma'(v)|$ and $|\Gamma_j(v)| \leq 0.1 \cdot |\Gamma'(v)|$ for every $j$). Let $C_v = \bigcup_{j\in J'} \Gamma_j(v)$ (note that the vertices in $C_v$ belong to several cliques $V_j$). In this case we let $W_v$ be a subset of $C_v$, and for each $u \in C_v$ we let $W_{v,u}$ be a subset of $C_v \setminus \Gamma(u)$. We shall show that the case hypothesis implies that there are many missing edges between pairs of vertices in $C_v$. Intuitively, this holds because $C_v$ essentially covers $\bigcup_{j\in J'} V_j$, whereas (by Observation 2) for any $j_1 \neq j_2$ there are many non-edges in $V_{j_1} \times V_{j_2}$. This ensures that we have many witnesses of the form $(v,u,w)$, where $u \in W_v$ and $w \in W_{v,u}$. Details follow.

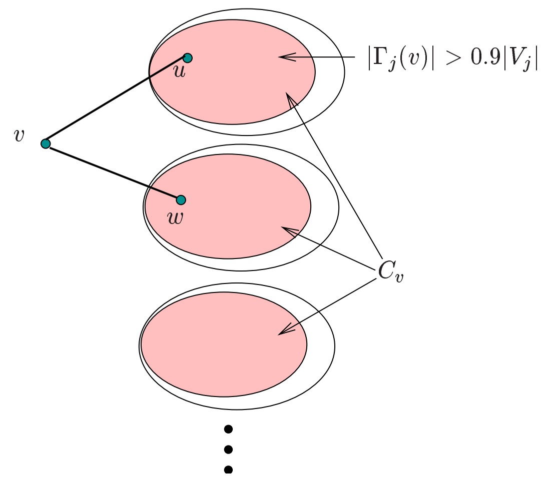


Figure 6: An Illustration for Case 2.1 in the proof of Claim 4.4.2.

For every $j_1 \neq j_2 \in J'$, by Observation 2 (and since $|\Gamma_j(v)| > 0.9|V_j|$ for every $j \in J'$), it holds that

$$|E(\Gamma_{j_1}(v), \Gamma_{j_2}(v))| \leq \frac{1}{2} \cdot |V_{j_1}| \cdot |V_{j_2}| < 0.7 \cdot |\Gamma_{j_1}(v)| \cdot |\Gamma_{j_2}(v)| \tag{47}$$

(cf. the derivation of Eq. (45) from Eq. (42)).

Letting $M \stackrel{\rm def}{=} \sum_{j_1\neq j_2\in J'} |(\Gamma_{j_1}(v) \times \Gamma_{j_2}(v)) \setminus E|$, we first observe that

$$M = \sum_{j_1\neq j_2\in J'} (|\Gamma_{j_1}(v)| \cdot |\Gamma_{j_2}(v)| - |E(\Gamma_{j_1}(v), \Gamma_{j_2}(v))|) \tag{48}$$

[8]We note that the threshold for relative density is also different in the current case.

$$\geq \sum_{j_1 \neq j_2 \in J'} (1-0.7) \cdot |\Gamma_{j_1}(v)| \cdot |\Gamma_{j_2}(v)| \tag{49}$$

$$= 0.3 \cdot \left( \left( \sum_{j \in J'} |\Gamma_j(v)| \right)^2 - \sum_{j \in J'} |\Gamma_j(v)|^2 \right) \tag{50}$$

$$\geq 0.3 \cdot \left( (0.5 \cdot |\Gamma'(v)|)^2 - 0.1 \cdot |\Gamma'(v)|^2 \right), \tag{51}$$

where the last inequality uses the hypotheses of Cases 2 and 2.1. Therefore, $|(C_v \times C_v) \setminus E| \geq M > 0.04 \cdot |\Gamma'(v)|^2$.

Defining

$$W_v \stackrel{\rm def}{=} \{u \in C_v : |C_v \setminus \Gamma(u)| \geq 0.02 \cdot |\Gamma'(v)|\}, \tag{52}$$

we note that $|W_v| \geq 0.02 \cdot |\Gamma'(v)|$. Next, we let $W_{v,u}$ be any $0.02 \cdot |W_v|$-size subset of $C_v \setminus \Gamma(u) \subseteq \Gamma'(v) \setminus \Gamma(u)$. As in the previous cases, Part 1 follows by the definition of these sets.

Establishing Part 2 (or rather Part 2b) is slightly more complicated in the current case, and so we first make the simplifying assumption that $|F(v)| < 0.01|W_v|$, for every vertex $v$. This simplifying assumption implies that $|F(v)| < 0.01|\Gamma'(v)|$, which means that for every $u \in W_v$ it holds that $|(C_v \setminus \Gamma(u)) \setminus F(v)| > 0.01|\Gamma'(v)|$. Now, we omit from $W_v$ all vertices in $F(v)$ (i.e., redefine $W_v$ as the set of $u \in C_v \setminus F(v)$ such that $|C_v \setminus \Gamma(u)| \geq 0.02 \cdot |\Gamma'(v)|$), and let each $W_{v,u}$ be a random $0.01|W_v|$-size subset of $(C_v \setminus \Gamma(u)) \setminus F(v)$. Part 2a follows, and so we turn to establishing Part 2b. Again, for any fixed $w$, the expected size of $U_w^{(2)}$ is upper-bounded by

$$\sum_{v \in [N]: C_v \ni w} \sum_{u \in W_v} \frac{0.01 \cdot |W_v|}{|(C_v \setminus \Gamma(u)) \setminus F(v)|} \leq \sum_{v \in [N]: (\Gamma'(v) \setminus F(v)) \ni w} \sum_{u \in W_v} \frac{0.01 \cdot |C_v|}{0.01 \cdot |C_v|} \tag{53}$$

$$= \sum_{v \in \Gamma'(w) \setminus F(w)} |W_v| \tag{54}$$

where the inequality uses $|(C_v \setminus \Gamma(u)) \setminus F(v)| \geq 0.01 \cdot |\Gamma'(v)|$ and $W_v \subseteq C_v \subseteq \Gamma'(v)$. We conclude that the existence of $w \in V_j$ such that $|U_w^{(2)}| > \epsilon^{4/3} N^2$ implies that $\sum_{v \in \Gamma'(w) \setminus F(w)} |W_v| > \epsilon^{4/3} N^2/2$, which in turn implies that either $|\Gamma'(w) \setminus F(w)| > \epsilon^{2/3} N/2$ or $|W_v| > \epsilon^{2/3} N$ for some $v \in \Gamma'(w) \setminus F(w)$. Thus, Part 2 follows (under the assumption that $|F(v)| < 0.01|W_v|$).

It remains to handle the case in which for some $v$ it holds that $|F(v)| \geq 0.01|W_v|$. In this case we just reset $W_v$ to the empty set, and the foregoing analysis still applies (establishing Part 2b). We need, however, to examine the effect of this modification on Part 2a. The key observation is that the sum of the sizes of the $W_v$'s decreases at most by $200|F|$, because the case of $|F(v)| \geq 0.01|W_v|$ (where $W_v$ is reset to empty) causes a loss of at most $|W_v| < 100|F(v)|$, whereas the case of $|F(v)| < 0.01|W_v|$ (in which we avoid $F(v)$) causes (as usual) a loss of at most $|F(v)|$. Thus, Part 2 holds in Case 2.1.

***Case 2.2:* $\sum_{j \in J \setminus J'} |\Gamma_j(v)| \geq 0.5 \cdot |\Gamma'(v)|$.** Let $J'' \stackrel{\rm def}{=} J \setminus J' = \{j : 1 \leq |\Gamma_j(v)| \leq 0.9|V_j|\}$, and note that for $j \in J''$ (as considered in this case) it may be that $|\Gamma_j(v)| \ll |V_j|$ and consequently for $j_1 \neq j_2 \in J''$ it may hold that $E(\Gamma_{j_1}(v), \Gamma_{j_2}(v)) \approx |\Gamma_{j_1}(v)| \cdot |\Gamma_{j_2}(v)|$. More generally, redefining $C_v \stackrel{\rm def}{=} \bigcup_{j \in J''} \Gamma_j(v)$, it may be that $|E(C_v, C_v)| \approx \binom{|C_v|}{2}$, and so the approach of Case 2.1 may not work in general (although it will work in the first subcase). Letting $J''' \stackrel{\rm def}{=} \{j \in J'' : |V_j| \leq |\Gamma'(v)|/10\}$, we consider two subcases:

1. If $\sum_{j\in J'''} |\Gamma_j(v)| \geq 0.4 \cdot |\Gamma'(v)|$, then we redefine $C_v \stackrel{\rm def}{=} \bigcup_{j\in J'''} \Gamma_j(v)$ and show that $|E(C_v, C_v)| \leq 0.9\binom{|C_v|}{2}$. Once the latter fact is established, we reach a situation as in Case 2.1 (where we only used $\binom{|C_v|}{2} - |E(C_v, C_v)| > 0.04|\Gamma'(v)|^2$) and proceed essentially as in that case. (The only modification is that here we only have $\binom{|C_v|}{2} - |E(C_v, C_v)| > 0.002|\Gamma'(v)|^2$, and so we let $W_v$ consists of all vertices $u \in C_v$ such that $|C_v \setminus \Gamma(u)| \geq 0.001|\Gamma'(v)|$, so that $|W_v| \geq 0.001 \cdot |\Gamma'(v)|$, and we let $W_{v,u}$ be a $0.001 \cdot |W_v|$-size random subset of $C_v \setminus \Gamma(u)$.) Thus, we focus on establishing that $|E(C_v, C_v)| \leq 0.9\binom{|C_v|}{2}$, by showing that otherwise one obtains a contradiction to the optimality of the partition (by replacing the sub-partition $(V_j)_{j\in J'''}$ with $(C_v, (V_j \setminus C_v)_{j\in J'''})$, where $V_j \setminus C_v = \overline{\Gamma}_j(v)$). Details follow.
   Assuming, towards the contradiction, that $|E(C_v, C_v)| > 0.9\binom{|C_v|}{2}$, we lower-bound the gain from the aforementioned replacement as follows. Combining all $C_v \cap V_j$'s (into $C_v$) and splitting each $V_j$ (to $(C_v \cap V_j, V_j \setminus C_v)$), yields a gain of at least
$$\begin{aligned}\Delta \stackrel{\rm def}{=} & \sum_{j_1<j_2\in J'''} |E(C_v \cap V_{j_1}, C_v \cap V_{j_2})| - \sum_{j_1<j_2\in J'''} |\overline{E}(C_v \cap V_{j_1}, C_v \cap V_{j_2})| \\ & - \sum_{j\in J'''} |E(C_v \cap V_j, V_j \setminus C_v)| \end{aligned} \tag{55}$$
   where $\overline{E}(Y, Z)$ denotes the set of pairs of vertices in $Y \times Z$ that do not have an edge between them. Thus:
$$\begin{aligned}\Delta &\geq |E(C_v, C_v)| - \sum_{j\in J'''} |E(C_v \cap V_j, C_v \cap V_j)| \\ &\quad -|\overline{E}(C_v, C_v)| - \sum_{j\in J'''} |E(C_v \cap V_j, V_j \setminus C_v)| \qquad (56)\\ &= |E(C_v, C_v)| - |\overline{E}(C_v, C_v)| - \sum_{j\in J'''} |E(C_v \cap V_j, V_j)| \qquad (57)\\ &\geq |E(C_v, C_v)| - |\overline{E}(C_v, C_v)| - |C_v| \cdot \max_{j\in J'''}\{V_j\}. \qquad (58)\end{aligned}$$
   By the contradiction hypothesis $|E(C_v, C_v)| > 0.9\binom{|C_v|}{2}$ (and $|\overline{E}(C_v, C_v)| < 0.1\binom{|C_v|}{2}$), whereas $\max_{j\in J'''}\{|V_j|\} \leq |\Gamma'(v)|/10$ and $|\Gamma'(v)| \leq 2.5|C_v|$ (by the definition of $J'''$ and the subcase hypothesis, respectively). Hence, $\Delta > 0.8\binom{|C_v|}{2} - 0.25|C_v|^2 > 0$, in contradiction to the optimality of the partition.
2. If $\sum_{j\in J''\setminus J'''} |\Gamma_j(v)| \geq 0.1 \cdot |\Gamma'(v)|$, then we proceed similarly to Case 1.1. That is, we try to obtain witnesses of the form $(v, u, w)$ such that $(u, w) \in \bigcup_{j\in J''\setminus J'''}(\Gamma_j(v) \times \overline{\Gamma}_j(v)) \cap E$; see Figure 7. Indeed, the only difference between Case 1.1 and the current subcase is that here $j \in J'' \setminus J'''$ may not be unique, but as we shall see this issue has little consequences. Specifically, we define
$$W_v \stackrel{\rm def}{=} \bigcup_{j\in J''\setminus J'''} \left\{u \in \Gamma_j(v) : |\Gamma_j(u) \cap \overline{\Gamma}_j(v)| \geq \frac{|\overline{\Gamma}_j(v)|}{4}\right\} \tag{59}$$

and note that $W_v \subseteq \Gamma'(v)$ and that for every $j \in J'' \setminus J'''$ it holds that $|W_v \cap V_j| \geq |\Gamma_j(v)|/4$ (since $E(\Gamma_j(v), V_j \setminus \Gamma_j(v)) \geq |\Gamma_j(v)| \cdot |V_j \setminus \Gamma_j(v)|/2$). Using the subcase hypothesis, it follows that $|W_v| \geq \sum_{j \in J'' \setminus J'''} |\Gamma_j(v)|/4 \geq |\Gamma'(v)|/40$, and using $j \in J'' \setminus J'''$ every $u \in W_v$ satisfies $|\Gamma_j(u) \cap \overline{\Gamma}_j(v)| \geq |\overline{\Gamma}_j(v)|/4 \geq |V_j|/40 \geq |\Gamma'(v)|/400$. Next, for every $j \in J'' \setminus J'''$ and every $u \in W_v \cap V_j$, we define $W_{v,u}$ to be a random subset of size $|\Gamma'(v)|/400$ of $\Gamma_j(u) \cap \overline{\Gamma}_j(v)$. Indeed, for every $u \in W_v$ and $w \in W_{v,u}$ it holds that $w \notin \Gamma'(v)$ and $w \in \Gamma(u) \setminus \Gamma'(u)$. Given the lower bounds on the sizes of the sets $W_v$ and $W_{v,u}$, Part 1 follows.

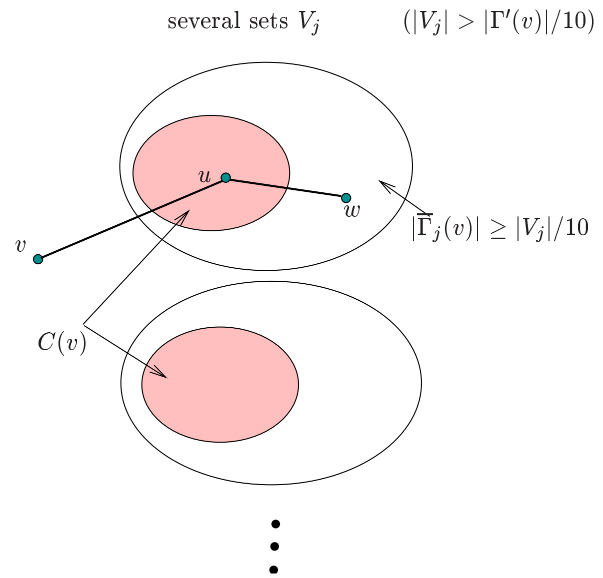


Figure 7: An Illustration for the 2nd subcase of Case 2.2 in the proof of Claim 4.4.2.

Again, proving Part 2 amounts to omitting from the foregoing $W_v$ all elements of $F(v)$; that is, we redefine $W_v$ as the set of all $u \in \bigcup_{j \in J'' \setminus J'''} \Gamma_j(v) \setminus F(v)$ such that $|\Gamma_j(u) \cap (\overline{\Gamma}_j(v))| \geq |\overline{\Gamma}_j(v)|/4\}$. Similarly, the sets $W_{v,u}$ are random subsets of size $|\Gamma'(v)|/400$ of $\Gamma_j(u) \cap \overline{\Gamma}_j(v)$. Thus, Part 2a follows.
To establish Part 2b, we note that, for any fixed $w \in V_j$, the expected size of $U_w^{(2)}$ is upper-bounded by

$$\sum_{v \in [N] \setminus V_j} \sum_{u \in W_v \cap V_j} \frac{|\Gamma'(v) \setminus F(v)|/400}{|\Gamma_j(u) \cap \overline{\Gamma}_j(v)|} \leq \sum_{v \in [N] \setminus V_j} \sum_{u \in \Gamma_j(v) \setminus F(v)} \frac{|\Gamma'(v) \setminus F(v)|}{10|V_j|} \tag{60}$$
$$= \sum_{v \in [N] \setminus V_j} \frac{|\Gamma_j(v) \setminus F(v)| \cdot |\Gamma'(v) \setminus F(v)|}{10|V_j|}$$

where the inequality uses $|\Gamma_j(u) \setminus \Gamma_j(v)| \geq |V_j \setminus \Gamma_j(v)|/4 \geq |V_j|/40$. Here too it is possible to choose the subsets $W_{v,u}$ so that the sizes of the sets $U_w^{(2)}$ are not much larger than (the upper bounds on the value of) their expected sizes. Again, it follows that if some $w \in V_j$ satisfies $|U_w^{(2)}| > \epsilon^{4/3} N^2$, then

$$\sum_{v \in [N] \setminus V_j} |\Gamma'(v) \setminus F(v)| \cdot |\Gamma_j(v) \setminus F(v)| > 5\epsilon^{4/3} N^2 |V_j|, \tag{61}$$

which implies that either for some $v \in [N] \setminus V_j$ it holds that $|\Gamma'(v) \setminus F(v)| > \epsilon^{2/3} N$ or that $\sum_{v \in [N] \setminus V_j} |\Gamma_j(v) \setminus F(v)| > \epsilon^{2/3} N |V_j|$. In the latter case, there must be a vertex $u \in V_j$ such that $|\Gamma'(u) \setminus F(v)| > \epsilon^{2/3} N$. Thus, Part 2b holds also in this subcase of Case 2.2.

Thus, we have established the claim for all subcases of Case 2.2.

Having completed the treatment of the two complementary cases of Case 2 (i.e., Cases 2.1 and 2.2), we complete the treatment of Case 2.

Having completed the treatment of the two complementary cases (i.e., Cases 1 and 2), the claim follows. ■

### 4.2.2 The existence of effective witnesses

Combining the hypothesis of Lemma 4.4 with (the basic parts of) Claims 4.4.1 and 4.4.2, we infer the existence of $\Omega(\epsilon^2 N^3)$ witnesses. Moreover, the elaborate parts of these claims provide us with some structure that will be useful towards proving that (with high probability) the sample taken by Algorithm 4.3 contains at least one effective witness (i.e., a witness whose three vertex-pairs are inspected by the algorithm). We shall use the following technical claim, which will be proved in Section 4.2.3. Essentially, the claim asserts that under some circumstances (i.e., those detailed in the conditions), a random set of adequate size (i.e., of size $O(\epsilon^{-2/3})$) contains a witness. Loosely speaking, the first condition means that the expected number of witnesses in the sample exceeds any desired constant, whereas the upper bounds on the sizes of the sets $W_v, W_{v,u}, U_v^{(1)}$ and $U_v^{(2)}$ (stated in the other conditions) guarantee sufficient concentration around the expected value.

**Claim 4.4.3** (on the existence of witnesses in a sample of vertices): *Suppose that the following conditions hold:*

1. $\sum_{v \in [N]} \sum_{u \in W_v} |W_{v,u}| = \Omega(\epsilon^2 \cdot N^3)$
2. *For every $v \in [N]$, it holds that $|W_v| < \epsilon^{2/3} N$ and for $U_v^{(1)} \stackrel{\text{def}}{=} \{x : v \in W_x\}$ and $U_v^{(2)} \stackrel{\text{def}}{=} \{(x,y) : v \in W_{x,y}\}$, it holds that $|U_v^{(1)}| < \epsilon^{2/3} N$ and $|U_v^{(2)}| < \epsilon^{4/3} N^2$.*
3. *For every $v \in [N]$ and $u \in W_v$, it holds that $|W_{v,u}| < \epsilon^{2/3} N$.*

*Then, for a sufficiently large constant $c$ that depends only on the constant in the Omega-notation, with probability at least $2/3$, a uniformly selected sample of $c \cdot \epsilon^{-2/3}$ vertices contains a triple $(v, u, w)$ such that $u \in W_v$ and $w \in W_{v,u}$.*

The proof of Claim 4.4.3 appears in Section 4.2.3. Using Claims 4.4.1, 4.4.2 and 4.4.3, we finally prove Lemma 4.4.

**Completing the proof of Lemma 4.4.** Recall that $\sum_{v \in [N]} (|\overline{\Gamma}(v)| + |\Gamma'(v)|) \geq \epsilon \cdot N^2$ (by the lemma's hypothesis). Thus, for any constant $\beta > 0$, either $\sum_{v \in [N]} |\overline{\Gamma}(v)| \geq \beta \cdot \epsilon \cdot N^2$ or $\sum_{v \in [N]} |\Gamma'(v)| \geq (1 - \beta) \cdot \epsilon \cdot N^2$. We analyze these two cases, while postponing the determination of the constant $\beta \in (0, 1)$ to the treatment of the second case.

Case of $\sum_{v \in [N]} |\overline{\boldsymbol{\Gamma}}(\boldsymbol{v})| \geq \boldsymbol{\beta} \cdot \boldsymbol{\epsilon} \cdot \boldsymbol{N}^2$. We consider two subcases (and use claim 4.4.3 only in the second one):

1. The easier subcase is when large sets $\overline{\Gamma}(\cdot)$ have a relatively large contribution to $\sum_{v\in[N]} |\overline{\Gamma}(v)|$. In this case, we apply Claim 4.4.1 with $F = \emptyset$ and obtain all that we need by using Part 1 of this claim, while observing that Algorithm 4.3 inspects all vertex pairs that arise from this analysis. (We note that this case refers to triples in $\bigcup_{k<(2\ell/3)+O(1)} R_k \times S_k \times S_k$, where the bound on $k$ is related to the bound on large sets.)

   Specifically, if $\sum_{v\in[N]:|\overline{\Gamma}(v)|\geq\epsilon^{2/3}N/2} |\overline{\Gamma}(v)| \geq (\beta/10)\cdot\epsilon\cdot N^2$, then applying Claim 4.4.1 with $F = \emptyset$ we obtain sets $W_v$'s and $W_{v,u}$'s such that Part 1 of Claim 4.4.1 holds. In particular, it follows that

$$\sum_{v\in[N]:|W_v|\geq\epsilon^{2/3}N/8} |W_v| \geq \sum_{v\in[N]:|\overline{\Gamma}(v)|\geq\epsilon^{2/3}N/2} \frac{|\overline{\Gamma}(v)|}{4} \quad (62)$$

$$\geq \frac{(\beta/10)\cdot\epsilon\cdot N^2}{4} = \Omega(\epsilon\cdot N^2). \quad (63)$$

   Recall that $\ell = \log_2(1/\epsilon)$. Thus, there exists $k \in \{1,\ldots,(2\ell/3)+3\}$ such that for $V^* \stackrel{\text{def}}{=} \{v \in [N] : 2^{-k}N \leq |W_v| < 2^{-k+1}N\}$ it holds that $\sum_{v\in V^*} |W_v| = \Omega(\epsilon\cdot N^2/\ell)$. Fixing this $k$, we note that $|V^*| = \Omega(2^k\epsilon\cdot N/\ell)$ and thus $\Pr[R_k \cap V^* \neq \emptyset] > 8/9$, where $R_k$ is as selected in Step 2 of Algorithm 4.3 (i.e., $R_k$ is a random set of size $\Omega((2^k\epsilon/\ell)^{-1})$). For the sake of the analysis, view $S_k$ (which is a uniformly selected subset of $S$ that has size $\Theta(2^k)$ and is also selected in Step 2) as the union of two independently selected subsets of equal size, denoted $S_k^1$ and $S_k^2$. Fixing any $v \in R_k \cap V^*$, we have $|W_v| \geq 2^{-k}N$ and so $\Pr[S_k^1 \cap W_v \neq \emptyset] > 8/9$. Finally, fixing any $u \in S_k^1 \cap W_v$, since $|W_{v,u}| = \Omega(|W_v|) = \Omega(2^{-k}N)$, we have $\Pr[S_k^2 \cap W_{v,u} \neq \emptyset] > 8/9$. Noting that all pairs $(R_k \times S_k) \cup (S_k \times S_k)$ are inspected by Algorithm 4.3, the claim follows (i.e., with probability at least 2/3, the sample taken by Algorithm 4.3 contains a witness).

2. The other subcase is when large sets $\overline{\Gamma}(\cdot)$ have a relatively small contribution to $\sum_{v\in[N]} |\overline{\Gamma}(v)|$. In this case, we apply Claim 4.4.1 while setting $F$ so to eliminate all large sets. Here we use both parts of the claim, where Part 2 provides the conditions required by the non-trivial probabilistic analysis captured in Claim 4.4.3. (We note that this case refers to triples in $S \times S \times S$.)

   Specifically, if $\sum_{v\in[N]:|\overline{\Gamma}(v)|\geq\epsilon^{2/3}N/2} |\overline{\Gamma}(v)| < (\beta/10)\cdot\epsilon\cdot N^2$, then we set $F = \{\{u,v\} : u \in \overline{\Gamma}(v),\ |\overline{\Gamma}(v)| \geq \epsilon^{2/3}N/2\}$, which means that $F(v) = \overline{\Gamma}(v)$ if $|\overline{\Gamma}(v)| \geq \epsilon^{2/3}N/2$ and $F(v) = \emptyset$ otherwise. Applying Claim 4.4.1 with this $F$, and noting that $|F| = \sum_{v\in[N]:|\overline{\Gamma}(v)|\geq\epsilon^{2/3}N/2} |\overline{\Gamma}(v)|$, we obtain sets $W_v$'s and $W_{v,u}$'s such that Claim 4.4.1 holds. In particular (by Part 1), we have that

$$\sum_{v\in[N]} |W_v| \geq \sum_{v\in[N]} \frac{|\overline{\Gamma}(v)|}{4} - 2|F| \quad (64)$$

$$\geq \left(\frac{\beta}{4} - 2\cdot\frac{\beta}{10}\right)\cdot\epsilon\cdot N^2 = \Omega(\epsilon\cdot N^2), \quad (65)$$

   whereas $|W_v| \leq |\overline{\Gamma}(v) \setminus F(v)| < \epsilon^{2/3}N/2$ holds for every $v \in [N]$. Recall that $|W_{v,u}| \leq |W_v|$ holds for every $u \in W_v$.

   Letting $U_w^{(1)} \stackrel{\text{def}}{=} \{v : w \in W_v\}$, for every $w$ it holds that $|U_w^{(1)}| < \epsilon^{2/3}N/2$ (because $v \in U_w^{(1)}$ implies $w \in \overline{\Gamma}(v)$ and $(v,w) \notin F$). Also, by Part 2, we get $|U_w^{(2)}| < \epsilon^{4/3}N^2$ for every $w$. Thus, all conditions of Claim 4.4.3 hold, and we conclude that (in this case), with high probability, the sample $S$ selected in Step 1 (of Algorithm 4.3) contains a witness (i.e., a triple $(v,u,w)$ such that $u \in W_v$ and $w \in W_{v,u}$).

This completes the treatement of the case in which $\sum_{v\in[N]} |\overline{\Gamma}(v)| \geq \beta \cdot \epsilon \cdot N^2$. The treatment of the case in which $\sum_{v\in[N]} |\Gamma'(v)| \geq (1-\beta)\cdot\epsilon\cdot N^2$ is analogous. Specifically, we consider analogous subcases (with different constants in the differentiating thresholds), and invoke Claim 4.4.2 (while setting $\alpha, \beta > 0$ to be sufficiently small such that all calculations work out).

Case of $\sum_{v\in[N]} |\boldsymbol{\Gamma'(v)}| \geq \boldsymbol{(1-\beta)\cdot\epsilon\cdot N^2}$. We may also assume that $\sum_{v\in[N]} |\overline{\Gamma}(v)| < \beta\cdot\epsilon\cdot N^2$, since otherwise the previous case applies. Thus, $\sum_{v\in[N]} |\Gamma'(v)| > \frac{1-\beta}{\beta}\cdot\sum_{v\in[N]} |\overline{\Gamma}(v)|$, which for $\beta \leq 1/2$ is at least $\frac{1}{2\beta}\cdot\sum_{v\in[N]} |\overline{\Gamma}(v)|$. Therefore, the premise of the (elaborate part of) Claim 4.4.2 holds with $\alpha = 250\beta$ and hence the conclusions of the claim hold as well. We consider two subcases, which are determined by a parameter $\gamma$ that will be set in the course of the analysis. In what follows, recall that $c_1, c_2, c_3$ and $c_4$ are constants that are defined by Claim 4.4.2.

1. If $\sum_{v\in[N]:|\Gamma'(v)|\geq\epsilon^{2/3}N/2} |\Gamma'(v)| \geq \gamma\cdot\sum_{v\in[N]} |\Gamma'(v)|$, then, by Part 1 of Claim 4.4.2, for every $v \in [N]$ we obtain a set $W_v \subseteq \Gamma'(v)$
   such that:
$$\sum_{v\in[N]:|\Gamma'(v)|\geq\epsilon^{2/3}N/2\,\&\,|W_v|\geq|\Gamma'(v)|/c_1} |\Gamma'(v)| \geq (\gamma-\alpha)\cdot(1-\beta)\cdot\epsilon\cdot N^2 \tag{66}$$
$$= (\gamma-250\beta)\cdot(1-\beta)\cdot\epsilon\cdot N^2\ , \tag{67}$$
   If $\beta$ and $\gamma$ are set so that $\gamma \geq 251\beta$, then,
$$\sum_{v\in[N]:|W_v|\geq\epsilon^{2/3}N/(2c_1)} |W_v| \geq \sum_{v\in[N]:|\Gamma'(v)|\geq\epsilon^{2/3}N/2\,\&\,|W_v|\geq|\Gamma'(v)|/c_1} \frac{|\Gamma'(v)|}{c_1} \tag{68}$$
$$\geq \frac{\beta\cdot(1-\beta)}{c_1}\cdot\epsilon\cdot N^2 = \Omega(\epsilon\cdot N^2)\ . \tag{69}$$
   Recall that $\ell = \log_2(1/\epsilon)$. Thus, there exists $k \in \{1,\ldots,(2\ell/3)+\log_2(2c_1)\}$ such that for $V^* \stackrel{\rm def}{=} \{v \in [N] : 2^{-k}N \leq |W_v| < 2^{-k+1}N\}$ it holds that $\sum_{v\in V^*} |W_v| = \Omega(\epsilon\cdot N^2/\ell)$. Fixing this $k$, we note that $|V^*| = \Omega(2^k\epsilon\cdot N/\ell)$ and thus $\Pr[R_k \cap V^* \neq \emptyset] > 8/9$, where $R_k$ is as selected in Step 2 of Algorithm 4.3 (i.e., $R_k$ is a random set of size $\Omega((2^k\epsilon/\ell)^{-1})$).

   For the sake of the analysis, consider viewing $S_k$ (which is a uniformly selected subset of $S$ that has size $\Theta(2^k)$ and is also selected in Step 2) as the union of two independently selected subsets of equal size, $S_k^1$ and $S_k^2$. Fixing any $v \in R_k \cap V^*$, we have $|W_v| \geq 2^{-k}N$ and so $\Pr[S_k^1 \cap W_v \neq \emptyset] > 8/9$. Finally, fixing any $u \in S_k^1 \cap W_v$, since $|W_{v,u}| \geq |W_v|/c_2 = \Omega(|W_v|) = \Omega(2^{-k}N)$, (where $W_{v,u} \subset (\Gamma(v)\setminus\Gamma(u)) \cup (\Gamma(u)\setminus\Gamma(v))$) we have $\Pr[S_k^2 \cap W_{v,u} \neq \emptyset] > 8/9$. Noting that all pairs $(R_k \times S_k) \cup (S_k \times S_k)$ are inspected by Algorithm 4.3, the claim follows for this subcase (i.e., with probability at least 2/3, Algorithm 4.3 finds a witness).

2. If $\sum_{v\in[N]:|\Gamma'(v)|\geq\epsilon^{2/3}N/2} |\Gamma'(v)| < \gamma\cdot\sum_{v\in[N]} |\Gamma'(v)|$, then we apply Part 2 of Claim 4.4.2 with $F = \{\{u,v\} : u \in \Gamma'(v),\ |\Gamma'(v)| \geq \epsilon^{2/3}N/2\}$. For every $v \in [N]$ we obtain a set $W_v \subseteq \Gamma'(v)\setminus F(v)$ (where $F(v) = \{u : \{v,u\} \in F\}$) such that if we set $\gamma = \frac{1}{2c_3c_4}$ then
$$\sum_{v\in[N]} |W_v| \geq \frac{1}{c_3}\cdot\left(\sum_{v\in[N]} |\Gamma'(v)|\right) - c_4\cdot|F| \tag{70}$$

$$\geq \ (1/c_3 - c_4 \cdot \gamma) \cdot \sum_{v \in [N]} |\Gamma'(v)| \tag{71}$$

$$\geq \ \frac{1}{2c_3} \cdot (1-\beta)\epsilon N^2 \ = \ \Omega(\epsilon \cdot N^2) \ . \tag{72}$$

Observe that by the constraint on the relation between $\beta$ and $\gamma$ that was imposed by the previous subcase, it suffices to set $\beta \leq \frac{1}{502 c_3 c_4}$. Since for every $v \in [N]$ and for every $u \in W_v$ we have that $|W_{v,u}| \geq |W_v|/c_2$, Equation (72) implies that

$$\sum_{v \in [N]} \sum_{u \in W_v} |W_{v,u}| \geq \sum_{v \in [N]} |W_v|^2/c_2 = \Omega(\epsilon^2 N^3) \ . \tag{73}$$

On the other hand, we have the following upper bound on the size of each $W_v$: $|W_v| \leq |\Gamma'(v) \setminus F(v)| < \epsilon^{2/3}N/2$, and $|W_{v,u}| \leq |W_v|$ holds (for every $u \in W_v$). Letting $U_w^{(1)} \stackrel{\text{def}}{=} \{v : w \in W_v\}$, for every $w$ it holds that $|U_w^{(1)}| < \epsilon^{2/3}N/2$ (because $v \in U_w^{(1)}$ implies $w \in \Gamma'(v)$ and $(v,w) \notin F$). Also, by Part 2b, we get $|U_w^{(2)}| \leq 10\epsilon^{4/3}N^2$ for every $w$. By applying Claim 4.4.3 (with $\epsilon$ set to $10\epsilon$ so that $|U_w^{(2)}| < \epsilon^{4/3}N^2$ for the new setting, while we still have that $\sum_{v\in[N]} \sum_{u \in W_v} |W_{v,u}| = \Omega(\epsilon^2 N^3)$ for this setting), we have that, with high probability, the sample $S$ selected in Step 1 of Algorithm 4.3 contains a witness (i.e., a triple $(v,u,w)$ such that $u \in W_v$ and $w \in W_{v,u}$).

Thus, based on Claim 4.4.3 (to be proven next), we completed the proof of Lemma 4.4. ■

### 4.2.3 Proof of Claim 4.4.3

We denote the random sample by $S$, and denote its elements by $v_1, \ldots, v_s, u_1, \ldots, u_s, w_1, \ldots, w_s$. We shall prove that, with probability at least $1 - O(s^{-1}\epsilon^{-2/3})$, there exists a triple $(i,j,k) \in [s]^3$ such that $u_j \in W_{v_i}$ and $w_k \in W_{v_i,w_j}$. The proof boils down to applying Chebyshev's Inequality to $\sum_{i,j,k \in [s]} \zeta_{i,j,k}$, where $\zeta_{i,j,k} = 1$ if $u_j \in W_{v_i}$ and $w_k \in W_{v_i,u_j}$, and $\zeta_{i,j,k} = 0$ otherwise. We first note that

$$\begin{aligned}
\mu \ &\stackrel{\text{def}}{=} \ \mathrm{Exp}_S\left[\sum_{i,j,k\in[s]} \zeta_{i,j,k}\right] && (74)\\
&= \ s^3 \cdot \mathrm{Pr}_{v,u,w\in[N]}[u \in W_v \ \wedge \ w \in W_{v,u}] && (75)\\
&= \ s^3 \cdot \frac{1}{N^3} \cdot \sum_{v\in[N]} \sum_{u \in W_v} |W_{v,u}| && (76)\\
&= \ \Omega(s^3 \cdot \epsilon^2) && (77)
\end{aligned}$$

where the last line follows by the first condition in the hypothesis. By Chebyshev's Inequality it follows that

$$\begin{aligned}
\mathrm{Pr}\left[\sum_{i,j,k\in[s]} \zeta_{i,j,k} = 0\right] \ &\leq \ \frac{\mathrm{Var}[\sum_{i,j,k\in[s]} \zeta_{i,j,k}]}{\mathrm{Exp}[\sum_{i,j,k\in[s]} \zeta_{i,j,k}]^2} && (78)\\
&= \ \mu^{-2} \cdot \left(\mathrm{Exp}\left[\left(\sum_{i,j,k\in[s]} \zeta_{i,j,k}\right)^2\right] - \mathrm{Exp}\left[\sum_{i,j,k\in[s]} \zeta_{i,j,k}\right]^2\right) && (79)
\end{aligned}$$

$$= \mu^{-2} \cdot \left( \left( \sum_{\overline{\ell} \in [s]^6} \mathrm{Exp}[\zeta_{i_1,j_1,k_1} \cdot \zeta_{i_2,j_2,k_2}] \right) - \mu^2 \right) \tag{80}$$

where $\overline{\ell} = (i_1, i_2, j_1, j_2, k_1, k_2)$. The upper bounds on $|W_v|, |W_{v,u}|, |U_w^{(1)}|$ and $|U_w^{(2)}|$ will be used in upper-bounding the large sum (i.e., $\sum_{\overline{\ell} \in [s]^6} \mathrm{Exp}[\zeta_{i_1,j_1,k_1} \cdot \zeta_{i_2,j_2,k_2}]$). We decompose the latter sum into partial sums that correspond to the following cases (regarding the relations between $i_1$-vs-$i_2$, $j_1$-vs-$j_2$, and $k_1$-vs-$k_2$).

***Case of $i \stackrel{\rm def}{=} i_1 = i_2$, $j \stackrel{\rm def}{=} j_1 = j_2$, and $k \stackrel{\rm def}{=} k_1 = k_2$.*** There are $s^3$ such terms, each having value $\mathrm{Exp}[\zeta_{i,j,k}^2] = \mathrm{Exp}[\zeta_{i,j,k}]$, which equals $\Pr_{v,u,w \in [N]}[u \in W_v \wedge w \in W_{v,u}] = \mu/s^3$. Thus, the total contribution of this case is $\mu$.

***Case of $i \stackrel{\rm def}{=} i_1 = i_2$, $j \stackrel{\rm def}{=} j_1 = j_2$, and $k_1 \neq k_2$.*** There are less than $s^4$ such terms, each having value $\mathrm{Exp}[\zeta_{i,j,k_1} \cdot \zeta_{i,j,k_2}]$, which equals

$$\begin{aligned} & \Pr_{v,u,w_1,w_2 \in [N]}[u \in W_v \wedge w_1, w_2 \in W_{v,u}] && (81) \\ & \leq \Pr_{v,u,w_1 \in [N]}[u \in W_v \wedge w_1 \in W_{v,u}] \cdot \max_{v,u,w_1 \in [N]} \left\{ \Pr_{w_2 \in [N]}[w_2 \in W_{v,u}] \right\} && (82) \\ & < \frac{\mu}{s^3} \cdot \epsilon^{2/3} && (83) \end{aligned}$$

where the last inequality is due to $|W_{v,u}| < \epsilon^{2/3} N$. Thus, the total contribution of this case is smaller than $s\epsilon^{2/3} \cdot \mu$.

***Case of $i \stackrel{\rm def}{=} i_1 = i_2$, $j_1 \neq j_2$, and $k \stackrel{\rm def}{=} k_1 = k_2$.*** There are less than $s^4$ such terms, each having value $\mathrm{Exp}[\zeta_{i,j_1,k} \cdot \zeta_{i,j_2,k}]$, which equals

$$\begin{aligned} & \Pr_{v,u_1,u_2,w \in [N]}[u_1, u_2 \in W_v \wedge w \in W_{v,u_1} \cap W_{v,u_2}] && (84) \\ & \leq \Pr_{v,u_1,w \in [N]}[u_1 \in W_v \wedge w \in W_{v,u_1}] \cdot \max_{v,u_1,w \in [N]} \left\{ \Pr_{u_2 \in [N]}[u_2 \in W_v] \right\} && (85) \\ & < \frac{\mu}{s^3} \cdot \epsilon^{2/3} && (86) \end{aligned}$$

where the last inequality is due to $|W_v| < \epsilon^{2/3} N$. Thus, the total contribution of this case is smaller than $s\epsilon^{2/3} \cdot \mu$.

***Case of $i \stackrel{\rm def}{=} i_1 = i_2$, $j_1 \neq j_2$, and $k_1 \neq k_2$.*** There are less than $s^5$ such terms, each having value $\mathrm{Exp}[\zeta_{i,j_1,k_1} \cdot \zeta_{i,j_2,k_2}]$, which equals

$$\begin{aligned} & \Pr_{v,u_1,u_2,w_1,w_2 \in [N]}[u_1, u_2 \in W_v \wedge w_1 \in W_{v,u_1} \wedge w_2 \in W_{v,u_2}] && (87) \\ & \leq \Pr_{v,u_1,w_1 \in [N]}[u_1 \in W_v \wedge w_1 \in W_{v,u_1}] \\ & \quad \cdot \max_{v,u_1,w_1 \in [N]} \left\{ \Pr_{u_2,w_2 \in [N]}[u_2 \in W_v \wedge w_2 \in W_{v,u_2}] \right\} && (88) \\ & < \frac{\mu}{s^3} \cdot (\epsilon^{2/3})^2 && (89) \end{aligned}$$

where the last inequality is due to $|W_v| < \epsilon^{2/3} N$ and $|W_{v,u_2}| < \epsilon^{2/3} N$. Thus, the total contribution of this case is smaller than $(s\epsilon^{2/3})^2 \cdot \mu$.

***Case of $i_1 \neq i_2$, $j \stackrel{\rm def}{=} j_1 = j_2$, and $k \stackrel{\rm def}{=} k_1 = k_2$.*** There are less than $s^4$ such terms, each having value $\mathrm{Exp}[\zeta_{i_1,j,k} \cdot \zeta_{i_2,j,k}]$, which equals

$$\begin{aligned}
&\mathrm{Pr}_{v_1,v_2,u,w\in[N]}[u \in W_{v_1} \cap W_{v_2} \ \wedge \ w \in W_{v_1,u} \cap W_{v_2,u}] && (90)\\
&\leq \ \mathrm{Pr}_{v_1,u,w\in[N]}[u \in W_{v_1} \ \wedge \ w \in W_{v_1,u}] \cdot \max_{v_1,u,w\in[N]} \left\{\mathrm{Pr}_{v_2\in[N]}[u \in W_{v_2}]\right\} && (91)\\
&< \ \frac{\mu}{s^3} \cdot \epsilon^{2/3} && (92)
\end{aligned}$$

where the inequality is due to $|U_u^{(1)}| < \epsilon^{2/3}N$ (and $u \in W_{v_2}$ iff $v_2 \in U_u^{(1)}$). Thus, the total contribution of this case is smaller than $s\epsilon^{2/3} \cdot \mu$.

***Case of $i_1 \neq i_2$, $j_1 \neq j_2$, and $k \stackrel{\rm def}{=} k_1 = k_2$.*** There are less than $s^5$ such terms, each having value $\mathrm{Exp}[\zeta_{i_1,j_1,k} \cdot \zeta_{i_2,j_2,k}]$, which equals

$$\begin{aligned}
&\mathrm{Pr}_{v_1,v_2,u_1,u_2,w\in[N]}[u_1 \in W_{v_1} \ \wedge \ u_2 \in W_{v_2} \ \wedge \ w \in W_{v_1,u_1} \cap W_{v_2,u_2}] && (93)\\
&\leq \ \mathrm{Pr}_{v_1,u_1,w\in[N]}[u_1 \in W_{v_1} \ \wedge \ w \in W_{v_1,u_1}] \cdot \max_{v_1,u_1,w\in[N]} \left\{\mathrm{Pr}_{u_2,v_2\in[N]}[w \in W_{v_2,u_2}]\right\} && (94)\\
&< \ \frac{\mu}{s^3} \cdot (\epsilon^{2/3})^2 && (95)
\end{aligned}$$

where the last inequality is due to $|U_w^{(2)}| < \epsilon^{4/3}N^2$ (and $w \in W_{v_2,u_2}$ iff $(v_2,u_2) \in U_w^{(2)}$). Thus, the total contribution of this case is smaller than $(s\epsilon^{2/3})^2 \cdot \mu$.

***Case of $i_1 \neq i_2$, $j \stackrel{\rm def}{=} j_1 = j_2$, and $k_1 \neq k_2$.*** There are less than $s^5$ such terms, each having value $\mathrm{Exp}[\zeta_{i_1,j,k_1} \cdot \zeta_{i_2,j,k_2}]$, which equals

$$\begin{aligned}
&\mathrm{Pr}_{v_1,v_2,u,w_1,w_2\in[N]}[u \in W_{v_1} \cap W_{v_2} \ \wedge \ w_1 \in W_{v_1,u} \ \wedge \ w_2 \in W_{v_2,u}] && (96)\\
&\leq \ \mathrm{Pr}_{v_1,u,w_1\in[N]}[u \in W_{v_1} \ \wedge \ w_1 \in W_{v_1,u}] \\
&\quad \cdot \max_{v_1,u,w_1\in[N]} \left\{\mathrm{Pr}_{v_2,w_2\in[N]}[u \in W_{v_2} \ \wedge \ w_2 \in W_{v_2,u}]\right\} && (97)\\
&< \ \frac{\mu}{s^3} \cdot (\epsilon^{2/3})^2 && (98)
\end{aligned}$$

where the last inequality is due to $|U_u^{(1)}| < \epsilon^{2/3}N$ and $|W_{v_2,u}| < \epsilon^{2/3}N$. Thus, the total contribution of this case is smaller than $(s\epsilon^{2/3})^2 \cdot \mu$.

***Case of $i_1 \neq i_2$, $j_1 \neq j_2$, and $k_1 \neq k_2$.*** There are less than $s^6$ such terms, each having value $\mathrm{Exp}[\zeta_{i_1,j_1,k_1} \cdot \zeta_{i_2,j_j,k_2}] = \mathrm{Exp}[\zeta_{i,j,k}]^2$, which equals $(\mu/s^3)^2$. Thus, the total contribution of this case is smaller than $\mu^2$.

Thus, we have one case (i.e., the first one) contributing $\mu$, three cases (each) contributing $s\epsilon^{2/3} \cdot \mu$, three cases (each) contributing $(s\epsilon^{2/3})^2 \cdot \mu$, and one case (i.e., the last one) contributing $\mu^2$. Using these upper bounds in Eq. (80), we obtain

$$\begin{aligned}
\mathrm{Pr}\left[\sum_{i,j,k\in[s]} \zeta_{i,j,k} = 0\right] \ &< \ \mu^{-2} \cdot \left(\left(\mu + 3 \cdot s\epsilon^{2/3} \cdot \mu + 3 \cdot (s\epsilon^{2/3})^2 \cdot \mu + \mu^2\right) - \mu^2\right) && (99)\\
&= \ \mu^{-1} \cdot \left(1 + 3s\epsilon^{2/3} + 3(s\epsilon^{2/3})^2\right). && (100)
\end{aligned}$$

Using $\mu = \Omega(s^3\epsilon^2)$ and a sufficiently large $s = O(\epsilon^{-2/3})$, we obtain an error bound of $O((s\epsilon^{2/3})^2/(s^3\epsilon^2)) = O(s^{-1}\epsilon^{-2/3}) < 1/3$, and the claim follows. ■

## 5 Larger Adaptive versus Non-adaptive Complexity Gaps

We start by establishing Theorem 1.2, which refers to the adaptive versus non-adaptive complexity gap of testing Bi-Clique Collections. We believe that the ideas underlying the adaptive algorithm and the non-adaptive lower bound (presented in Sections 5.1 and 5.2) can serve as a basis for establishing the larger gap stated in Conjecture 1.3. Indeed, as shown in Section 5.3, this is the case with respect to the non-adaptive lower bound (which indeed establishes Part 2 of Conjecture 1.3). In Section 5.4 we outline an adaptive algorithm that we believe to be suitable for Part 1 of Conjecture 1.3. In Section 5.4, we also state and prove a promise problem version of Conjecture 1.3.

### 5.1 The Adaptive Query Complexity of Bi-Clique Collection

The tester for $\mathcal{BCC}$ is obtained by extending the ideas that underly the tester for $\mathcal{CC}$ (i.e., Algorithm 3.1). The extension is relatively straightforward, but the analysis will have to address additional difficulties (i.e., beyond those encountered in the analysis of Algorithm 3.1). We mention, however, that the current algorithm uses two levels of adaptivity (e.g., inspecting the edge relation of selected neighbors) as compared with the single level of adaptivity employed by Algorithm 3.1 (which inspects, e.g., the edge relation of neighbors).

**Algorithm 5.1** (adaptive tester for $\mathcal{BCC}$): *On input $N$ and $\epsilon$ and oracle access to a graph $G = ([N], E)$, set $\ell = \log_2(1/\epsilon) + 2$, $t = \Theta(\ell^4)$, and proceed in $\ell$ iterations as follows: For $i = 1, \ldots, \ell$, uniformly select $100 \cdot 2^i$ start vertices and for each selected vertex $v \in [N]$ perform the following sub-test, denoted* sub-test$_i(v)$*:*

1. *Select, uniformly at random, a sample, $S$, of $t/(2^i\epsilon)$ vertices, and determine $\Gamma_S(v) = S \cap \Gamma(v)$ by making the queries $(v, w)$ for each $w \in S$. If $\Gamma_S(v) \neq \emptyset$ then select $u$ at random in $\Gamma_S(v)$ and continue to the following steps.* (Otherwise, halt and accept $v$.)
2. *Determine $\Gamma_S(u) = S \cap \Gamma(u)$ by making the queries $(u, w)$ for each $w \in S$.*
3. *If $|\Gamma_S(v) \times \Gamma_S(u)| \leq t/(2^i\epsilon)$ then check that for every $(w_1, w_2) \in \Gamma_S(v) \times \Gamma_S(u)$ it holds that $(w_1, w_2) \in E$. Otherwise* (i.e., $|\Gamma_S(v) \times \Gamma_S(u)| > t/(2^i\epsilon)$)*, uniformly select a sample of $t/(2^i\epsilon)$ pairs in $\Gamma_S(v) \times \Gamma_S(u)$ and check that each selected pair is in $E$.*
4. *Let $B = (\Gamma_S(v) \times \Gamma_S(v)) \cup (\Gamma_S(u) \times \Gamma_S(u))$. If $|B| \leq t/(2^i\epsilon)$ then check that for every $(w_1, w_2) \in B$ it holds that $(w_1, w_2) \notin E$. Otherwise* (i.e., $|B| > t/(2^i\epsilon)$)*, uniformly select a sample of $t/(2^i\epsilon)$ pairs in $B$ and check that each selected pair is in* not *$E$.*
5. *Select a sample of $t/(2^i\epsilon)$ pairs in $(\Gamma_S(v) \cup \Gamma_S(u)) \times (S \setminus (\Gamma_S(v) \cup \Gamma_S(u)))$ and check that each selected pair is* not *in $E$.*

*The sub-test* (i.e., sub-test$_i(v)$) *accepts if and only if all checks were positive* (i.e., no edges were missed in Step 3 and no edges were detected in Steps 4 and 5)*. The tester itself accepts if and only if all $\sum_{i=1}^{\ell} 10 \cdot 2^i$ invocations of the sub-test accepted.*

The query complexity of this algorithm is $\sum_{i=1}^{\ell}(100 \cdot 2^i) \cdot O(t/(2^i\epsilon)) = O(\ell \cdot t/\epsilon) = \widetilde{O}(1/\epsilon)$. Clearly, this algorithm accepts (with probability 1) any graph that is in $\mathcal{BCC}$. It remains to analyze its behavior on graphs that are $\epsilon$-far from $\mathcal{BCC}$.

**Lemma 5.2** *If $G = ([N], E)$ is $\epsilon$-far from $\mathcal{BCC}$, then on input $N, \epsilon$ and oracle access to $G$, Algorithm 5.1 rejects with probability at least $2/3$.*

Part 1 of Theorem 1.2 follows.

**Proof:** We proceed as in the proof of Lemma 3.2; that is, we will show that if Algorithm 5.1 accepts with probability at least 1/3 then the graph is $\epsilon$-close to $\mathcal{BCC}$. The proof makes use of a revised notion of $i$-good start vertices, which is defined on top of the notion of $i$-good edges. The definition refers to a parameter $\gamma$, which will be determined so that $\gamma = \Theta(1/t) = \Theta(\log^{-4}(1/\epsilon))$. Similarly to the analysis in the proof of Lemma 3.2, it is instructive to think of first setting $\gamma$ (whose setting is determined by another parameter, $\beta_2$, which is introduced subsequently), and then $t$ is set to be a (sufficiently large) constant factor larger than $1/\gamma$.

**Definition 5.2.1** *An edge $(v, u)$ is* $i$-good *if the following three conditions hold.*

1. *The number of missing edges in $\Gamma(v) \times \Gamma(u)$ is at most $\gamma \cdot 2^i\epsilon \cdot |\Gamma(v, u)| \cdot N$ edges, where $\Gamma(v, u) \stackrel{\rm def}{=} \Gamma(v) \cup \Gamma(u)$; that is, $|(\Gamma(v) \times \Gamma(u)) \setminus E| \leq \gamma \cdot 2^i\epsilon \cdot |\Gamma(v, u)| \cdot N$.*
2. *The number of edges in $(\Gamma(v) \times \Gamma(v)) \cup (\Gamma(u) \times \Gamma(u))$ is at most $\gamma \cdot 2^i\epsilon \cdot |\Gamma(v, u)| \cdot N$.*
3. *For every positive integer $j \leq j_0 \stackrel{\rm def}{=} \log_2(|\Gamma(v, u)|/(\gamma \cdot 2^i\epsilon N))$, the number of vertices in $\Gamma(v, u)$ that have at least $\gamma \cdot 2^{i+j}\epsilon \cdot N$ edges going out of $\Gamma(v, u)$ is at most $2^{-j} \cdot |\Gamma(v, u)|$.*

*A vertex $v$ is* $i$-good *if at least $0.8 \cdot |\Gamma(v)|$ of its neighbors yield an edge that is $i$-good; that is, if $|\{u \in \Gamma(v) : (v, u) \text{ is } i\text{-good}\}| \geq 0.8 \cdot |\Gamma(v)|$.*

**Claim 5.2.2** *If $v$ has degree at least $\gamma \cdot 2^i\epsilon \cdot N$ and is not $i$-good, then the probability that* sub-test$_i(v)$ *rejects is at least* 0.1.

Proof: By the hypothesis $|\Gamma(v)| \geq \gamma \cdot 2^i\epsilon \cdot N$, with high constant probability, Step 1 of sub-test$_i(v)$ generates a non-empty sample of vertices in $\Gamma(v)$. Conditioned on this event, since these vertices are uniformly distributed in $\Gamma(v)$, (and using the hypothesis that $v$ is not $i$-good), with probability at least 0.2 the vertex $u \in \Gamma(v)$ selected in this sample is such that $(v, u)$ is not $i$-good. We fix such an edge $(v, u)$ for the rest of this proof.

Assume that Condition 1 of $i$-goodness does not hold for $(v, u)$, and let

$$\rho \stackrel{\rm def}{=} \frac{\gamma \cdot 2^i\epsilon \cdot |\Gamma(v, u)| \cdot N}{|\Gamma(v)| \cdot |\Gamma(u)|} \geq \frac{\gamma \cdot 2^i\epsilon \cdot N}{\min(|\Gamma(v)|, |\Gamma(u)|)} \quad (101)$$

denote a lower bound on the fraction of missing edges in $\Gamma(v) \times \Gamma(u)$. (Note that the foregoing violation of Condition 1 may occur only if $\min(|\Gamma(v)|, |\Gamma(u)|) \geq \gamma \cdot 2^i\epsilon \cdot N$.) Then, with high constant probability, it holds that $\min(|\Gamma_S(v)|, |\Gamma_S(u)|) > m/2$, where

$$m \stackrel{\rm def}{=} \frac{t}{\epsilon 2^i} \cdot \frac{\min(|\Gamma(v)|, |\Gamma(u)|)}{N} \quad (102)$$

is the mininum of the expected sizes of $|\Gamma_S(v)|$ and $|\Gamma_S(u)|$, and is lower bounded by $t \cdot \gamma$ which is a (sufficiently large) constant. Also note that the members of $\Gamma_S(v)$ and $\Gamma_S(u)$ are distributed uniformly in $\Gamma(v)$ and $\Gamma(u)$, respectively. Considering $n = m/2$ uniformly distributed vertices in $\Gamma(v)$ and $n$ uniformly distributed vertices in $\Gamma(u)$, it follows (as in the proof of Claim 3.2.2) that, with high constant probability, the fraction of edges that are missing in the subgraph induced by the said sample is at least $\rho/2$. This implies that Step 3 rejects with high constant probability (regardless if it examines all pairs in $\Gamma_S(v) \times \Gamma_S(u)$ or just examines a random sample of $\frac{t}{2^i\epsilon} \geq \frac{t\gamma}{\rho}$ pairs).

The treatment of Condition 2 is similar, except that here we refer to the number of edges (in $(\Gamma(v)\times\Gamma(v))\cup(\Gamma(u)\times\Gamma(u)))$ over $|\Gamma(v)|^2+|\Gamma(u)|^2 = \Theta(|\Gamma(v,u)|^2)$. We conclude that if Condition 2 (of $i$-goodness of $(v,u)$) is violated, then Step 4 of the test rejects with high constant probability.

Finally, we turn to Condition 3 of $i$-goodness. Assuming that this condition does not hold for $(v,u)$, we claim that Step 5 of the test rejects with high constant probability. The proof is analogous to the analysis of Condition 2 in Claim 3.2.2, except that $\Gamma(v,u)$ replaces $\Gamma(v)$.

Thus (recalling the simple probabilistic assertions made at the start of the proof), sub-test$_i(v)$ rejects with probability at least $(1-\delta)\cdot 0.2$, where $\delta \in (0,1)$ is an arbitrary small constant, and the current claim follows. ■

**Claim 5.2.3** *If Algorithm 5.1 accepts with probability at least* $1/3$*, then for every* $i \in [\ell]$ *the number of vertices of degree at least* $\gamma \cdot 2^i\epsilon \cdot N$ *that are not* $i$*-good is at most* $2^{-i}\cdot N/4$*.*

Proof: Assuming to the contrary that the number of these vertices exceeds $2^{-i}\cdot N/4$, Claim 5.2.2 implies that a single invocation of sub-test$_i$ rejects with probability at least $0.025\cdot 2^{-i}$. Recalling that Algorithm 5.1 invokes sub-test$_i$ on $100\cdot 2^i$ uniformly selected random vertices, the claim follows. ■

Additional difficulties. As stated up-front, the current proof faces additional difficulties that were not encountered in the proof of Lemma 3.2. These difficulties refer to the partition reconstruction procedure, which is supposed to provide an approximately good partition of the graph to bi-cliques. The first problem refers to the case that $(v,u)$ is $i$-good, but most of $\Gamma(v,u)$ belongs to previously identified bi-cliques and furthermore these vertices reside in $\Gamma(u)$ (rather than in $\Gamma(v)$). Thus, we cannot "charge" these vertices to edges that are adjacent to $v$, but rather develop a charging rule that allows us to charge $v$ indirectly via its typical neighbors $u$. The second problem refers to the treatment of low-degree vertices, and it arises from the fact that vertices in $\Gamma(v,u)$ may have vastly different degrees (which, indeed, occurs in the case that $\Gamma(v)$ has a significantly different cardinality than $\Gamma(u)$). Our solution is based on using two different degree thresholds (depending on the relation between the degree of a vertex and the degree of most of its neighbors). With this motivation in mind, we turn to the actual description of the (iterative) partition-reconstruction procedure.

The partition reconstruction procedure. The iterative procedure is initiated with $C = L_0 = L_0^{(1)} = L_0^{(2)} = L_0^{(I)} = \emptyset$, $R_0 = [N]$ and $i = 1$, where $C$ denotes the set of vertices "covered" (by bi-cliques) so far, $R_{i-1}$ denotes the set of "remaining" vertices after iteration $i-1$ and $L_{i-1}$ denotes the set of vertices cast aside (as having "low degree") in iteration $i-1$. The set $L_{i-1}$ is the union of three sets, $L_{i-1}^{(1)}$, $L_{i-1}^{(2)}$, and $L_{i-1}^{(I)}$, where the first two sets correspond to two degree thresholds, denoted $\beta_1$ and $\beta_2$, and the third set consists of many subsets that use intermediate thresholds (for avoiding a non-smooth transition). In each iteration, a set $F_i$ of edges is constructed, where each edge in $F_i$ is used to determine a biclique (or, more precisely, a pair of subsets that are close to being a biclique). We shall set $\beta_1 = \Theta(1/\ell) = \Theta(\log^{-1}(1/\epsilon))$ and $\beta_2 = \Theta(\beta_1/\ell) = \Theta(1/\ell^2)$. Recall that $\gamma = \Theta(\log^{-4}(1/\epsilon)$, so that $\gamma = O(\beta_2/\ell^2)$ (and in the analysis we shall determine the sufficient size of the constant $c$ such that $\gamma = \beta_2/(c\ell^2)$).

The $i^{\text{th}}$ iteration proceeds as follows, where $i = 1,\ldots,\ell$ and $F_i$ is initialized to $\emptyset$.

1. Pick an arbitrary vertex $v \in R_{i-1}\setminus C$ that satisfies the following three conditions
   (a) $v$ is $i$-good.

(b) $v$ has sufficiently high degree in the following sense: either $|\Gamma(v)| \geq \beta_1 \cdot 2^i \epsilon \cdot N$ or for some $k \in [\ell']$, where $\ell' = \log_{0.9}(\beta_2/\beta_1) = O(\log \ell)$, both $|\Gamma(v)| \geq 0.9^k \cdot \beta_1 \cdot 2^i \epsilon \cdot N$ and $\phi_k(v)$ hold, where $\phi_k(v)$ represents the condition that a significant fraction of $v$'s neighbors have a significantly higher degree than $v$ itself; specifically, $\phi_k(v)$ holds if

$$\left|\left\{w \in \Gamma(v) \colon |\Gamma(w)| > \left(1.1 + \frac{k}{10\ell'}\right) \cdot |\Gamma(v)|\right\}\right| \;>\; \frac{|\Gamma(v)|}{100\ell}\,. \tag{103}$$

Note that $\phi_{\ell'}(v)$ holds if $|\{w \in \Gamma(v) : |\Gamma(w)| > 1.2 \cdot |\Gamma(v)|\}|$ is greater than $|\Gamma(v)|/100\ell$, and the corresponding degree bound is $\beta_2 \cdot 2^i \epsilon \cdot N$ (because $0.9^{\ell'} = \beta_2/\beta_1$).

(c) There exists $u \in \Gamma(v) \setminus C$ such that the edge $(v, u)$ is $i$-good and

$$\left|(\Gamma(v,u) \setminus C) \setminus \left(\bigcup_{j \leq i-1} L_j\right)\right| \;\geq\; \frac{|\Gamma(v,u)|}{5}$$

(i.e., relatively few vertices of $\Gamma(v,u)$ are covered by $C$ or cast aside in previous iterations due to having low degree).

If no such vertex $v$ exists, then define

$$L_i^{(1)} \;=\; \{v \in R_{i-1} \setminus C : \neg\phi_1(v) \wedge (|\Gamma(v)| < \beta_1 \cdot 2^i \epsilon \cdot N)\}, \tag{104}$$

$$L_i^{(I)} \;=\; \bigcup_{k \in [\ell'-1]} \{v \in R_{i-1} \setminus C : \phi_k(v) \wedge \neg\phi_{k+1}(v) \wedge (|\Gamma(v)| < 0.9^k \beta_1 \cdot 2^i \epsilon \cdot N)\}, \tag{105}$$

$$L_i^{(2)} \;=\; \{v \in R_{i-1} \setminus C : \phi_{\ell'}(v) \wedge (|\Gamma(v)| < \beta_2 \cdot 2^i \epsilon \cdot N)\}, \tag{106}$$

$L_i = L_i^{(1)} \cup L_i^{(I)} \cup L_i^{(2)}$, and $R_i = R_{i-1} \setminus (L_i \cup C)$.

If $i < \ell$ then proceed to the next iteration, and otherwise terminate.

2. For a vertex $v$ as selected in Step 1, pick an arbitrary $u \in \Gamma(v) \setminus C$ satisfying Condition 1c. Let $C_{v,u} = \{w \in \Gamma(v,u) : |\Gamma(w) \setminus \Gamma(v,u)| < |\Gamma(v,u)|\}$. Form a new bi-clique with the vertex set $C'_{v,u} \leftarrow C_{v,u} \setminus C$, and update $F_i \leftarrow F_i \cup \{(v,u)\}$ and $C \leftarrow C \cup C'_{v,u}$. This bi-clique will have $\Gamma'(v) \stackrel{\text{def}}{=} \Gamma(v) \cap C'_{v,u}$ on one side and $\Gamma'(u) \stackrel{\text{def}}{=} \Gamma(u) \cap C'_{v,u}$ on the other side.

Note that by Condition 1c (and the definition of $i$-goodness), for every $(v,u) \in F_i$, it holds that $|C_{v,u}| > (1 - o(1)) \cdot |\Gamma(v,u)|$ and $|\Gamma(v,u) \setminus C| \geq |\Gamma(v,u)|/5$. Thus, $|C'_{v,u}| \geq |C_{v,u}| - |\Gamma(v,u) \cap C| \geq |\Gamma(v,u)|/6$, which allows translating quality guarantees that are quantified in terms of $|\Gamma(v,u)|$ to similar guarantees in terms of $|C'_{v,u}|$. In fact, $|C'_{v,u} \setminus (\bigcup_{j \leq i-1} L_j)| \geq |\Gamma(v,u)|/6$, which enables further translation of these guarantees to quantification in terms of $|C'_{v,u} \cap R_{i-1}|$.

**Claim 5.2.4** *Referring to the partition reconstruction procedure, for every $i \in [\ell]$, the following holds.*

1. *The number of missing edges inside the bi-cliques formed in iteration $i$ is at most $12\gamma\epsilon \cdot N^2$; that is,*

$$\left|\bigcup_{(v,u) \in F_i} \{(w_1, w_2) \in \Gamma'(v) \times \Gamma'(u) : (w_1, w_2) \notin E\}\right| \leq 12\gamma\epsilon \cdot N^2.$$

2. *The number of* "superfluous" *edges inside the bi-cliques formed in iteration $i$ is at most* $12\gamma\epsilon \cdot N^2$*; that is,*

$$\left| \bigcup_{(v,u)\in F_i} \{(w_1,w_2) \in (\Gamma'(v)\times\Gamma'(v)) \cup (\Gamma'(u)\times\Gamma'(u)) : (w_1,w_2)\in E\} \right| \le 12\gamma\epsilon \cdot N^2.$$

3. *The number of* "superfluous" *edges between bi-cliques formed in iteration $i$ and either $R_i$ or other bi-cliques formed in the same iteration is at most* $36\ell \cdot \gamma\epsilon \cdot N^2$*; actually,*

$$\left| \bigcup_{(v,u)\in F_i} \{(w_1,w_2) \in C'_{v,u} \times (R_{i-1} \setminus C'_{v,u}) : (u,w)\in E\} \right| \le 36\ell \cdot \gamma\epsilon \cdot N^2.$$

4. $|R_i| \le 2^{-i} \cdot N$ *and* $|L_i| \le 2^{-(i-1)} \cdot N$.

Thus, the total number of violations caused by the bi-cliques that are formed by the foregoing procedure is upper-bounded by $(36+o(1))\ell^2 \cdot \gamma\epsilon \cdot N^2 = o(\epsilon N^2)$.

Proof: We prove all items simultaneously, by induction from $i=0$ to $i=\ell$. Needless to say, all items hold vacuously for $i=0$, and thus we focus on the induction step.

Starting with Item 1, we note that every $(v,u)\in F_i$ is $i$-good and thus the number of edges missing in $\Gamma'(v)\times\Gamma'(u) \subseteq \Gamma(v)\times\Gamma(u)$ is at most $\gamma 2^i\epsilon\cdot|\Gamma(v,u)|\cdot N$. As in the proof of Claim 3.2.4, we need to relate $|\Gamma(v,u)|$ to $|C'_{v,u}\cap R_{i-1}|$ (in order to upper-bound the contribution of all pairs in $F_i$). We recall that $C'_{v,u} = C_{v,u}\setminus C$, where $C$ is the set of vertices that are already covered when $(v,u)$ is added to $F_i$. Also recall that $|\Gamma(v,u)\setminus C_{v,u}| = o(1)\cdot|\Gamma(v,u)|$ and $|(\Gamma(v,u)\setminus C)\setminus L| \ge |\Gamma(v,u)|/5$, where $L \stackrel{\rm def}{=} \bigcup_{j\in[i-1]} L_j$. Using $C'_{v,u} = (C'_{v,u}\cap R_{i-1}) \cup (C'_{v,u}\cap L)$, we get that $C'_{v,u}\cap R_{i-1} = (C_{v,u}\setminus C)\setminus L$ and it follows that $|C'_{v,u}\cap R_{i-1}| \ge |(\Gamma(v,u)\setminus C)\setminus L| - o(|\Gamma(v,u)|) > |\Gamma(v,u)|/6$. Combining all the above (and recalling that the sets $C'_{v,u}$ are disjoint), we obtain

$$\begin{aligned}
\left| \bigcup_{(v,u)\in F_i} \{(w_1,w_2)\in \Gamma'(v)\times\Gamma'(u) : (w_1,w_2)\notin E\} \right| &\le \gamma 2^i\epsilon \cdot \sum_{(v,u)\in F_i} |\Gamma(v,u)|\cdot N && (107)\\
&\le \gamma 2^i\epsilon \cdot 6|R_{i-1}|\cdot N. && (108)
\end{aligned}$$

Using the induction hypothesis regarding $R_{i-1}$ (i.e., $|R_{i-1}| < 2^{-(i-1)}\cdot N$), Item 1 follows.

Item 2 is proved in a similar fashion. As for Item 3, we adapt the proof of Item 2 of Claim 3.2.4. Specifically, the number of edges in $C_{v,u}\times([N]\setminus C_{v,u})$ is upper-bounded by the sum of $|C_{v,u}\times(\Gamma(v,u)\setminus C_{v,u})|$ and the number of edges in $C_{v,u}\times([N]\setminus\Gamma(v,u))$. Using Condition 3 of $i$-goodness (of $(v,u)$), we upper-bound both $|\Gamma(v,u)\setminus C_{v,u}|$ and the number of edges of the second type. Hence, the number of edges in $C'_{v,u}\times(R_{i-1}\setminus C'_{v,u}) \subseteq C_{v,u}\times([N]\setminus C_{v,u})$ is at most $3\ell\cdot\gamma 2^i\epsilon\cdot|\Gamma(v,u)|\cdot N$. Using again $\sum_{(v,u)\in F_i}|\Gamma(v,u)| < 6|R_{i-1}|$ and $|R_{i-1}| < 2^{-(i-1)}\cdot N$, we establish Item 3.

Turning to Item 4, we first note that $L_i \subseteq R_{i-1}$ and thus $|L_i| \le |R_{i-1}| \le 2^{-(i-1)}\cdot N$. As for $R_i$, let us consider all the cases that might lead to placing a vertex $v$ in $R_i$; that is, the various violations of the three conditions in Step 1.

Violation of Condition (b): not having sufficiently high degree. We observe that vertices that violate Condition (b) do not contribute to $R_i$, because each such vertex is either covered in iteration

$i$ or ends-up in $L_i$. Specifically, let $v$ be an arbitrary vertex that violates Condition (b), and let $k(v) \in \{0,1,\ldots,\ell'\}$ be the largest index $k$ such that $\phi_k(v)$ holds (where $\phi_0$ is fictitiously defined such that it always holds). Then, Condition (b) is equivalent to requiring that $|\Gamma(v)| \geq 0.9^{k(v)} \cdot \beta_1 \cdot 2^i\epsilon \cdot N$ holds. Indeed, if the latter condition does not hold, then $v$ is placed in $L_i$ (and the converse holds as well).

In the subsequent cases, we shall assume that Condition (b) holds with respect to the vertex $v$.

Violation of Condition (a): not being $i$-good. Here we refer to vertices that are not $i$-good although they have degree at least $\beta_2 \cdot 2^i\epsilon \cdot N > \gamma \cdot 2^i\epsilon \cdot N$. By Claim 5.2.3, the number of vertices of this type is at most $2^{-i} \cdot N/4$.

Violation of Condition (c). Here we refer to vertices that satisfy both Conditions (a) and (b) but violate Condition (c), which refers to the existence of a good edge that yields a bi-clique with sufficiently many new vertices. The rest of the proof is devoted to upper-bounding the number of such vertices. Loosely speaking, this is done by using the upper bound established in Item 3, while relying on the hypothesis that these vertices satisfy both Conditions (a) and (b).

Recalling that we refer to vertices that satisfy both Conditions (a) and (b), we first upper-bound the number of vertices that have relatively many neighbors in the current $C$, i.e., vertices $v$ such that $|\Gamma(v) \cap C| \geq |\Gamma(v)|/8$. As in the proof of Claim 3.2.4, each such vertex $v$ requires at least $|\Gamma(v)|/8 \geq \beta_2 \cdot 2^i\epsilon \cdot N/8$ edges from $C' \stackrel{\text{def}}{=} \bigcup_{(v',u') \in \bigcup_{j\in[i]} F_j} C'_{v',u'}$ to it, whereas by Item 3 the total number of edges going out from $C'$ to $R_i$ is at most $i \cdot 36\ell \cdot \gamma\epsilon \cdot N^2 \leq 36\ell^2 \cdot \gamma\epsilon \cdot N^2$. Hence, the number of vertices of this type is upper-bounded by

$$\frac{36\ell^2 \cdot \gamma\epsilon \cdot N^2}{\beta_2 \cdot 2^i\epsilon \cdot N} = \frac{36\ell^2 \cdot \gamma}{\beta_2} \cdot 2^{-i}N < 0.1 \cdot 2^{-i}N, \tag{109}$$

where the last inequality uses $\gamma < \beta_2/(360\ell^2)$.

In the rest of the proof we consider only vertices that have relatively few neighbors in the current $C$ (i.e., $|\Gamma(v) \cap C| \leq |\Gamma(v)|/8$). In particular, by the case hypothesis (i.e., $v$ is $i$-good), there exists $u \notin C$ such that $(v,u)$ is $i$-good (because the fraction of "non-good" pairs $(v,u)$ is at most 0.2). Thus, we focus on the condition $|(\Gamma(v,u) \setminus C) \setminus L| > |\Gamma(v,u)|/5$, where $L \stackrel{\text{def}}{=} \bigcup_{j \leq i-1} L_j$ and $C$ denotes the current set of covered vertices. We distinguish three cases with respect to the relation between $|\Gamma(v)|$ and $|\Gamma(u)|$. Actually, *letting $U_v$ denote the set of vertices $u \in \Gamma(v) \setminus C$ such that $(v,u)$ is $i$-good*, we consider three cases regarding the relations of $|\Gamma(v)|$ and $\{|\Gamma(u)| : u \in U_v\}$.

***Case 1: there exists $u \in U_v$ such that $|\Gamma(v)| > 1.3|\Gamma(u)|$.*** We just pick an arbitrary such $u$, and note that, using the case hypothesis (which implies $|\Gamma(v)| > |\Gamma(v,u)|/2$), it suffices to show that $|(\Gamma(v) \setminus C) \setminus L| > |\Gamma(v)|/2$. Since $|\Gamma(v) \cap C| \leq |\Gamma(v)|/8$, we focus on upper-bounding $|\Gamma(v) \cap L|$ for but a small number of vertices $v$ (that fall under this case). The intuition is that in the current case $\neg\phi_1(v)$ holds, and so the fact that $v \notin L_i$ implies that $|\Gamma(v)| \geq \beta_1 \cdot 2^i\epsilon N$. On the other hand, each vertex in $\Gamma(v) \cap L_j$ has at most $\beta_2 \cdot 2^j\epsilon N$ neighbors of degree at least $\beta_1 \cdot 2^i\epsilon N$, which yields a total count of $2\beta_2\epsilon N^2$ edges in $L_j \times (R_{i-1} \setminus L_i)$. Thus, the number of vertices $v \in R_{i-1} \setminus L_i$ for which $|\Gamma(v) \cap L| > |\Gamma(v)|/8$ holds is sufficiently small. Details follow.

Using the hypothesis that $(v,u)$ is $i$-good (and referring to Condition 2 of Definition 5.2.1), we note that the number of edges with both endpoints in $\Gamma(v)$ is at most $\gamma \cdot 2^i\epsilon \cdot |\Gamma(v,u)| \cdot N \leq$

$\gamma \cdot 2^{i+1}\epsilon \cdot |\Gamma(v)| \cdot N$. Thus, less than a $1/(200\ell)$ fraction of the vertices in $\Gamma(v)$ have more than $200\ell \cdot \gamma \cdot 2^{i+1}\epsilon \cdot N < \beta_2 \cdot 2^i\epsilon \cdot N/100 \leq |\Gamma(v)|/100$ such edges, where the inequalities are due to $\gamma \leq \beta_2/40000\ell$ and $|\Gamma(v)| \geq \beta_2 \cdot 2^i\epsilon \cdot N$ (since $v \notin L_i$). By Condition 3 of Definition 5.2.1, at most a $1/(200\ell)$ fraction of the vertices in $\Gamma(v)$ have at least $200\ell \cdot \gamma \cdot 2^i\epsilon \cdot N < |\Gamma(v)|/100$ edges going out of $\Gamma(v,u)$. We conclude that less than a $1/(100\ell)$ fraction of the vertices in $\Gamma(v)$ have degree exceeding $|\Gamma(u)| + 0.02|\Gamma(v)| < |\Gamma(v)|$, and so $\neg\phi_1(v)$ holds. The latter fact allows us to increase our lower bound on $|\Gamma(v)|$ (from $|\Gamma(v)| \geq \beta_2 \cdot 2^i\epsilon N$) to $|\Gamma(v)| \geq \beta_1 \cdot 2^i\epsilon N$ (using again $v \notin L_i$). Thus, if $|\Gamma(v) \cap L| > |\Gamma(v)|/8$ then there exist at least $\beta_1 \cdot 2^i\epsilon N/8$ edges from $L = \bigcup_{j \leq i-1} L_j$ to $v$.

We upper-bound the number of such vertices $v$ (i.e., for which $|\Gamma(v) \cap L| > |\Gamma(v)|/8$), by upper-bounding the number of edges that may go from $L$ to any vertex of degree at least $\beta_1 \cdot 2^i\epsilon N$. The contribution of each vertex in $L_j^{(2)}$ to this number is at most $\beta_2 \cdot 2^j\epsilon N$, because vertices in $L_j^{(2)}$ have degree at most $\beta_2 \cdot 2^j\epsilon N$. As for the vertices in $L_j \setminus L_j^{(2)}$, each such vertex $u'$ violates $\phi_{\ell'}$ and thus can contribute at most $|\Gamma(u')|/100\ell$ to this number, because at most a $1/(100\ell)$ fraction of its neighbors have degree exceeding $1.2|\Gamma(u')| < \beta_1 \cdot 2^i\epsilon N$ (since $|\Gamma(u')| < \beta_1 \cdot 2^j\epsilon N$ and $j \leq i-1$), whereas we count edges to vertices of degree at least $\beta_1 \cdot 2^i\epsilon N$. Thus, the contribution of each vertex in $u' \in L_j$ to the count is at most $\max(\beta_2 \cdot 2^j\epsilon N, |\Gamma(u')|/100\ell) \leq \beta_1 \cdot 2^j\epsilon N/100\ell$ (since $\beta_2 \leq \beta_1/100\ell$ and $|\Gamma(u')| < \beta_1 \cdot 2^j\epsilon N$). Recalling that $|L_j| \leq |R_{j-1}| \leq 2^{-(j-1)}N$, it follows that the number of bad vertices (i.e., vertices $v$ of degree at least $\beta_1 \cdot 2^i\epsilon N$ with at least $|\Gamma(v)|/8$ neighbors in $L$) is at most

$$\frac{\sum_{j \leq i-1} |L_j| \cdot \beta_1 \cdot 2^j\epsilon \cdot N/100\ell}{\beta_1 \cdot 2^i\epsilon N/8} \leq \frac{(i-1) \cdot \beta_1 \cdot 2\epsilon \cdot N^2/100\ell}{\beta_1 \cdot 2^i\epsilon N/8} \tag{110}$$

$$< 0.16 \cdot 2^{-i}N, \tag{111}$$

whereas the rest of the vertices $v \in R_{i-1} \setminus L_i$ satisfy $|\Gamma(v) \cap L| \leq |\Gamma(v)|/8$. Recalling that $|\Gamma(v) \cap C| \leq |\Gamma(v)|/8$, we conclude that $|(\Gamma(v) \setminus C) \setminus L| > |\Gamma(v)|/2$, and the claim follows; that is, the current case is only responsible for $0.16 \cdot 2^{-i}N$ vertices violating Condition (c).

***Case 2: for every $u \in U_v$ it holds that $|\Gamma(v)| < 0.7|\Gamma(u)|$.*** We first show that for every such $u$ it holds that $|\Gamma(u) \cap L| \leq |\Gamma(u)|/8$, and later consider two subcases. In the first subcase $|\Gamma(u) \cap C| \leq |\Gamma(u)|/8$ holds (for some relevant $u$), and so we obtain $|(\Gamma(u) \setminus C) \setminus L| > |\Gamma(u)|/2$ and use $|\Gamma(u)| > |\Gamma(v,u)|/2$ to conclude that $v$ satisfies Condition (c). In the other subcase, where $|\Gamma(u) \cap C| > |\Gamma(u)|/8$ holds for all relevant $u$, we bound the number of vertices $v$ for which this may occur.

The proof that $|\Gamma(u) \cap L| \leq |\Gamma(u)|/8$ is supported by the intuition that almost all vertices in $\Gamma(u)$ have approximately the same degree as $v$ and satisfy $\phi_{\ell'}$ (since most of their neighbors have degree approximately $|\Gamma(u)| > (10/7)|\Gamma(v)|$), which implies that they cannot be in $L$ (because vertices in $L$ that satisfy $\phi_{\ell'}$ have degree at most $\beta_2 \cdot 2^{i-1}\epsilon N$, whereas $v \in R_{i-1} \setminus L_i$ has degree at least $\beta_2 \cdot 2^i\epsilon N$). Details follow.

We start by showing that almost all vertices in $\Gamma(u)$ satisfy $\phi_{\ell'}$. Analogously to the previous case, at most a 0.01 fraction of the vertices in $\Gamma(u)$ have more than $0.02 \cdot |\Gamma(v)|$ neighbors not in $\Gamma(v)$. On the other hand, by using Condition 1 of Definition 5.2.1, at least a 0.99 fraction of the vertices in $\Gamma(u)$ have at least $0.99 \cdot |\Gamma(v)|$ neighbors in $\Gamma(v)$, whereas at least a 0.99 fraction of the vertices in $\Gamma(v)$ have degree at least $0.99 \cdot |\Gamma(u)|$. Let us denote by $Y$ the subset of $\Gamma(u)$ containing vertices $v'$ such that $|\Gamma(v')| \leq 1.02 \cdot |\Gamma(v)|$ and $\Gamma(v') \cap \Gamma(v)$

contains at least $0.98 \cdot |\Gamma(v)|$ vertices of degree at least $0.99 \cdot |\Gamma(u)|$. Then, $|Y| > 0.98|\Gamma(u)|$, because a 0.98 fraction of the vertices in $\Gamma(u)$ have both degree at most $1.02 \cdot |\Gamma(v)|$ and at least $0.99 \cdot |\Gamma(v)|$ neighbors in $\Gamma(v)$ (whereas at most a 0.01 fraction of the vertices in $\Gamma(v)$ have degree smaller than $0.99 \cdot |\Gamma(u)|$). We note that each vertex in $Y$ has degree at most $1.02 \cdot |\Gamma(v)| < 0.72 \cdot |\Gamma(u)|$, whereas at least a $0.98/1.02$ fraction (which is significantly greater than $(100\ell)^{-1}$) of its neighbors have degree at least $0.99 \cdot |\Gamma(u)| > 1.2 \cdot 0.72 \cdot |\Gamma(u)|$, which implies that each vertex in $Y$ satisfies $\phi_{\ell'}$. Using the latter fact and recalling that each vertex in $Y$ has degree at least $0.99 \cdot |\Gamma(v)| \geq 0.99 \cdot \beta_2 \cdot 2^i \epsilon N$ (since $v \notin L_i$), we show that $Y \cap L = \emptyset$. The latter claim follows by noting that for every $v' \in L$ that satisfies $\phi_{\ell'}$ it holds that $|\Gamma(v')| < \beta_2 \cdot 2^{i-1} \epsilon N$, whereas every $v' \in Y$ satisfies both $\phi_{\ell'}$ and $|\Gamma(v')| > 0.99 \cdot \beta_2 \cdot 2^i \epsilon N$. Finally, using $Y \cap L = \emptyset$ and $|Y| \geq 0.98|\Gamma(u)|$, we get $|\Gamma(u) \cap L| \leq |\Gamma(u) \setminus Y| \leq 0.02|\Gamma(u)|$.

Having established $|\Gamma(u) \cap L| \leq |\Gamma(u)|/8$, one may attempt to provide a similar upper bound for $|\Gamma(u) \cap C|$. However, unlike in the previous case (or rather in the preliminary proof that $\Gamma(v) \cap C$ is small), here we cannot *directly* charge the vertices in $\Gamma(u) \cap C$ to edges going out from $C$ to $v$. Still, *an indirect charging rule will work*; that is, *we first charge such vertices to $u$, and then distribute the charge to $u$'s neighbors.* This will yield an upper bound on the number of vertices $v$ for which there exists no $u \in U_v$ such that $|\Gamma(u) \cap C| \leq |\Gamma(u)|/8$. In light of the foregoing, we consider two subcases.

1. The easy subcase is the one where there exists $u \in U_v$ such that $|\Gamma(u) \cap C| \leq |\Gamma(u)|/8$ (and $|\Gamma(u)| > |\Gamma(v)|/0.7$, by the case hypothesis). In this subcase, we conclude that $v$ satisfies Condition (c), since $|(\Gamma(v,u) \setminus C) \setminus L| > |\Gamma(u)|/2 > |\Gamma(v,u)|/2$. That is, this subcase does not contribute any vertices that violate Condition (c).

2. The other subcase refers to the case that for every $u \in U_v$ it holds that $|\Gamma(u) \cap C| > |\Gamma(u)|/8$. This means that there are at least $|\Gamma(u)|/8$ edges going out from $C$ to $u$. Wishing to charge these edges to the initial vertex $v$ (while considering all initial $v \in R_{i-1} \setminus L_i$), we charge each neighbor of $u$ by one eighth of an edge (i.e., 1/8 unit) as its share in the total number of edges going from $C$ to $u$. That is, these $|\Gamma(u) \cap C|$ edges generate a charging of $|\Gamma(u)|/8$ units, which is distributed equally among all vertices in $\Gamma(u)$. (No overcharging occurs since $|\Gamma(u) \cap C| > |\Gamma(u)|/8$.)

   (Indeed, an important observation is that we are not concerned with the existence of a specific $u \in U_v$ that violates $|\Gamma(u) \cap C| \leq |\Gamma(u)|/8$, but should be concerned only if this violation occurs for all $u \in U_v$ (such that $|\Gamma(u)| > |\Gamma(v)|/0.7$), since otherwise we are done by the first subcase. Thus, we get into trouble with $v$ only if, for every $u \in U_v$ both $|\Gamma(u)| > |\Gamma(v)|/0.7$ and $|\Gamma(u) \cap C| > |\Gamma(u)|/8$ holds.)[9]

   Let us denote the set of such bad (initial) vertices by $B$; that is, $v \in B$ *if for every $u \in U_v$ both $|\Gamma(u)| > |\Gamma(v)|/0.7$ and $|\Gamma(u) \cap C| > |\Gamma(u)|/8$ holds.* Note that each vertex $v \in B$ is charged with at least $(|\Gamma(v)|/2) \cdot (1/8) > \beta_2 \cdot 2^i \epsilon N/16$ (units that account for) edges going from $C$ to $\Gamma(v)$, where $|\Gamma(v)|/2$ is a lower bound on the number of vertices $u \in \Gamma(v)$ such that $u \notin C$ and $(v,u)$ is $i$-good.[10] Since the total number of edges going out from $C$ is at most $36\ell^2 \cdot \gamma\epsilon \cdot N^2$, we upper-bound $|B|$ by $0.1 \cdot 2^{-i} N$ (as in Eq. (109),

[9]Again, these conditions are guaranteed by the case and subcase hypotheses.

[10]Recall that the fraction of vertices $u \in \Gamma(v)$ such that $u \in C$ is at most $1/8$, whereas the fraction of vertices $u \in \Gamma(v)$ such that $(v,u)$ is not $i$-good is at most $0.2 < 3/8$.

except that here we use $\gamma < \beta_2/(6000\ell^2)$).[11]

To re-cap, note that we showed that the current case is only responsible for $0.1 \cdot 2^{-i}N$ vertices that violating Condition (c).

***Case 3: there exists $u \in U_v$ such that $0.7|\Gamma(u)| \leq |\Gamma(v)| \leq 1.3|\Gamma(u)|$.*** In addition, we assume here that Case 1 does not hold. We first note that the analysis of $|\Gamma(u) \cap C|$ (for all $u \in U_v$) as presented in Case 2 still holds. Thus, for all but $0.1 \cdot 2^{-i}N$ vertices $v$, there exists a vertex $u$ such that for every $u \in U_v$ it holds that $|\Gamma(u) \cap C| \leq |\Gamma(u)|/8$. These vertices will contribute to violation of Condition (c), but we shall show that all other vertices satisfy Condition (c).

Thus, we consider any arbitrary $v$ such that there there exists a vertex $u \in U_v$ that satisfies $|\Gamma(u) \cap C| \leq |\Gamma(u)|/8$ (and $|\Gamma(v)| \leq 1.3|\Gamma(u)|$). We shall show, below, that $|\Gamma(u) \cap L| \leq |\Gamma(u)|/8$, and conclude that $|(\Gamma(u) \setminus C) \setminus L| \geq |\Gamma(u)|/2$, which in turn is lower-bounded by $|\Gamma(v,u)|/5$ (since $|\Gamma(u)| \geq |\Gamma(v,u)|/2.3$, which follows from $|\Gamma(v)| \leq 1.3|\Gamma(u)|$).

The claim $|\Gamma(u) \cap L| \leq |\Gamma(u)|/8$ is supported by the intuition that almost all vertices in $\Gamma(u)$ have approximately the same degree as $v$. However, in the current case these vertices do not necessarily satisfy $\phi_{\ell'}$ and so their being in $L$ does not necessarily mean their having degree below $\beta_2 \cdot 2^{i-1}\epsilon N$, which is significantly smaller than $|\Gamma(v)| \geq \beta_2 \cdot 2^i\epsilon N$. So we need a different method to argue that being in $L$ is inconsistent with having degree approximately $|\Gamma(v)|$. Indeed, the source of trouble is that for two different thresholds $\beta' > \beta''$ it may be the case that $v \notin L_i$ holds because $|\Gamma(v)| \geq \beta'' \cdot 2^i\epsilon N$, whereas $v' \in L_j$ holds because $|\Gamma(v')| < \beta' \cdot 2^j\epsilon N$. Here is where the intermediate thresholds (and the different $\phi_k$) come into play: we shall show that whenever the foregoing happens it holds that $\beta'$ is very close to $\beta''$ (rather than $\beta' > 2\beta''$, which would have not given anything). Specifically, we shall show that if $\phi_k(v)$ holds then $\phi_{k-1}(v')$ must hold for almost all $v' \in \Gamma(u)$. Thus, if $v \notin L_i$ due to $|\Gamma(v)| \geq 0.9^k\beta_1 \cdot 2^i\epsilon N$ (and $\phi_k(v)$ holds), then $v' \in L_j$ implies that $|\Gamma(v')| < 0.9^{k-1}\beta_1 \cdot 2^j\epsilon N$, which yields the desired contradiction. Details follow.

Using arguments as in the previous two cases, we first establish that at least a 0.99 fraction of the vertices in $\Gamma(u)$ have degree at most $(1+\ell^{-2}) \cdot |\Gamma(v)|$ and have at least $(1-\ell^{-2}) \cdot |\Gamma(v)|$ neighbors in $\Gamma(v)$. (Here the argument relies on $\gamma \leq \beta_2/(500\ell^2)$ and $|\Gamma(u)| \geq |\Gamma(v)|/1.3 \geq \beta_2 \cdot 2^i\epsilon N/1.3$.) Let us denote this (large) subset of $\Gamma(u)$ by $Y$, and note that $v \in Y$. Similarly, one can show that at least $1-(200\ell)^{-1}$ of the vertices in $\Gamma(v)$ have degrees in the interval $[(1-(300\ell')^{-1}) \cdot |\Gamma(u)|, (1+(300\ell')^{-1}) \cdot |\Gamma(u)|]$, which we denote in short by $[(1 \pm (300\ell')^{-1}) \cdot |\Gamma(u)|]$. Hence, for every $v' \in Y$, it holds that $|\Gamma(v')|$ is in the interval $(1 \pm (300\ell')^{-1}) \cdot |\Gamma(v)|$, whereas at least $\frac{1-(200\ell)^{-1}}{1+\ell^{-2}} > 1-(100\ell)^{-1}$ of its neighbors (i.e., the vertices in $\Gamma(v')$) have degrees in the interval $[(1 \pm (300\ell')^{-1}) \cdot |\Gamma(u)|]$. Denoting (for every $v' \in Y$),

$$\rho(v') \stackrel{\text{def}}{=} \max_{S \subseteq \Gamma(v') \text{ s.t. } |S| = |\Gamma(v')|/100\ell} \left\{ \min_{u' \in S} \left\{ \frac{|\Gamma(u')|}{|\Gamma(v')|} \right\} \right\} \tag{112}$$

we infer that for every $v' \in Y$ (including $v$) it holds that $\rho(v') = \frac{(1 \pm (300\ell')^{-1}) \cdot |\Gamma(u)|}{(1 \pm (300\ell')^{-1}) \cdot |\Gamma(v)|} = (1 \pm (100\ell')^{-1}) \cdot \frac{|\Gamma(u)|}{|\Gamma(v)|}$. It follows that $\rho(v') \geq \frac{1-(100\ell')^{-1}}{1+(100\ell')^{-1}} \cdot \rho(v) > (1-(30\ell')^{-1}) \cdot \rho(v)$.

[11] Specifically, here we have

$$\frac{36\ell^2 \cdot \gamma\epsilon \cdot N^2}{\beta_2 \cdot 2^i\epsilon \cdot N/16} = \frac{576\ell^2 \cdot \gamma}{\beta_2} \cdot 2^{-i}N < 0.1 \cdot 2^{-i}N,$$

where the last inequality uses $\gamma < \beta_2/(6000\ell^2)$.

Recall that $k(v') \in \{0, 1, \ldots, \ell'\}$ is the largest index $k$ such that $\phi_k(v')$ holds (where $\phi_0$ always holds). Indeed, $\rho(v) > 1.1 + \frac{k(v)}{10\ell'}$ and $|\Gamma(v)| \geq 0.9^{k(v)} \cdot \beta_1 \cdot 2^i\epsilon \cdot N$ (because $v \notin L_i$). Combining $\rho(v') > (1 - (30\ell')^{-1}) \cdot \rho(v)$ and $\rho(v) > 1.1 + \frac{k(v)}{10\ell'}$, it follows that for every $v' \in Y$ it holds that $\rho(v') > 1.1 + \frac{k(v)-1}{10\ell'}$, which implies $k(v') \geq k(v) - 1$. It follows that $Y \cap L = \emptyset$, because otherwise we obtain, for some $j \leq i-1$, a vertex $v' \in Y \cap L_j$ such that $|\Gamma(v')| < 0.9^{k(v')} \cdot \beta_1 \cdot 2^j\epsilon \cdot N \leq 0.9^{k(v)-1} \cdot \beta_1 \cdot 2^{i-1}\epsilon \cdot N \leq |\Gamma(v)|/1.8$, which contradicts $|\Gamma(v')| \geq (1 - (300\ell')^{-1}) \cdot |\Gamma(v)| > \Gamma(v)|/1.8$. Recalling that $|Y| \geq 0.99 \cdot |\Gamma(u)|$, we conclude that $|\Gamma(u) \cap L| \leq 0.01|\Gamma(u)|$.

Combining the preliminary bound (of Eq. (109)) and the bounds of the foregoing three cases, we conclude that at most $(0.1 + 0.16 + 0.1 + 0.1) \cdot 2^{-i}N < 0.5 \cdot 2^{-i}N$ vertices satisfy Conditions (a) and (b) but violate Condition (c).

Recall that $R_i$ only contains vertices that satisfy Condition (b) but violate either Condition (a) or Condition (c). The number of the former was upper-bounded by $2^{-i}N/4$, whereas the number of the latter was just upper-bounded by $0.5 \cdot 2^{-i}N$. Thus, $|R_i| \leq (0.25 + 0.5) \cdot 2^{-i} \cdot N$, and Item 4 follows. This completes the proof of the current claim. ■

Completing the reconstruction and its analysis. The foregoing construction leaves "unassigned" the vertices in $R_\ell$ as well as some of the vertices in $L_1, \ldots, L_\ell$. (Note that some vertices in $\bigcup_{i=1}^{\ell-1} L_i$ may be placed in bi-cliques constructed in later iterations, but there is no guarantee that this actually happens.) We assign each of these remaining vertices to a two-vertex bi-clique (i.e., an isolated pair of vertices connected by an edge). Ignoring the number of edges used in these bi-cliques (which is negligible), the number of violations caused by this assignment equals the number of edges with both endpoints in $R' \stackrel{\text{def}}{=} R_\ell \cup (\bigcup_{i=1}^{\ell} L_i)$, because edges with a single endpoint in $R'$ were already accounted for in Item 3 of Claim 5.2.4. Nevertheless, we upper-bound the number of violations by the total number of edges incident to $R'$, which in turn is upper-bounded by

$$\sum_{v \in R_\ell \cup (\bigcup_{i \in [\ell]} L_i)} |\Gamma(v)| \leq |R_\ell| \cdot N + \sum_{i=1}^{\ell} \sum_{v \in L_i} |\Gamma(v)| \tag{113}$$

$$\leq \frac{\epsilon N}{4} \cdot N + \sum_{i=1}^{\ell} 2^{-(i-1)}N \cdot \beta_1 2^i \epsilon N \tag{114}$$

$$= \frac{\epsilon}{4} \cdot N^2 + 2\ell \cdot \beta_1 \cdot \epsilon N^2. \tag{115}$$

By the foregoing setting of $\beta_1$ (i.e., $\beta_1 \leq 1/4\ell$), it follows that the number of these edges is smaller than $\epsilon N^2/2$. Combining this with the bounds on the number of violating edges (or non-edges) as provided by Claim 5.2.4, the lemma follows. ■

## 5.2 Non-Adaptive Lower-Bound for Bi-Clique Collection

In this section we establish Part 2 of Theorem 1.2 by adapting the proof presented in Section 4.1. Specifically, for every value of $\epsilon > 0$, we consider two different classes of graphs, one consisting of graphs in $\mathcal{BCC}$ and the other consisting of graphs that are $\epsilon$-far from $\mathcal{BCC}$, and show that a non-adaptive algorithm of query complexity $o(\epsilon^{-3/2})$ cannot distinguish between graphs selected at random in these classes.

### 5.2.1 The two sets

The first class, denoted $\mathcal{BCC}_\epsilon$, contains all $N$-vertex graphs such that each graph consists of $(16\epsilon)^{-1}$ bi-cliques, and each bi-clique has $8\epsilon \cdot N$ vertices on each side. It will be instructive to partition these $(16\epsilon)^{-1}$ bi-cliques into $(32\epsilon)^{-1}$ pairs (each consisting of two bi-cliques), and view each of these bi-cliques as a super-cycle of length four with $4\epsilon \cdot N$ vertices in each of its four independent sets. The second class, denoted $\mathcal{SC}_8\mathcal{C}_\epsilon$, contains all $N$-vertex graphs such that each graph consists of $(32\epsilon)^{-1}$ super-cycles of length 8, and each of these super-cycles has $4\epsilon \cdot N$ vertices in each of its eight independent sets. For an illustration, see Figure 8. Indeed, $\mathcal{BCC}_\epsilon \subseteq \mathcal{BCC}$, whereas, as we show next, each graph in $\mathcal{SC}_8\mathcal{C}_\epsilon$ is $\epsilon$-far from $\mathcal{BCC}$. Note that *both classes contain only bipartite graphs.*

**Claim 5.3** *Every graph in $\mathcal{SC}_8\mathcal{C}_\epsilon$ is $\epsilon$-far from $\mathcal{BCC}$.*

**Proof:** Let $G = ([N], E)$ be a graph in $\mathcal{SC}_8\mathcal{C}_\epsilon$, let $(V_j^1, \ldots, V_j^8)$ be the eights sets of vertices in its $j^{\text{th}}$ super-cycle, and let $V_j = \bigcup_{s=1}^8 V_j^s$. For any partition $\mathcal{P} = ((X_1^1, X_1^2) \ldots, (X_\ell^1, X_\ell^2))$ into "potential bicliques", we let $\Delta_G(\mathcal{P})$ denote the number of edge modifications that are required in order to convert the pairs of sets $(X_i^1, X_i^2), \ldots, (X_\ell^1, X_\ell^2)$ into a collection of bicliques (with no edges between the bicliques). Then,

$$\Delta_G(\mathcal{P}) = \sum_{i=1}^{\ell} \left( |E(X_i^1, X_i^1)| + |E(X_i^2, X_i^2)| + |\overline{E}(X_i^1, X_i^2)| \right) + \sum_{i<i'} |E(X_i, X_{i'})| \,, \qquad (116)$$

where $X_i = X_i^1 \cup X_i^2$ and $\overline{E}(X_i^1, X_i^2)$ denotes the set of pairs of vertices in $X_i^1 \times X_i^2$ that do not have an edge between them. Thus, the distance between $G$ and $\mathcal{BCC}$ is $N^{-2}$ times the minimum, taken over all partitions $\mathcal{P}$, of $\Delta_G(\mathcal{P})$. We need to show that $\Delta_G(\mathcal{P}) > \epsilon N^2$, for every partition $\mathcal{P}$.

Similarly to the proof of Claim 4.1, we first observe that, without loss of generality, we may assume that each set $X_i$ intersects at most one $V_j$. This is true since otherwise, by refining the partition (i.e., replacing each $(X_i^1, X_i^2)$ with the collection of all nonempty $(X_i^1 \cap V_j, X_i^2 \cap V_j)$), the value of $\Delta_G(\cdot)$ can only decrease (because there are no edges between the different $V_j$'s).

Our next observation is that we may assume, without loss of generality, that each $V_j^s$ intersects at most one $X_i$ (i.e., one pair $(X_i^1, X_i^2)$). This is true because (for every $j \in [(32\epsilon)^{-1}]$, $s \in [8]$, $i \in [\ell]$, and $b \in \{1,2\}$) the contribution of each vertex $v \in V_j^s \cap X_i^b$ to Eq. (116) comes only from pairs $(v, u)$ such that either $u \notin V_j^s$ or $u \in V_j^s \cap X_i^{3-b}$. In particular, this contribution does not depend on $|V_j^s \cap X_i^b|$ nor on $|V_j^s \cap X_{i'}^{b'}|$ for any $i \neq i'$ and $b, b' \in \{1,2\}$. Therefore, if $V_j^s$ contains vertices of both $X_i^b$ and $X_{i'}^{b'}$ for some $i \neq i'$ and $b, b' \in \{1,2\}$, then it is possible to either move all $V_j^s \cap X_{i'}^{b'}$ to $V_j^s \cap X_i^b$ or the other way around without increasing $\Delta_G(\mathcal{P})$ (and possibly even decreasing it).

Having concluded that each $V_j^s$ is contained in some $X_i$, we observe that either $X_i^1 \cap V_j^s = \emptyset$ or $X_i^2 \cap V_j^s = \emptyset$. This holds because, using the same reasoning as above, if both $V_j^s \cap X_i^1 \neq \emptyset$ and $V_j^s \cap X_i^2 \neq \emptyset$ then by combining both sets into either $V_j^s \cap X_i^1$ or $V_j^s \cap X_i^2$ we only decrease $\Delta_G(\mathcal{P})$.

We have shown that, for every $i \in [\ell]$ and $b \in \{1,2\}$, there exists $j \in [(32\epsilon)^{-1}]$ and $S \subseteq [8]$ such that $X_i^b = \bigcup_{s \in S} V_j^s$. Thus, we may think of assigning pairs of the form $(i, b)$ to the eight slots on the cycle (i.e., $V_j^1, ..., V_j^8$), and we note that assigning $(i, b)$ and $(i', b')$ to (cyclically) adjacent slots incurs a cost of $K^2$ if and only if $(i', b') \neq (i, 3-b)$. In addition, assigning $(i, b)$ and $(i, 3-b)$ to non-adjacent slots also incurs a cost of $K^2$. Noting that it is impossible to assign these pairs at a cost of less than $3K^2$, it follows that the assignment to each $V^j$ contributed to $\Delta_G(\mathcal{P})$ at least $3K^2 = 48\epsilon^2 N^2$ violating vertex pairs. Combining the contribution of all $j \in [(32\epsilon)^{-1}]$, the claim follows. ■

### 5.2.2 The indistinguishability result

In order to motivate the claim that a non-adaptive algorithm of query complexity $o(\epsilon^{-3/2})$ cannot distinguish between graphs selected at random in these classes, consider the algorithm that selects $o(\epsilon^{-3/4})$ vertices and inspects the induced subgraph. Consider the partition of a graph in $\mathcal{SC}_8\mathcal{C}_\epsilon$ into $(32\epsilon)^{-1}$ pairs of bi-cliques (equiv., super-cycles of length 4), and correspondingly the partition of a graph in $\mathcal{SC}_8\mathcal{C}_\epsilon$ into $(32\epsilon)^{-1}$ super-cycles of length 8. Then, the probability that a sample of $o(\epsilon^{-3/4})$ vertices contains at least four vertices that reside in the same part (of $32\epsilon \cdot N$ vertices) is $o(\epsilon^{-3/4})^4 \cdot (32\epsilon)^3 = o(1)$. On the other hand, one may show that if this event does not occur, then the answers obtained from both graphs are indistinguishable. As will be shown below, this intuition extends to an arbitrary non-adaptive algorithm.

As in Section 4.1, it suffices to consider deterministic algorithms. We shall show that, for every set of $o(\epsilon^{-3/2})$ queries, the answers provided by a randomly selected element of $\mathcal{BCC}_\epsilon$ are statistically close to the answers provided by a randomly selected element of $\mathcal{SC}_8\mathcal{C}_\epsilon$. As in Section 4.1, for an $N$-vertex graph $G$ and a query $(u, v)$, we denote the corresponding answer by $\text{ans}_G(u, v)$.

**Lemma 5.4** *Let $G_1$ and $G_2$ be random $N$-vertex graphs uniformly distributed in $\mathcal{BCC}_\epsilon$ and $\mathcal{SC}_8\mathcal{C}_\epsilon$, respectively. Then, for every sequence $(v_1, v_2), \ldots, (v_{2q-1}, v_{2q}) \in [N] \times [N]$, where the $v_i$'s are not necessarily distinct, it holds that the statistical difference between* $\text{ans}_{G_1}(v_1, v_2), \ldots, \text{ans}_{G_1}(v_{2q-1}, v_{2q})$ *and* $\text{ans}_{G_2}(v_1, v_2), \ldots, \text{ans}_{G_2}(v_{2q-1}, v_{2q})$ *is $O(q^2\epsilon^3)$.*

Part 2 of Theorem 1.2 follows.

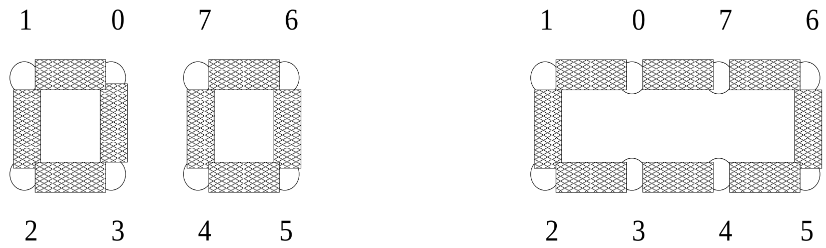


Figure 8: A single part, consisting of eight independent sets in $\mathcal{BCC}_\epsilon$ and $\mathcal{SC}_8\mathcal{C}_\epsilon$ (that is, either two bicliques, viewed as two super-cycles of length 4, or a single super-cycle of length 8).

**Proof:** We adapt the proof of Lemma 4.2. Here, we consider a 1-1 correspondence, denoted $\phi$, between the vertices of an $N$-vertex graph in $\mathcal{BCC}_\epsilon \cup \mathcal{SC}_8\mathcal{C}_\epsilon$ and triples in $[(32\epsilon)^{-1}] \times \{0, 1, \ldots, 7\} \times [4\epsilon \cdot N]$. Specifically, $\phi(v) = (i, j, w)$ indicates that $v$ resides in the $(j + 1)^{\text{st}}$ independent set of the $i^{\text{th}}$ part of the graph, and it is vertex number $w$ in this set. Recall that in the case of a graph in $\mathcal{BCC}_\epsilon$ the eight independent sets are arranged in two super-cycles (each of length 4), whereas in the case of a graph in $\mathcal{SC}_8\mathcal{C}_\epsilon$ the eight independent sets are arranged in a single super-cycle of length 8. (See Figure 8.) Consequently, the answers provided by uniformly distributed $G_1 \in \mathcal{BCC}_\epsilon$ and $G_2 \in \mathcal{SC}_8\mathcal{C}_\epsilon$ can be emulated by the following two corresponding random processes.

1. The process $A_1$ selects uniformly a bijection $\phi : [N] \to [(32\epsilon)^{-1}] \times \{0, 1, \ldots, 7\} \times [4\epsilon \cdot N]$ and answers each query $(u, v) \in [N] \times [N]$ by 1 if and only if for $\phi(u) = (i_1, j_1, w_1)$ and $\phi(v) = (i_2, j_2, w_2)$ it holds that both $i_1 = i_2$ and $j_1 = (j_2 \pm 1 \bmod 4) + \lfloor j_2/4 \rfloor \cdot 4$.

2. The process $A_2$ selects uniformly a bijection $\phi : [N] \to [(32\epsilon)^{-1}] \times \{0, 1, \ldots, 7\} \times [4\epsilon \cdot N]$ and answers each query $(u, v) \in [N] \times [N]$ by 1 if and only if for $\phi(u) = (i_1, j_1, w_1)$ and $\phi(v) = (i_2, j_2, w_2)$ it holds that both $i_1 = i_2$ and $j_1 = j_2 \pm 1 \bmod 8$.

Let us denote by $\phi'(v)$ (resp., $\phi''(v)$ and $\phi'''(v)$) the first (resp., second and third) coordinates of $\phi(v)$; that is, $\phi(v) = (\phi'(v), \phi''(v), \phi'''(v))$. Then, both processes answer the query $(u, v)$ with 0 if $\phi'(u) \neq \phi'(v)$, and the difference between the processes is confined to the case that $\phi'(u) = \phi'(v)$. Specifically, conditioned on $\phi'(u) = \phi'(v)$, it holds that $A_1(u, v) = 1$ if and only if $\phi''(u) = (\phi''(v) \pm 1 \bmod 4) + \lfloor \phi''(v)/4 \rfloor \cdot 4$, whereas $A_2(u, v) = 1$ if and only if $\phi''(u) = \phi''(v) \pm 1 \bmod 8$. However, since the (random) value of $\phi''$ is not present at the answer, the foregoing difference may go unnoticed. These considerations apply to a single query, but things may change in case of several queries. In general, the event that allows distinguishing the two processes is a simple cycle of at least four vertices that have the same $\phi'$ value. Minor differences may also be due to equal $\phi'''$ values, and so we also consider these in our "bad" event.

**Definition 5.4.1** *We say that $\phi$ is* bad (w.r.t. the sequence $(v_1, v_2), \ldots, (v_{2q-1}, v_{2q}) \in [N] \times [N]$), *if any of the following two conditions hold:*

1. *For some $i \in [(32\epsilon)^{-1}]$, the subgraph $Q_i = (V_i, E_i)$, where $V_i = \{v_k : k \in [2q] \wedge \phi'(v) = i\}$ and $E_i = \{\{v_{2k-1}, v_{2k}\} : v_{2k-1}, v_{2k} \in V_i\}$, contains a simple cycle of length at least four.*

2. *There exists $i \neq j \in [2q]$ such that $\phi'''(v_i) = \phi'''(v_j)$.*

Indeed, the query sequence $(v_1, v_2), \ldots, (v_{2q-1}, v_{2q})$ will be fixed throughout the rest of the proof, and so we shall omit it from our terminology.

**Claim 5.4.2** *The probability that a uniformly distributed bijection $\phi$ is bad is upper bounded by*

$$O(q^2\epsilon^3) + \frac{q^2}{16\epsilon N} .$$

Proof: We start by upper-bounding the probability that the second event in Definition 5.4.1 holds. We have $\binom{2q}{2}$ sub-events, and each holds with probability $1/(32\epsilon \cdot N)$. As for the first event, for every $t \geq 4$, we upper-bound the probability that some $Q_i$ contains a simple cycle of length $t$. As in the proof of Claim 4.2.2, we observe that the query graph contains at most $(2q)^{t/2}$ cycles of length $t$ (cf. [A81, Thm. 3]), whereas the probability that a specific simple $t$-cycle is contained in some $Q_i$ is $(32\epsilon)^{t-1}$. Thus, the probability of the first event is upper-bounded by

$$\sum_{t \geq 4} (2q)^{t/2} \cdot (32\epsilon)^{t-1} < \sum_{t \geq 4} \left(\sqrt{2q} \cdot 32 \cdot \epsilon^{(t-1)/t}\right)^t < \sum_{t \geq 4} \left(50\sqrt{q} \cdot \epsilon^{3/4}\right)^t ,$$

which is upper-bounded by $2 \cdot (50\sqrt{q} \cdot \epsilon^{3/4})^4 = O(q^2\epsilon^3)$, provided that $50\sqrt{q} \cdot \epsilon^{3/4} < 1/2$ (and the claim holds trivially otherwise). ■

**Claim 5.4.3** *Conditioned on the bijection $\phi$ not being bad, the sequences $(A_1(v_1, v_2), \ldots, A_1(v_{2q-1}, v_{2q}))$ and $(A_2(v_1, v_2), \ldots, A_2(v_{2q-1}, v_{2q}))$ are identically distributed.*

Proof: Noting that Definition 5.4.1 only refers to $\phi'$ and $\phi'''$, we fix any choice of $\phi'$ and $\phi'''$ that yields a good $\phi$ and consider the residual random choice of $\phi''$. Referring to the foregoing subgraphs $Q_i$'s, recall that pairs with endpoints in different $Q_i$'s are answered by 0 in both processes. Note that (by the second condition in Definition 5.4.1) the hypothesis implies that $\phi'''$ assigns different values to the different vertices in $\{v_k : k \in [2q]\}$, and it follows that $\phi''$ assigns these vertices values that are uniformly and independently distributed in $\{0, 1, \ldots, 7\}$. Now, using the first condition in Definition 5.4.1, the hypothesis implies that the only simple cycles appearing in $Q_i = (V_i, E_i)$

have length three. We shall show that this implies that (in each of the two processes) the answer assigned to each edge in $Q_i$ is independent of the answer given to other edges of $Q_i$.

We first note that, in each of the two processes, every query $(v_{2k-1}, v_{2k})$ such that $\phi''(v_{2k-1}) \equiv \phi''(v_{2k-1}) \pmod 2$ is answered negatively (i.e., in such a case, $A_1(v_{2k-1}, v_{2k}) = A_2(v_{2k-1}, v_{2k}) = 0$). Thus, fixing any (random) values of $(\phi''(v_k) \bmod 2 : k \in [2q])$, we may omit from $Q_i = (V_i, E_i)$ all edges that connect vertices that have the same value of $\phi'' \pmod 2$, because the answers to these queries are already determined (as 0, in each of the two processes). This omission eliminates (from $Q_i$) all cycles of length three, which are the only simple cycles in the original $Q_i$, and thus each modified $Q_i$ is a forest. We can now proceed analogously to the proof of Claim 4.2.3, although things are slightly more complex here. Specifically, we consider the residual random values of $\phi''$ (conditioned on $\phi'' \bmod 2$); that is, we augments the fixed values of $\phi'' \bmod 2$ with the random values of $\lfloor \phi''/2 \rfloor$, which are uniformly distributed in $\{0, 1, 2, 3\}$. We view these random selections as taking place in an order determined by some fixed traversal of each tree (of the aforementioned forest), and note that at each step (and in each of the processes) the new random value (uniformly distributed in $\{0, 1, 2, 3\}$) yields answer 1 (to the corresponding query) with probability 1/2.

1. In the case of $A_1$, the query/edge $(u, v) \in E_i$ (which satisfies $\phi'(u) = i = \phi'(v)$ and $\phi''(u) \equiv \phi''(v) + 1 \pmod 2$) is answered 1 if and only if $\phi''(u) = (\phi''(v) \pm 1 \bmod 4) + \lfloor \phi''(v)/4 \rfloor \cdot 4$ holds (which means that $\lfloor \phi''(u)/4 \rfloor = \lfloor \phi''(v)/4 \rfloor$). Thus, $A_1(u, v) = 1$ with probability 1/2.

2. In the case of $A_2$, the query/edge $(u, v) \in E_i$ (which satisfies $\phi'(u) = i = \phi'(v)$ and $\phi''(u) \equiv \phi''(v) + 1 \pmod 2$) is answered 1 if and only if $\phi''(u) = \phi''(v) \pm 1 \bmod 8$ holds. Thus, $A_2(u, v) = 1$ with probability 2/4.

Thus, in each of the two processes, each query is answered by the value 1 with probability exactly 1/2, independently of the answers to all other queries. The claim follows. ■

Combining Claims 5.4.2 and 5.4.3, it follows that the statistical distance between the sequences $(A_1(v_1, v_2), \ldots, A_1(v_{2q-1}, v_{2q}))$ and $(A_2(v_1, v_2), \ldots, A_2(v_{2q-1}, v_{2q}))$ is at most $O(q^2\epsilon^3 + q^2(\epsilon N)^{-1})$, and the lemma follows for sufficiently large $N$. ■

## 5.3 Non-Adaptive Lower-Bound for Super-Cycle Collection

In this section we establish a lower bound on the non-adaptive query complexity of testing Super-Cycle Collections. We do so by generalizing the ideas presented in Section 5.2.

Specifically, fixing any $t \geq 4$, for every value of $\epsilon > 0$, we consider two different classes of graphs, one consisting of graphs in $\mathcal{S}\mathtt{C}_t\mathcal{C}$ and the other consisting of graphs that are $\epsilon$-far from $\mathcal{S}\mathtt{C}_t\mathcal{C}$, and show that a non-adaptive algorithm of query complexity $o(\epsilon^{-(2t-2)/t})$ cannot distinguish between graphs selected at random in these classes.

### 5.3.1 The two sets

The first class, denoted $\mathcal{S}\mathtt{C}_t\mathcal{C}_\epsilon$, contains all $N$-vertex graphs such that each graph consists of $(t^2\epsilon)^{-1}$ super-cycles of length $t$, and each super-cycle has $t\epsilon \cdot N$ vertices in each of its $t$ independent sets. It will be instructive to partition these $(t^2\epsilon)^{-1}$ super-cycles into $(2t^2\epsilon)^{-1}$ pairs. The second class, denoted $\mathcal{S}\mathtt{C}_{2t}\mathcal{C}_\epsilon$, contains all $N$-vertex graphs such that each graph consists of $(2t^2\epsilon)^{-1}$ super-cycles of length $2t$, and each super-cycle has $t\epsilon \cdot N$ vertices in each of its $2t$ independent sets. For an illustration, see Figure 9. Indeed, $\mathcal{S}\mathtt{C}_t\mathcal{C}_\epsilon \subseteq \mathcal{S}\mathtt{C}_t\mathcal{C}$, whereas, as we show next, each graph in $\mathcal{S}\mathtt{C}_{2t}\mathcal{C}_\epsilon$ is $\epsilon$-far from $\mathcal{S}\mathtt{C}_t\mathcal{C}$.

**Claim 5.5** *Every graph in $\mathcal{S}\mathtt{C}_{2t}\mathcal{C}_\epsilon$ is $\epsilon$-far from $\mathcal{S}\mathtt{C}_t\mathcal{C}$.*

**Proof:** Let $G = ([N], E)$ be a graph in $\mathcal{S}\mathtt{C}_{2t}\mathcal{C}_\epsilon$, let $(V_j^1, \ldots, V_j^{2t})$ be the $2t$ sets of vertices in its $j^{\text{th}}$ super-cycle, and let $V_j = \bigcup_{s=1}^{2t} V_j^s$. For any partition $\mathcal{P} = ((X_1^1, ..., X_1^t) \ldots, (X_\ell^1, ..., X_\ell^t))$ into "potential $t$-super-cycles", we let $\Delta_G(\mathcal{P})$ denote the number of edge modifications that are required in order to convert $\mathcal{P}$ into a collection of $t$-super-cycles (with no edges between the $t$-super-cycles).

Similarly to the proof of Claim 5.3, we observe that, without loss of generality, we may assume that (1) each set $X_i = \bigcup_r X_i^r$ intersects at most one $V_j$, and (2) each $V_j^s$ intersects at most one $X_i$. Furthermore, we may also assume, without loss of generality, that each $V_j^s$ intersects at most one $X_i^b$ (but in this case the argument does not necessarily decrease $\Delta_G(\mathcal{P})$, although it never increases it).

Thus, for every $i \in [\ell]$ and $b \in \{1, ..., t\}$, there exists $j \in [(2t^2\epsilon)^{-1}]$ and $S \subseteq [2t]$ such that $X_i^b = \bigcup_{s \in S} V_j^s$. Now, we may think of assigning pairs of the form $(i, b)$ to the $2t$ slots on the cycle (i.e., $V_j^1, ..., V_j^{2t}$), and we note that assigning $(i, b)$ and $(i', b')$ to (cyclically) adjacent slots incurs a cost of $K^2$ if and only if $(i', b') \notin \{(i, b-1), (i, b+1)\}$, where $K \stackrel{\text{def}}{=} t\epsilon N$ and addition is modulo $t$. In addition, assigning $(i, b)$ and $(i, b \pm 1)$ to non-adjacent slots also incurs a cost of $K^2$. Note that it is impossible to assign these pairs at a cost of less than $3K^2$, because such an assignment mandates having at most two adjacent pairs that are assigned different values of $i$, whereas a consecutive run of $t$ values of any $i$ contains either an adjacent pair that does not have the form $(i, b)$ and $(i, b \pm 1)$ or a non-adjacent pair of $(i, b)$ and $(i, b \pm 1)$. It follows that the assignment to each $V^j$ contributed to $\Delta_G(\mathcal{P})$ at least $3K^2 = 3t^2\epsilon^2N^2$ violating vertex pairs. Combining the contribution of all $j \in [(2t^2\epsilon)^{-1}]$, the claim follows. ■

### 5.3.2 The indistinguishability result

As in Section 5.2, we motivate the claim that a non-adaptive algorithm of query complexity $o(\epsilon^{-(2t-2)/t})$ cannot distinguish between graphs selected at random in these classes by considering a specific algorithm that inspects the subgraph induced by a random set of $o(\epsilon^{-(t-1)/t})$ vertices. The probability that a sample of $o(\epsilon^{-(t-1)/t})$ vertices contains at least $t$ vertices that reside in the same part (of $(2t^2\epsilon) \cdot N$ vertices) is $\binom{o(\epsilon^{-(t-1)/t})}{t} \cdot (2t^2\epsilon)^{t-1} = o(1)$, where the $o$-notation refers to a fixed value of $t$ and a varying value of $\epsilon > 0$. On the other hand, one may show that if this event does not occur, then the answers obtained from both graphs are indistinguishable. As will be shown below, this intuition extends to an arbitrary non-adaptive algorithm. Following the same conventions as in Section 5.2, it suffices to prove the following

**Lemma 5.6** (Lemma 5.4, generalized): *For every fixed $t \geq 4$, let $G_1$ and $G_2$ be random $N$-vertex graphs uniformly distributed in $\mathcal{S}\mathtt{C}_t\mathcal{C}_\epsilon$ and $\mathcal{S}\mathtt{C}_{2t}\mathcal{C}_\epsilon$, respectively. Then, for every sequence $(v_1, v_2), \ldots, (v_{2q-1}, v_{2q}) \in [N] \times [N]$, where the $v_i$'s are not necessarily distinct, it holds that the statistical difference between $\text{ans}_{G_1}(v_1, v_2), \ldots, \text{ans}_{G_1}(v_{2q-1}, v_{2q})$ and $\text{ans}_{G_2}(v_1, v_2), \ldots, \text{ans}_{G_2}(v_{2q-1}, v_{2q})$ is $O(q^{t/2}\epsilon^{t-1})$.*

Part 2 of Conjecture 1.3 follows. Indeed, Lemma 5.4 can be obtained as a special case of Lemma 5.6 by setting $t = 4$. The following proof is slightly different from the proof provided in Section 5.2.

**Proof:** We generalize the proof of Lemma 5.4. We consider a bijection, denoted $\phi$, between the vertices of an $N$-vertex graph in $\mathcal{S}\mathtt{C}_t\mathcal{C}_\epsilon \cup \mathcal{S}\mathtt{C}_{2t}\mathcal{C}_\epsilon$ and triples in $[(2t^2\epsilon)^{-1}] \times \{0, 1, \ldots, 2t-1\} \times [t\epsilon \cdot N]$. Specifically, $\phi(v) = (i, j, w)$ indicates that $v$ resides in the $(j+1)^{\text{st}}$ independent set of the $i^{\text{th}}$ part of the graph, and that it is vertex number $w$ in this set. Recall that in the case of a graph in $\mathcal{S}\mathtt{C}_t\mathcal{C}_\epsilon$ the

$2t$ independent sets in each part are arranged in two super-cycles (each of length $t$), whereas in the case of a graph in $\mathcal{S}\mathrm{C}_{2t}\mathcal{C}_\epsilon$ the $2t$ independent sets are arranged in a single super-cycle of length $2t$. Consequently, the answers provided by uniformly distributed $G_1 \in \mathcal{S}\mathrm{C}_t\mathcal{C}_\epsilon$ and $G_2 \in \mathcal{S}\mathrm{C}_{2t}\mathcal{C}_\epsilon$ can be emulated by the following two corresponding random processes.

1. The process $A_1$ selects uniformly a bijection $\phi : [N] \to [(2t^2\epsilon)^{-1}] \times \{0, 1, \ldots, 2t-1\} \times [t\epsilon \cdot N]$ and answers each query $(u, v) \in [N] \times [N]$ by 1 if and only if for $\phi(u) = (i_1, j_1, w_1)$ and $\phi(v) = (i_2, j_2, w_2)$ it holds that both $i_1 = i_2$ and $j_1 = (j_2 \pm 1 \bmod t) + \lfloor j_2/t \rfloor \cdot t$.

2. The process $A_2$ selects uniformly a bijection $\phi : [N] \to [(2t^2\epsilon)^{-1}] \times \{0, 1, \ldots, 2t-1\} \times [t\epsilon \cdot N]$ and answers each query $(u, v) \in [N] \times [N]$ by 1 if and only if for $\phi(u) = (i_1, j_1, w_1)$ and $\phi(v) = (i_2, j_2, w_2)$ it holds that both $i_1 = i_2$ and $j_1 = j_2 \pm 1 \bmod 2t$.

Again, let us denote by $\phi'(v)$ (resp., $\phi''(v)$ and $\phi'''(v)$) the first (resp., second and third) coordinates of $\phi(v)$; that is, $\phi(v) = (\phi'(v), \phi''(v), \phi'''(v))$. Then, both processes answer the query $(u, v)$ with 0 if $\phi'(u) \neq \phi'(v)$, and the difference between the processes is confined to the case that $\phi'(u) = \phi'(v)$. Specifically, conditioned on $\phi'(u) = \phi'(v)$, it holds that $A_1(u, v) = 1$ if and only if $\phi''(u) = (\phi''(v) \pm 1 \bmod t) + \lfloor \phi''(v)/t \rfloor \cdot t$, whereas $A_2(u, v) = 1$ if and only if $\phi''(u) = \phi''(v) \pm 1 \bmod 2t$. In general, the event that allows distinguishing the two processes is a simple cycle of at least $t$ vertices that have the same $\phi'$ value. Minor differences may also be due to equal $\phi'''$ values, and so we also consider these in our "bad" event.

**Definition 5.6.1** (Definition 5.4.1, generalized): *We say that $\phi$ is* bad (w.r.t. the sequence of queries $(v_1, v_2), \ldots, (v_{2q-1}, v_{2q}) \in [N] \times [N]$), *if any of the following two conditions hold:*

1. *For some $i \in [(2t^2\epsilon)^{-1}]$, the subgraph $Q_i = (V_i, E_i)$, where $V_i = \{v_k : k \in [2q] \wedge \phi'(v) = i\}$ and $E_i = \{\{v_{2k-1}, v_{2k}\} : v_{2k-1}, v_{2k} \in V_i\}$, contains a simple cycle of length at least $t$.*

2. *There exists $i \neq j \in [2q]$ such that $\phi'''(v_i) = \phi'''(v_j)$.*

The query sequence $(v_1, v_2), \ldots, (v_{2q-1}, v_{2q})$ will be fixed throughout the rest of the proof, and so we shall omit it from our terminology.

**Claim 5.6.2** (Claim 5.4.2, generalized): *The probability that a uniformly distributed bijection $\phi$ is bad is upper bounded by*

$$(3t)^{2t} \cdot q^{t/2} \cdot \epsilon^{t-1} + \frac{q^2}{t^2\epsilon N} .$$

Proof: We start by upper-bounding the probability that the second event in Definition 5.6.1 holds. We have $\binom{2q}{2}$ sub-events, and each holds with probability $1/(2t^2\epsilon \cdot N)$. As for the first event, for every $\ell \geq t$, we upper-bound the probability that some $Q_i$ contains a simple cycle of length $\ell$ by $(2q)^{\ell/2} \cdot (2t^2\epsilon)^{\ell-1}$ (once again using the fact that a subgraph with $2q$ edges contains at most $(2q)^{\ell/2}$ cycles of length $\ell$ (cf. [A81, Thm. 3])). Thus, the probability of the first event is upper-bounded by

$$\sum_{\ell \geq t} (2q)^{\ell/2} \cdot (2t^2\epsilon)^{\ell-1} \;<\; \sum_{\ell \geq t} \left(3t^2\sqrt{q} \cdot \epsilon^{(t-1)/t}\right)^{\ell} .$$

If $3t^2\sqrt{q} \cdot \epsilon^{(t-1)/t} < 1/2$, then this expression is upper bounded by $2 \cdot (3t^2\sqrt{q} \cdot \epsilon^{(t-1)/t})^t \leq (3t)^{2t} \cdot q^{t/2} \cdot \epsilon^{t-1}$. But if $3t^2\sqrt{q} \cdot \epsilon^{(t-1)/t} \geq 1/2$, so that $q^{t/2} \geq 6^{-t} \cdot t^{-2t} \cdot \epsilon^{-(t-1)}$, then $(3t)^{2t} \cdot q^{t/2} \cdot \epsilon^{t-1} > 1$, so that the claim hold trivially. ■

**Claim 5.6.3** (Claim 5.4.3, generalized): *Conditioned on the bijection $\phi$ not being bad, the sequences $(A_1(v_1,v_2),\ldots,A_1(v_{2q-1},v_{2q}))$ and $(A_2(v_1,v_2),\ldots,A_2(v_{2q-1},v_{2q}))$ are identically distributed.*

Proving this claim is the only difficulty in extending the proof of Lemma 5.4 to the current setting. Indeed, the following proof yields a slightly different proof of Claim 5.4.3.

Proof: Again, we fix any choice of $\phi'$ and $\phi'''$ that yields a good $\phi$, and consider the residual random choice of $\phi''(v_1),\ldots,\phi''(v_{2q})$, which (by the second hypothesis in Definition 5.6.1) are uniformly and independently distributed in $\{0,1,\ldots,2t-1\}$. Considering any of the aforementioned graphs $Q_i=(V_i,E_i)$, we note that this graph does not contain simple cycles of length greater than $t-1$.

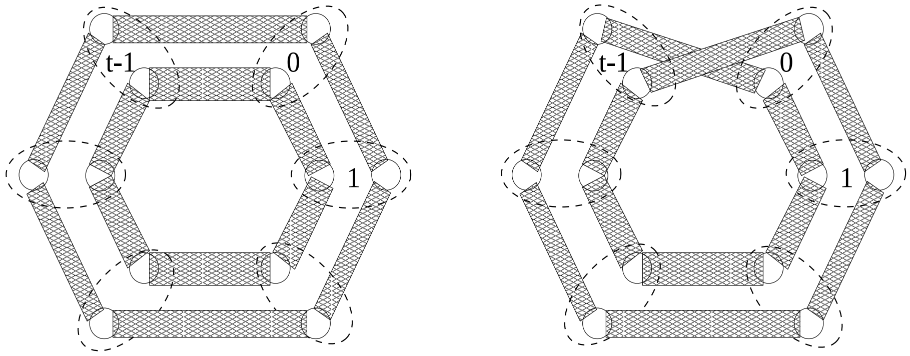


Figure 9: A single part, consisting of $2t$ independent sets, in $\mathcal{S}\mathrm{C}_t\mathcal{C}_\epsilon$ and $\mathcal{S}\mathrm{C}_{2t}\mathcal{C}_\epsilon$, respectively. The ellipses indicate the values of $\psi''$.

We now consider $\phi'':V_i\to\{0,1,\ldots,2t-1\}$ as being selected at random in two stages. In the first stage we assign each vertex a random value mod $t$, and in the second stage we assign each vertex a random bit representing its most significant bit; that is, for each vertex $v\in V_i$, we first determine (at random) the value $\phi''(v)$ mod $t$, which we denote by $\psi''(v)$, and next determine (at random) the bit $\lfloor\phi''(v)/t\rfloor$, which we denote by $\pi''(v)$. Thus, $\phi''(v)=\psi''(v)+\pi''(v)\cdot t$, and it will be instructive to depict the graphs as in Figure 9. Fixing an arbitrary setting of values for the first stage, we shall consider what may happen in the second stage.

For every fixed setting of $\psi''$, we consider the residual graph $Q_i'=(V_i,E_i')$, where $E_i'$ contains only the queries in $E_i$ that are still undetermined (given $\psi''$); that is, $(u,v)\in E_i$ is placed in $E_i'$ if and only if $\psi''(u)\equiv\psi''(v)\pm 1 \pmod t$, whereas all the other queries (or rather the answers to them) are already determined (as being answered by 0). We shall consider the connected components of $Q_i'$, and show that (conditioned on the foregoing setting of $\psi''$) the answers provided to the queries in $E_i'$ under $A_1$ are distributed identically to the answers provided under $A_2$. Specifically, for each possible sequence of answers, we shall show a 1-1 correspondence between the assignments of $\pi''$ that yield these answers under $A_1$ and the assignments of $\pi''$ that yield these answers under $A_2$. (Recall that $\phi''(v)=\psi''(v)+\pi''(v)\cdot t$.) That is, for each possible sequence of answers and each connected component of $Q_i'$, we shall show that the number of assignments of $\pi''$ that yield these answers under $A_j$ is independent of $j\in\{1,2\}$.

Let $C=(V_i'',E_i'')$ be an arbitrary connected component of $Q_i'=(V_i,E')$, and let $A'':E_i''\to\{0,1\}$ describe an arbitrary sequence of answers to the queries $E_i''$. Our aim is proving that the number of assignments of $\pi''$ that yield these answers under $A_j$ (i.e., satisfy $A_j(u,w)=A''(u,w)$ for every $(u,w)\in E_i''$) is independent of $j\in\{1,2\}$. Furthermore, we shall show that this number is either two or zero (when considering only the assignment of $\pi''$ to $V_i''$). Consider any spanning tree $T$ of $C$, rooted at an arbitrary vertex $v\in V_i''$. For each choice of $\sigma\in\{0,1\}$, we shall prove that there exists a unique assignment $\pi'' : V_i'' \to \{0,1\}$ such that $\pi''(v) = \sigma$ and $\pi''$ is consistent with $A''$ and $A_1$ (resp., $A_2$) on the edges of $T$. That is, the resulting $\pi''$ is such that the answers as mandated by $A''$ for the edges of $T$ fit the answers that $A_1$ (resp., $A_2$) provides with respect to $\phi'' = \psi'' + t \cdot \pi''$. As we shall see, these assignments might be inconsistent with the value of $A''$ on edges that do not belong to the spanning tree. However, we shall show that there is an inconsistency when fitting $A_1$ if and only if there is an inconsistency when fitting $A_2$. Details follow.

Fitting the process $A_1$: Recall that the value of $\pi''$ on the root of $T$ was set to $\sigma$. The value of $\pi''$ on all other vertices is set, by traversing the tree $T$, in the following manner. When traversing the tree edge $(u, w)$ from a vertex $u$ for which $\pi''(u)$ was already determined to a new $w$ (for which $\pi''(w)$ is still undetermined), we set $\pi''(w) \leftarrow \pi''(u)$ if $A''(u, w) = 1$ and $\pi''(w) \leftarrow 1 - \pi''(u)$ otherwise (i.e., if $A''(u, w) = 0$).

Note that this process determines the values of the bits $\pi''(w)$ for all $w \in V_i''$ such that the tree-neighbors $u$ and $w$ are assigned the same bit if and only if $A''(u, w) = 1$. This is indeed consistent with the definition of $A_1$. Furthermore, the setting of the values of $\pi''$ is uniquely determined by the requirement to be consistent with $A_1$.

Fitting the process $A_2$: We assign values exactly as in the case of fitting $A_1$, with a single exception that refers to the case that the tree-edge $(u, w) \in E_i''$ satisfies $\{\psi''(u), \psi''(w)\} = \{0, t-1\}$. In this case (where vertex $u$ has already been assigned a value), we set $\pi''(w) \leftarrow 1 - \pi''(u)$ if $A''(u, w) = 1$ and $\pi''(w) \leftarrow \pi''(u)$ otherwise (i.e., if $A''(u, w) = 0$).

That is, in this case (i.e., $\{\psi''(u), \psi''(w)\} = \{0, t-1\}$), the process determines the value of $\pi''(w)$ such that the tree-neighbors $u$ and $w$ are assigned the opposite bits if and only if $A''(u, w) = 1$.

As noted in the foregoing discussion, while each of the two assignments is consistent with $A''$ (and the corresponding $A_j$) on the edges of the spanning tree $T$, there may be inconsistencies with the edges of $E_i''$ that are not tree edges. It remains to show that there is an inconsistency with respect to the process $A_1$ if and only if there is an inconsistency with respect to the process $A_2$.

We shall say that an edge $(u, w) \in E_i''$ (e.g., an edge of the spanning tree $T$) is a crossing edge if $\{\psi''(u), \psi''(w)\} = \{0, t-1\}$. By definition of the two assignments, the only difference between them is caused when traversing a tree edge that is a crossing edge. For such an edge, the value of $\pi''$ is flipped when fitting the process $A_2$ if and only if it is *not* flipped when fitting the process $A_1$. Thus, for each $u \in V_i''$, the value assigned to $\pi''(u)$ when fitting $A_2$ is the XOR of the value assigned to $\pi''(u)$ when fitting $A_1$ and the *parity* of the number of crossing edges that belong to the tree path from (the root) $v$ to $u$.

Now, consider an edge $(u, w) \in E_i''$ that is not an edge in the spanning tree $T$. Consider the simple tree paths from the root $v$ to vertices $u$ and $w$, respectively, and let us denote their branching point by $v'$. Let $p_u$ (resp., $p_w$) be the path on the spanning tree $T$ leading from $v'$ to $u$ (resp., $w$), and $p_u'$ be the path from $v'$ to $u$ obtained by augmenting $p_w$ with the (non-tree) edge $(w, u)$. Then, the union of $p_u$ and $p_u'$ constitutes a simple cycle, which by the hypothesis has length smaller than $t$. As we shall show in the next paragraph, it follows that *the parity of the number of crossing edges on $p_u$ equals the parity of the number of crossing edges on $p_u'$*. In other words, the parity of the number of crossing edges on $p_u$ equals the parity of the number of crossing edges on $p_w$ if and only if $(u, w)$ is not a crossing edge. Assuming that $(u, w)$ is not a crossing edge, consider the value assigned to $\pi''(u)$ and $\pi''(w)$ when fitting $A_1$ (by following the paths from the root to $u$ and $w$, respectively). Then, $A''(u, w)$ is inconsistent with $\pi''(u)$ and $\pi''(w)$ as determined when fitting the process $A_1$ if and only if $A''(u, w)$ is inconsistent with $\pi''(u)$ and $\pi''(w)$ as determined when fitting the process $A_2$, because in both cases $\pi''(u) \oplus \pi''(w)$ is the same value (since the total number of crossing edges on $p_v$ and $p_w$ is even). A similar argument holds when $(u, w)$ is a crossing edge (since then $\pi''(u) \oplus \pi''(w)$ flips from $A_1$ to $A_2$), and the claim follows.

To verify the assertion regarding the parity of the number of crossing edges on $p_u$ and on $p'_u$, consider the values assigned by $\psi''$ to the vertices in the union of $p_u$ and $p'_u$. Since the union of $p_u$ and $p'_u$ is a cycle of length less than $t$, these values must belong to a proper subset, $S$, of $\{0, \ldots, t-1\}$. If this set does not contain $\{0, t-1\}$, then we are done (since neither of the paths may contain a crossing edge). Otherwise, for some $j$, it holds that $S$ is a subset of the union of $S_1 = \{j+1, \ldots, t-1\}$ and $S_2 = \{0, \ldots, j-1\}$. If $\psi''(v')$ and $\psi''(u)$ belong to the same $S_k$, then the parity of the number of crossing edges on both $p_u$ and $p'_u$ is even (since these paths can only move from one subset to the other via a crossing edge).[12] Similarly, if $\psi''(v')$ and $\psi''(u)$ do not belong to the same subset then the parity on each of these paths must be odd. ■

Combining Claims 5.6.2 and 5.6.3, the lemma follows. ■

## 5.4 A Candidate Adaptive Tester for Super-Cycle Collection

In this section we outline an adaptive $\widetilde{O}(\epsilon^{-1})$-query algorithm that we conjecture to be a tester for $\mathcal{S}\mathtt{C}_t\mathcal{C}$, where $t \geq 5$ is fixed. The algorithm is a generalization of Algorithm 5.1, and we focus on outlining the corresponding sub-test, denoted sub-test$_i(v)$.

Recall that in Algorithm 5.1 this sub-test consists, essentially, of finding an edge $(v, u)$ and checking the potential bi-clique induced by it (i.e., $\Gamma(u) \times \Gamma(v)$). In the current context we try to find a $t$-cycle $(v_0, v_1, \ldots, v_{t-1})$ such that $v_0 = v$ and for every $j \in \{0, \ldots, t-1\}$ it holds that $v_j \in \Gamma(v_{j-1 \bmod t}) \cap \Gamma(v_{j+1 \bmod t})$ and $\Gamma(v_{j-1 \bmod t}) \neq \Gamma(v_{j+1 \bmod t})$. Given such a candidate $t$-cycle $\overline{v}$, letting $I_j(\overline{v}) \stackrel{\text{def}}{=} \Gamma(v_{j-1 \bmod t}) \cap \Gamma(v_{j+1 \bmod t})$, we check that $I_j(\overline{v}) \times I_{j+1 \bmod t}(\overline{v})$ is a bi-clique, and that $\Gamma(v_j) = I_{j-1 \bmod t}(\overline{v}) \cup I_{j+1 \bmod t}(\overline{v})$. Each of these tasks is to be performed by making $\text{poly}(\log(1/\epsilon))/(2^i\epsilon)$ queries. The implementation of the various checks is similar to the implementation of similar checks performed in Algorithm 5.1, and so we focus on finding the aforementioned $t$-cycle.

Starting with $v_0 \stackrel{\text{def}}{=} v$, we obtain $v_1 \in \Gamma(v)$ just as ($u$ was obtained) in Algorithm 5.1. In fact, we may obtain $v_{t-1} \in \Gamma(v)$ in the same way, except that we need to verify that the latter vertex is actually in a different independent set than $v_1$. This is done by checking that $\Gamma(v_{t-1})$ is different from $\Gamma(v_1)$, where any $w$ in the symmetric difference of $\Gamma(v_1)$ and $\Gamma(v_{t-1})$ can serve as a witness. (Indeed, $w \in \Gamma(v_1) \setminus \Gamma(v_{t-1})$ can be used as $v_2$.) Similarly, when holding a partial path $(v_{t-j}, \ldots, v_0, \ldots, v_k)$, we seek a vertex $v_{k+1}$ (resp., $v_{t-(j+1)}$) such that $\Gamma(v_{k+1})$ and $\Gamma(v_{k-1})$ (resp., $\Gamma(v_{t-(j+1)})$ and $\Gamma(v_{t-(j-1)})$) are different. When the path reaches length $t-1$ (i.e., holds $t$ vertices), we treat it as a candidate $t$-cycle.

We note that, as in the case of Algorithm 5.1, it may happen that the foregoing algorithm fails to find a $t$-cycle, $(v_0, \ldots, v_{t-1})$. In this case, the algorithm performs only a subset of the checks mentioned above. Specifically, suppose that the algorithm failed to extend the partial path $\overline{v} \stackrel{\text{def}}{=} (v_{t-j}, \ldots, v_0, \ldots, v_k)$ any further. Then, for intermediate vertices, the checks are as before, but for the extremes we should proceed with more care. For example, assuming the path contains at least four vertices, we let $I_{t-j}(\overline{v}) \stackrel{\text{def}}{=} \Gamma(v_{t-j+1 \bmod t}) \setminus I_{t-j+2 \bmod t}(\overline{v})$.

Clearly, the foregoing algorithm always accepts any graph in $\mathcal{S}\mathtt{C}_t\mathcal{C}$, and we conjecture that it (or possibly a slight variant of it) rejects with high probability graphs that are $\epsilon$-far from $\mathcal{S}\mathtt{C}_t\mathcal{C}$.

[12]Note that the $\psi''$-values of intermediate vertices along any path must be "adjacent" modulo $t$, and so moving between $\{j+1, \ldots, t-1\}$ and $\{0, \ldots, j-1\}$ is only possible via $(t-1, 0)$.

In the next theorem we prove that a simplified version of this algorithm can distinguish with high probability between graphs in $\mathcal{S}\mathtt{C}_t\mathcal{C}$ and graphs in $\mathcal{S}\mathtt{C}_{2t}\mathcal{C}' \stackrel{\rm def}{=} \bigcup_{i\geq 5}\mathcal{S}\mathtt{C}_{2t}\mathcal{C}_{2^{-i}}$ that are $\epsilon$-far from $\mathcal{S}\mathtt{C}_t\mathcal{C}$. We refer to this promise problem as $\Pi_t$.

**Theorem 5.7** (an almost-quadratic complexity gap for promise problems): *For every positive integer $t\geq 3$, the promise problem $\Pi_t$ satisfies the following:*

1. *There exists an adaptive tester of query complexity $O(\epsilon^{-1})$ for $\Pi_t$. Furthermore, this tester has one-sided error and runs in time $O(\epsilon^{-1})$.*
2. *Any non-adaptive tester for $\Pi_t$ must have query complexity $\Omega(\epsilon^{-2+(2/t)})$.*
3. *There exists a non-adaptive tester of query complexity $O(\epsilon^{-2+(2/t)})$ for $\Pi_t$. Furthermore, this tester has one-sided error and runs in time $O(\epsilon^{-2+(2/t)})$.*

Indeed, in light of Theorems 1.1 and 1.2, the cases of $t\in\{3,4\}$ are of little interest, but there are given here for the sake of uniformity (and since Theorem 1.2 lacks Part 3). We also stress that the hidden constants in the O-notation may depend on (the constant) $t$.

**Proof:** As noted above, Part 2 follows from Lemma 5.6 (which actually holds also in the case of $t=3$). Specifically, for $\ell=\log_2(1/\epsilon)$, Lemma 5.6 asserts that an algorithm of query complexity $q\stackrel{\rm def}{=}o(\epsilon^{-2+(2/t)})$ cannot distinguish between graphs that are uniformly distributed in $\mathcal{S}\mathtt{C}_t\mathcal{C}_{2^{-\ell}}$ and graphs that are uniformly distributed in $\mathcal{S}\mathtt{C}_{2t}\mathcal{C}_{2^{-\ell}}$, since its distinguishing gap is $O(q^{t/2}\epsilon^{t-1})=o(1)$. Part 2 follows since $\mathcal{S}\mathtt{C}_t\mathcal{C}_{2^{-\ell}}\subset\mathcal{S}\mathtt{C}_t\mathcal{C}$, whereas all graphs in $\mathcal{S}\mathtt{C}_{2t}\mathcal{C}_{2^{-\ell}}$ are both in $\mathcal{S}\mathtt{C}_{2t}\mathcal{C}'$ and $\epsilon$-far from $\mathcal{S}\mathtt{C}_t\mathcal{C}$.

Turning to Part 1, as noted above, this part can be proved by using the algorithm outlined above. Actually, for the current task of testing the promise problem $\Pi_t$, a degenerate version of the foregoing algorithm will do, and we detail and analyze such a version next. The key observation underlying this simplified version is that in the current context the input is guaranteed (by the promise problem formulation, cf. [ESY]) to consist of a collection of super-cycles. On input $G=([N],E)$ and proximity parameter $\epsilon>0$, our (simplified) algorithm proceeds as follows.

1. Select arbitrarily a vertex $v_0$.
2. Select at random a sample $S_1$ of $\Theta(1/\epsilon)$ vertices, and query all pairs $(v_0,u)$ for $u\in S_1$. If $S_1\cap\Gamma(v_0)=\emptyset$, then accept. Otherwise, select arbitrarily a vertex $v_1\in S_1\cap\Gamma(v_0)$.
3. For $i=1,...,t-1$, attempt to find a vertex $v_{i+1}\in\Gamma(v_i)$ such that $v_{i+1}$ does not reside in the same independent set as $v_{i-1}$ (i.e., $\Gamma(v_{i+1})=\Gamma(v_{i-1})$). This is done as follows.
   (a) Select at random a sample $S_{i+1}$ of $\Theta(1/\epsilon)$ vertices, and query all pairs $(v_i,u)$ for $u\in S_i$.
   (b) If $S_{i+1}\cap\Gamma(v_i)=\emptyset$, then accept. Otherwise, we let $T\stackrel{\rm def}{=}S_{i+1}\cap\Gamma(v_i)$, and proceed as follows.
   (c) Select at random a set $U$ of $O(1)$ vertices in $T$, and an auxiliary sample $R$ of $O(1/\epsilon)$ vertices of $G$. Query all pairs $(U\cup\{v_{i-1}\})\times R$, and determine $\Gamma_R(u)\stackrel{\rm def}{=}R\cap\Gamma(u)$ for each $u\in U\cup\{v_{i-1}\}$. If for every $u\in U$, it holds that $\Gamma_R(u)=\Gamma_R(v_{i-1})$, then accept. Otherwise, select arbitrarily a vertex $v_{i+1}\in U$ such that $\Gamma_R(v_{i+1})\neq\Gamma_R(v_{i-1})$.
4. Select at random an auxiliary sample $R$ of $O(1/\epsilon)$ vertices, and query all pairs $(\{v_0,v_t\})\times R$. Accept if and only if $\Gamma_R(v_t)=\Gamma_R(v_0)$.

This algorithm has query complexity $O(1/\epsilon)$ (recall that $t$ is a constant) and it accepts any graph in $\mathcal{S}\mathtt{C}_t\mathcal{C}$, since whenever a path $(v_0, v_1, ..., v_t)$ is found, it is the case that $v_0$ and $v_t$ reside in the same independent set (and hence satisfy $\Gamma(v_0) = \Gamma(v_t)$). On the other hand, if $G \in \mathcal{S}\mathtt{C}_{2t}\mathcal{C}'$ is $\epsilon$-far from $\mathcal{S}\mathtt{C}_t\mathcal{C}$, then it must be that $G \in \mathcal{S}\mathtt{C}_{2t}\mathcal{C}_{2^{-j}}$, for some $j \leq \log_2 4/\epsilon$. In this case, with high constant probability, the algorithm does not accept $G$ in Step 2, since the sample $S_1$ is likely to hit the set $\Gamma(v_0)$ (which has cardinality $2^{-j}N \geq \epsilon N/4$). Similarly, with high constant probability, the algorithm does not accept $G$ in any iteration of Step 3, since the sample $S_{i+1}$ is likely to contain at least one vertex $u$ in $\Gamma(v_i)$ that does not reside in the same independent set as $v_{i-1}$ (and the auxiliary sample $R$ is likely to contain some vertex in $(\Gamma(u) \cup \Gamma(v_{i-1})) \setminus (\Gamma(u) \cap \Gamma(v_{i-1}))$). Lastly, observe that for the constructed path $(v_0, v_1, ..., v_t)$ it holds that $v_0$ and $v_t$ do not reside in the same independent set, and furthermore $\Gamma(v_0) \cap \Gamma(v_t) = \emptyset$. Thus, Step 4 rejects with high constant probability.

Finally, we turn to Part 3, which is established by a (canonical) tester that inspects the subgraph induced by a uniformly selected set of $O(\epsilon^{-1+(1/t)})$ vertices, and *rejects* if and only if this set contains $t$ vertices such that the subgraph induced by these $t$ vertices is a simple path (i.e., contains only the $t-1$ edges of this path). This algorithm never rejects any graph $G \in \mathcal{S}\mathtt{C}_t\mathcal{C}$, because if the subgraph of $G$ induced by some set of $t$ vertices contains a $t$-vertex path, denoted $(v_1, ..., v_t)$, then either each $v_i$ resides in a different independent set of the same super-cycle (which implies that $v_t$ and $v_1$ are connected) or some $v_i$ and $v_{i+2}$ reside in the same independent set (which yields the 4-cycle containing $(v_{i-1}, v_i, v_{i+1}, v_{i+2})$). In contrast, for every $j \leq \log_2 4/\epsilon$, every graph in $G \in \mathcal{S}\mathtt{C}_{2t}\mathcal{C}_{2^{-j}}$ contains sets of $t$ vertices such that the subgraph induced by each such set is a simple $t$-vertex path. Furthermore, with high constant probability, a random sample of $O(\epsilon^{-1+(1/t)})$ vertices contains such a set, because such a sample contains at least $(3t)!$ random $t$-vertex sets that are each contained in the same super-cycle,[13] and with probability at least $1/(2t)!$ each such $t$-vertex set induces a path. ■

## 6 Non-Adaptive Testing with $\widetilde{O}(1/\epsilon)$ Complexity

We first note that $\Omega(1/\epsilon)$ (adaptive) queries are required for testing any graph property that is non-trivial for testing, where a graph property $\Pi$ is non-trivial for testing if there exists $\epsilon_0 > 0$ such that for infinitely many $N \in \mathbb{N}$ there exist $N$-vertex graphs $G_1$ and $G_2$ such that $G_1 \in \Pi$ and $G_2$ is $\epsilon_0$-far from $\Pi$. We note that all properties considered in this work are non-trivial for testing. On the other hand, the negation of this (non-triviality) condition means that for every $\epsilon > 0$ and all sufficiently large $N \in \mathbb{N}$ either $\Pi$ contains no $N$-vertex graph or all $N$-vertex graphs are $\epsilon$-close to $\Pi$. In such a case (for every such $\epsilon$ and $N$), the tester may decide without even looking at the graph.[14] Turning back to properties that are non-trivial for testing, we prove that any tester for such a property must have query complexity $\Omega(1/\epsilon)$.

**Proposition 6.1** *Let $\Pi$ be a property that is non-trivial for testing. Then, any tester for $\Pi$ has query complexity $\Omega(1/\epsilon)$.*

[13] The claim follows by using a generalized birthday problem. In our case we have $B = 2^j$ bins and claim that, with high constant probability, assigning at random $b = \widetilde{O}(t) \cdot B^{1-(1/t)}$ balls to these bins results in having some bin contain $t$ balls. This can be proved by considering a $t$-step process, so that at each step $O(\log t) \cdot B^{1-(1/t)}$ balls are assigned. For $j = 1, ..., t$, we claim that after the $j^{\text{th}}$ step, with probability at least $1 - o(1/t)$, there are at least $B^{1-(j/t)}$ bins that contain $j$ balls each. This claim is easily proved by induction on $j$.

[14] Indeed, there exists natural graph properties that are trivial for testing (e.g., connectivity, non-planarity, having no vertex of odd degree); see [GGR, Sec. 10.2.1].

Note that the claim holds also for general properties (i.e., arbitrary sets of functions).

**Proof:** Let $\epsilon_0 > 0$ be as in the definition, and consider any $N \in \mathbb{N}$ such that $\Pi$ contains some $N$-vertex graphs and there exist some $N$-vertex graphs that are $\epsilon_0$-far from $\Pi$. Let $G_0$ be any $N$-vertex graph that is $\epsilon_0$-far from $\Pi$, let $G_1 \in \Pi$ be an $N$-vertex graph closest to $G_0$, and let $\delta > \epsilon_0$ denote the relative distance between $G_0$ and $G_1$. Let $D$ denote the set of vertex pairs on which $G_0$ and $G_1$ differ; indeed, $|D| = \delta \cdot N^2$. Now, for every $\epsilon \leq \epsilon_0$ (and $\epsilon > N^{-2}$), consider a graph, $G$, obtained at random from $G_0$ and $G_1$ by uniformly selecting a random $R \subseteq D$ of cardinality $\epsilon \cdot N^2$ and letting $G$ agree with $G_0$ on all pairs in $R$ and agree with $G_1$ otherwise. Clearly, any tester that makes $o(\epsilon_0/\epsilon)$ queries cannot distinguish $G$ from $G_1$ (because regardless of its query selection strategy, its next query resides in $R$ with probability at most $|R|/|D| \leq \epsilon/\epsilon_0$). Thus, such a tester cannot decide correctly on both $G$ and $G_1$ (because $G$ is $\epsilon$-far from $\Pi$ whereas $G_1 \in \Pi$). Recalling that $\epsilon_0$ is a fixed constant, the proposition follows. ■

To justify the fact that all our testers are inherently non-canonical, we show that (for any property that is non-trivial for testing) canonical testers must use $\Omega(\epsilon^{-2})$ queries.

**Proposition 6.2** *Let $\Pi$ be a property that is non-trivial for testing. Then, any canonical tester for $\Pi$ has query complexity $\Omega(1/\epsilon^2)$.*

**Proof:** We adapt the proof of Proposition 6.1 so as to force any canonical tester to sample $\Omega(1/\epsilon)$ vertices. Let $\epsilon_0 > 0$, $G_0 = ([N], E_0)$ and $G_1 = ([N], E_1) \in \Pi$ be as in that proof. Then, there exists a set of at least $\epsilon_0 N/2$ vertices, denoted $B$, such that for every $v \in B$ the symmetric difference between the sets $\{u : \{v,u\} \in E_0\}$ and $\{u : \{v,u\} \in E_1\}$ has size at least $\epsilon_0 N/2$. Now, for every $\epsilon \leq \epsilon_0/2$ (and $\epsilon > N^{-1}$), consider a graph, $G$, obtained from $G_0$ and $G_1$ by arbitrarily selecting a subset $D \subseteq B$ of cardinality $(2\epsilon/\epsilon_0) \cdot N$ and letting $G$ agree with $G_0$ on all vertex pairs that intersect $D$ and agree with $G_1$ otherwise. Clearly, $G$ is $\epsilon$-far from $\Pi$, but any canonical tester that selects $o(\epsilon_0/\epsilon)$ random vertices cannot distinguish $G$ from $G_1$. Thus, such a tester cannot decide correctly on both $G$ and $G_1$ (because $G$ is $\epsilon$-far from $\Pi$ whereas $G_1 \in \Pi$). Recalling that $\epsilon_0$ is a fixed constant, the proposition follows. ■

## 6.1 Clique and Bi-Clique

We start with the problem of testing whether the given graph is a clique (or, equivalently, an independent set). The algorithm consists of selecting uniformly $O(1/\epsilon)$ vertex-pairs and checking whether each of these pairs is connected by an edge. Clearly, if the graph is $\epsilon$-far from being a clique, then a randomly selected pair of vertices is connected with probability at most $1 - \epsilon$. The foregoing algorithm and analysis seem to provide the simplest example of a graph property that can be tested by $O(1/\epsilon)$ non-adaptive queries. A somewhat less simple example is provided by testing the property of being a bi-clique.

**Algorithm 6.3** (non-adaptive test of bi-cliqueness): *On input $N$ and $\epsilon$ and oracle access to a graph $G = ([N], E)$, set $t = \Theta(1/\epsilon)$ and select arbitrarily a start vertex $s$* (e.g., $s = 1$). *For $i = 1, \ldots, t$, uniformly select a pair of vertices $(u_i, v_i)$, and make the queries $(s, u_i)$, $(s, v_i)$, and $(u_i, v_i)$. Accept if and only if for every $i$ an even number of the answers are positive* (i.e., indicate the existence of an edge).

Clearly, if $G$ is a bi-clique then for every $i$ either all vertices reside on the same side (and so $(s, u_i)$, $(s, v_i)$, and $(u_i, v_i)$ are all non-edges) or a single vertex is in solitude (and is thus adjacent to the other two vertices). To analyze what happens when $G$ is $\epsilon$-far from being a bi-clique we observe that $s$ induces a partition of the graph to neighbors and non-neighbors (i.e., the 2-partition $(\Gamma(s), [N] \setminus \Gamma(s))$). That is, if $G$ were a bi-clique then every vertex $v \in \Gamma(s)$ (resp., $v \in [N] \setminus \Gamma(s)$) would have satisfied $\Gamma(v) = [N] \setminus \Gamma(s)$ (resp., $\Gamma(v) = \Gamma(s)$).[15] However, since $G$ is $\epsilon$-far from being a bi-clique, it follows that either there are at least $\frac{\epsilon}{2} \cdot N^2$ edges in $(\Gamma(s) \times \Gamma(s)) \cup (([N] \setminus \Gamma(s)) \times ([N] \setminus \Gamma(s)))$ or at least $\frac{\epsilon}{2} \cdot N^2$ edges are missing from $\Gamma(s) \times ([N] \setminus \Gamma(s))$. Thus, the sample of $t$ pairs will hit such an edge with probability at least 2/3.

### 6.2 Collection of a Constant Number of Cliques

For any constant $c$, we consider the set of graphs that each consists of a collection of (up to) $c$ cliques; that is, the property $\mathcal{CC}^{\leq c}$. Note that the special case of $\mathcal{CC}^{\leq 2}$ is analogous to bi-clique, because a graph $G = ([N], E)$ is in $\mathcal{CC}^{\leq 2}$ if and only if its complement graph $([N], ([N] \times [N]) \setminus E)$ is a bi-clique. Here we deal with the general case of a constant $c \geq 3$.

To motivate the following non-adaptive tester (Algorithm 6.4), consider first the case in which the input graph consists of $c + 1$ cliques such that the smallest clique has size $2\sqrt{\epsilon}N$. In this case, with high probability, a sample of $O(\epsilon^{-1/2})$ random vertices contains an independent set of size $c + 1$, which will be discovered if we probe the entire induced subgraph. This case will be detected in Step 1 of the algorithm. To motivate Step 2, consider the case that, for some $\alpha \in (3\epsilon, o(\sqrt{\epsilon}))$, the graph consists of two cliques of size $(1 - \alpha)N/2$ and a third clique of size $\alpha N$ such that each vertex in the third clique is connected to an $\epsilon/\alpha$ fraction of the vertices in each of the large cliques. In this case, Step 1 is unlikely to sample a vertex of the small clique (and will thus fail to detect that this graph is $\epsilon$-far from $\mathcal{CC}^{\leq c}$), but a sample as in Step 2 (with $i = \log_2(\alpha/\epsilon)$) is likely to contain a vertex of the small clique as well as a neighbor from each of the two large cliques.

**Algorithm 6.4** (non-adaptive test for $\mathcal{CC}^{\leq c}$): *On input $N$ and $\epsilon$ and oracle access to a graph $G = ([N], E)$, set $\ell = \log_2(8c^2/\epsilon)$ and proceed as follows.*

1. *Select a uniform sample of $\Theta(\epsilon^{-1/2})$ vertices, denoted $S$, and examine all vertex pairs in $S$.*

2. *For $i = 1, \ldots, \ell$ select, uniformly at random, samples of $\Theta(\log(1/\epsilon)/(2^i\epsilon))$ and $\Theta(2^i)$ vertices in $[N]$ denoted $T_i^1$ and $T_i^2$, respectively, and a sample of $\Theta(\min\{2^i, 1/(2^i\epsilon)\})$ vertices in $S$, denoted $S_i$. Examines all the vertex pairs in $S_i \times (T_i^1 \cup T_i^2)$ and in $T_i^1 \times T_i^2$.*

3. *Accept if and only if the view of the subgraph as obtained in Steps 1-2 is consistent with some graph in $\mathcal{CC}^{\leq c}$. Namely, let $g' : \left((S \times S) \cup \left(\bigcup_{i=1}^{\ell} \left((S_i \times (T_i^1 \cup T_i^2)) \cup (T_i^1 \times T_i^2)\right)\right)\right) \to \{0,1\}$ be the function determined by the answers obtained in Steps 1-2. Then, the test accepts if and only if $g'$ can be extended to a function over $S' \times S'$ that represents a graph in $\mathcal{CC}^{\leq c}$, where $S' \stackrel{\text{def}}{=} S \cup \left(\bigcup_{i=1}^{\ell} (T_i^1 \cup T_i^2)\right)$.*

Step 3 can be implemented efficiently by constructing the connected components of the graph defined by the positive answers (cf. discussion following Algorithm 4.3). It is instructive to spell out several implications of the acceptance criterion that underlies Step 3. Indeed, this criterion implies that the following four conditions hold (or equivalently, if any one of them is violated, then the algorithm will reject):

[15] Indeed, this is a simple application of the "induced partition" idea, which underlies the analysis of many of the testers of [GGR].

(i) The subgraph induced by $S$ is in $\mathcal{CC}^{\leq c}$.

In such a case, we denote the corresponding cliques by $C_1, \ldots, C_{c'}$, where $c' \leq c$.

(ii) For every $i \in [\ell]$ and every $v \in T_i^1 \cup T_i^2$, either $\Gamma(v) \cap S_i = \emptyset$ or, for some $j \in [c']$, it holds that $\Gamma(v) \cap S_i = C_j \cap S_i$.

(iii) For every $i \in [\ell]$, if $|\{j : C_j \cap S_i \neq \emptyset\}| = c$ then every $v \in T_i^1 \cup T_i^2$ has at least one neighbor in $S_i$.

(iv) For every $i \in [\ell]$ and for every $v \in T_i^1$ and $u \in T_i^2$ such that $\Gamma(v) \cap S_i \neq \emptyset$ and $\Gamma(u) \cap S_i \neq \emptyset$ the following holds. If $\Gamma(v) \cap S_i = \Gamma(u) \cap S_i$ then $(v, u) \in E$, while if $\Gamma(v) \cap S_i \neq \Gamma(u) \cap S_i$, then $(v, u) \notin E$.

(We mention that it is considerably easier to design and analyze an adaptive tester of query complexity $O(1/\epsilon)$ for $\mathcal{CC}^{\leq c}$; see a more general result in [A09, Sec. 4].) Algorithm 6.4 has query complexity

$$\begin{aligned} |S|^2 + \sum_{i=1}^{\ell} \left( |S_i| \cdot (|T_i^1| + |T_i^2|) + |T_i^1| \cdot |T_i^2| \right) &= O(1/\epsilon) + \log(1/\epsilon) \cdot O(\log(1/\epsilon)/\epsilon) && (117) \\ &= \widetilde{O}(1/\epsilon) && (118) \end{aligned}$$

and accepts every graph in $\mathcal{CC}^{\leq c}$ with probability 1. We thus turn to analyze the case that the input graph $G = ([N], E)$ is $\epsilon$-far from $\mathcal{CC}^{\leq c}$. Namely, we show:

**Lemma 6.5** *If $G$ is $\epsilon$-far from $\mathcal{CC}^{\leq c}$ then Algorithm 6.4 rejects with probability at least* $2/3$.

Theorem 1.4 follows.

**Proof:** The analysis relies on the fact that $\mathcal{CC}^{\leq c}$ is a hereditary property (i.e., any induced subgraph of any graph in $\mathcal{CC}^{\leq c}$ is also in $\mathcal{CC}^{\leq c}$), which implies that any independent set of size $c+1$ is a witness for the input graph not being in $\mathcal{CC}^{\leq c}$. Thus, considering only the sample $S$ (selected in Step 1), we show that, with high constant probability, either $S$ contains such an independent set (and the algorithm rejects) or $S$ induces a partition of almost all the graph's vertices. In the latter case, with high constant probability, the auxiliary samples and queries made in Step 2 will cause the algorithm to reject. Details follow.

We start by considering the choice of $S$ (in Step 1 of the algorithm). We think of $S$ as being selected in $c+1$ phases (where $c$ is a constant), such that in phase $t \in [c+1]$, a new uniform sample $S^t$, of $\Theta(\epsilon^{-1/2})$ vertices, is selected. Intuitively, the objective of the first $c$ phases is to ensure, with high (constant) probability, that as long as the number of vertices that do not have any neighbor among the vertices selected so far is relatively big, we obtain such a vertex in the next phase. After $c$ phases we use the selected vertices to define a partition of the graph vertices into at most $c$ subsets with some *exceptional* vertices (which either do not have any neighbor among the vertices selected in the previous phases or are somehow inconsistent with these vertices). The objective of phase $c+1$ is to ensure that (with high probability) the number of exceptional vertices is relatively small (or else, cause rejection).

For each $t \in [c+1]$, let $S^{\leq t} = \bigcup_{k=1}^{t} S^k$. Recall that the algorithm queries all vertex pairs in $S \times S$. Hence, if for any $t \in [c+1]$, the subgraph induced by $S^{\leq t}$ is not a collection of at most $c$ cliques, then the algorithm rejects, and we are done. Otherwise, let $C_1^t, \ldots, C_{b^{(t)}}^t$ denote the $b^{(t)} \leq c$

cliques in the subgraph induced by $S^{\leq t}$. For each $t \in [c+1]$, we define the following partition of the set $[N]$ of all graph vertices:

$$V_j^t \stackrel{\rm def}{=} \{v : \Gamma(v) \cap S^{\leq t} = C_j^t\} \quad \text{for } 1 \leq j \leq b^{(t)}, \tag{119}$$

$$R_0^t \stackrel{\rm def}{=} \{v : \Gamma(v) \cap S^{\leq t} = \emptyset\}, \tag{120}$$

$$R_1^t \stackrel{\rm def}{=} [N] \setminus \Big(R_0^t \cup \Big(\bigcup_{1 \leq j \leq b^{(t)}} V_j^t\Big)\Big). \tag{121}$$

That is, for every $j \in [b^{(t)}]$, the subset $V_j^t$ consists of the vertices that neighbor all vertices in $C_j^t$ and no other vertex in $S^{\leq t}$, the subset $R_0^t$ consists of all vertices that have no neighbor in $S^{\leq t}$, and $R_1^t$ consists of all vertices that either neighbor only some of the vertices in one of the cliques $C_j^t$ (but not all) or have neighbors in more than one of the cliques.

Given the above notation, we make the following observations. First, *for any choice of $S$, it holds that $V_j^{t+1} \subseteq V_j^t$ for every $j \in [b^{(t)}]$*, and likewise $R_0^{t+1} \subseteq R_0^t$ *while* $R_1^{t+1} \supseteq R_1^t$. Next, we turn to probabilistic assertions, which refer to random choices of $S$.

1. For any $t \in [c]$ and any fixing of $S^{\leq t}$, if $|R_1^t| > \frac{1}{4}\epsilon^{1/2}N$, then the algorithm rejects with high probability (where the probability is taken over the choice of $S^{t+1}$).

   This holds because, under the hypothesis, it is very likely that $S^{t+1}$ will contain some vertex in $R_1^t$, whereas in this case the subgraph induced by $S^{\leq(t+1)}$ is not a collection of (at most $c$) cliques, and the algorithm rejects.

2. For any $t \in [c]$ and any fixing of $S^{\leq t}$, if $|R_0^t| > \frac{1}{4}\epsilon^{1/2}N$, then, with high probability, $b^{(t+1)} \geq b^{(t)} + 1$ (where the probability is taken over the choice of $S^{t+1}$).

   This holds because, under the hypothesis, it is very likely that $S^{t+1}$ will contain some vertex in $R_0^t$, whereas such a vertex cannot fit to any of the existing cliques.

3. For any $t \in [c]$ and any fixing of $S^{\leq t}$, for every $j \in [b^{(t)}]$ such that $|V_j^t| \geq \frac{\epsilon^{-1/2}}{2c}N$, with high probability (over the choice of $S^{t+1}$), it holds that

$$\frac{|C_j^{t+1}|}{|S^{t+1}|} \geq 0.9 \cdot \frac{|V_j^t|}{N}. \tag{122}$$

   This follows by an application of the standard multiplicative Chernoff bound.

Combining the foregoing observations, we infer that for, say, a 0.99 fraction of the possible choices of $S$ either the subgraph induced by $S$ is not in $\mathcal{CC}^{\leq c}$ or there exists $t^* \in [c]$ such that

**(1)** $|R_1^{t^*+1}|, |R_0^{t^*+1}| \leq \frac{1}{4}\epsilon^{1/2}N$,

**(2)** $b^{(t^*+1)} = b^{(t^*)}$, and

**(3)** for every $j \in [b^{(t^*+1)}]$ such that $|V_j^{t^*+1}| \geq \frac{\epsilon^{-1/2}}{2c}N$ it holds that $|C_j^{t^*+1}|/|S| \geq (2c)^{-1} \cdot |V_j^{t^*+1}|/N$.

Thus, throughout the rest of our analysis, we shall assume that the latter three conditions hold. (We later take into account the small constant probability that this is not the case and the algorithm did not reject.)[16]

Fixing $t^*$ as above, we simplify the notation by using the following shorthands: $C_j$ for $C_j^{t^*+1}$, $V_j$ for $V_j^{t^*+1}$, $R_0$ for $R_0^{t^*+1}$, $R_1$ for $R_1^{t^*+1}$, and $c'$ for $b^{(t^*+1)}$. We also denote $R_0 \cup R_1$ by $R$.

Recall that $G$ is $\epsilon$-far from $\mathcal{CC}^{\leq c}$. This means that for every partition of the graph vertices into at most $c$ subsets, the total number of vertex pairs that either belong to the same subset but do not have an edge between them, or belong to different subsets but do have an edge between them, is greater than $\epsilon N^2$. In particular, this holds for the partition of $[N]$, denoted $(\widetilde{V}_j)_{j\in\{0,1,\ldots,c'\}}$, that we define as follows:

- For every $j \in [c']$, it holds that $V_j \subseteq \widetilde{V}_j$.
- The vertices in $R$ are partitioned among the $\widetilde{V}_j$'s so as to minimize the number of violations caused by pairs of the form $(v,w) \in R \times ([N] \setminus R)$. Specifically, for every vertex $v \in R$ and $j \in [c']$, let $e_j(v) = |\Gamma(v) \cap V_j|$ (resp., $\bar{e}_j = |V_j \setminus \Gamma(v)|$) denote the number of neighbors (resp., non-neighbors) that $v$ has in $V_j$. If $c' = c$ then each vertex $v \in R$ is placed in the subset $\widetilde{V}_j$ for which $\bar{e}_j(v) + \sum_{k\in[c']\setminus\{j\}} e_k(v)$ is minimized. If $c' < c$ then we do the same, except that every vertex $v \in R$ that satisfies $\sum_{k=1}^{c'} e_k(v) < \min_{j\in[c']}\{\bar{e}_j(v) + \sum_{k\in[c']\setminus\{j\}} e_k(v)\}$ is placed in $\widetilde{V}_0$. In particular, $v$ is placed in $\widetilde{V}_0$ if and only if for every $j \in [c']$ it holds that $e_j(v) < \bar{e}_j(v)$ (which is equivalent to saying that for every $j \in [c']$ it holds that $\sum_{k=1}^{c'} e_k(v) < \bar{e}_j(v) + \sum_{k\in[c']\setminus\{j\}} e_k(v)$).

We note that it may be the case that $\widetilde{V}_0 = \emptyset$; indeed, this always happens when $c' = c$.
Recall that $|R| \leq \frac{1}{2}\epsilon^{1/2}N$. Therefore, the total number of vertex pairs in $R \times R$ is at most $\frac{1}{4}\epsilon N^2$. It follows that if $G$ is $\epsilon$-far from $\mathcal{CC}^{\leq c}$ then (at least) one of the following three events must occur:

1. There are at least $\frac{1}{4}\epsilon N^2$ missing edges between pairs of vertices that belong to the same subset $V_j$; that is, $\sum_{j=1}^{c'} |(V_j \times V_j) \setminus E| \geq \frac{\epsilon}{4}N^2$.
2. There are at least $\frac{1}{4}\epsilon N^2$ superfluous edges between pairs of vertices that belong to different subsets $V_j$ and $V_k$; that is, $\sum_{j=1}^{c'-1}\sum_{k=j+1}^{c'} |(V_j \times V_k) \cap E| \geq \frac{\epsilon}{4}N^2$.
3. The total number of missing and superfluous edges contributed by pairs of vertices in $R \times (\bigcup_{j=1}^{c'} V_j)$ is at least $\frac{1}{4}\epsilon N^2$. That is, if for each $j \in [c']$ and $v \in R \cap \widetilde{V}_j$ we let
$$x(v) = \bar{e}_j(v) + \sum_{k\in[c']\setminus\{j\}} e_k(v) \ , \tag{123}$$
and for $v \in R \cap \widetilde{V}_0$ we let
$$x(v) = \sum_{1\leq k\leq c'} e_k(v) \ , \tag{124}$$
then $\sum_{j=0}^{c'} \sum_{v\in R\cap\widetilde{V}_j} x(v) \geq \frac{\epsilon}{4}N^2$. (Recall that $\widetilde{V}_0 = \emptyset$ whenever $c' = c$.)

[16] Specifically, if the algorithm accepts with probability at least, say, 0.001, then (by Item 1) $|R_1^t| \leq \frac{1}{4}\epsilon^{1/2}N$ typically holds (for any $t$). By Item 2, $|R_0^t| > \frac{1}{4}\epsilon^{1/2}N$ typically implies $b^{(t+1)} > b^{(t)}$, and so $b^{(t+1)} = b^{(t)}$ indicates that $|R_0^t| \leq \frac{1}{4}\epsilon^{1/2}N$ (while $|R_0^{t+1}| \leq |R_0^t|$ always holds). Noting that we cannot have $b^{(t+1)} > b^{(t)}$ for every $t \in [c]$, it follows that for, say, a 0.99 fraction of the choices of $S$, there exists a $t^* \in [c]$ that satisfies conditions (1) and (2). On the other hand, for a 0.999 fraction of the choices of $S$, for every $t \in [c]$ and every $j \in [b^{(t+1)}]$ such that $|V_j^{t+1}| \geq \frac{\epsilon^{-1/2}}{2c}N$ it holds that $|C_j^{t+1}|/|S^{t+1}| \geq 0.9|V_j^{t+1}|/N$. Using $|S| = (c+1)\cdot|S^{t+1}|$ and $0.9/(c+1) > 1/2c$, the claim follows.

It remains to prove that in each of the three foregoing cases the algorithm rejects with probability at least 5/6. Specifically, we shall show that, with probability at least 5/6, there exists an $i \in [\ell]$ such that the sample $S_i \cup T_i^1 \cup T_i^2$ contains a set of vertices which induces a subgraph not in $\mathcal{CC}^{\leq c}$ that is inspected by the algorithm. More specifically, this set will contain at most one vertex from each $T_i^b$, and we shall use the fact that the algorithm inspects all pairs in $(S_i \times (T_i^1 \cup T_i^2)) \cup (T_i^1 \times T_i^2) \cup (S_i \times S_i)$. In what follows let $\epsilon' = \frac{\epsilon}{8\ell c^2}$ (and recall that $\ell = \log_2(8c^2/\epsilon)$).

Case 1: $\sum_{j=1}^{c'} |(V_j \times V_j) \setminus E| \geq \frac{\epsilon}{4} N^2$. In this case there must be an index $j^* \in [c']$ such that the number of missing edges with both endpoints in $V_{j^*}$ is at least $\frac{\epsilon}{4c} N^2$; that is,

$$\sum_{v \in V_{j^*}} |V_{j^*} \setminus (\{v\} \cup \Gamma(v))| \geq \frac{\epsilon}{4c} N^2 . \tag{125}$$

In particular, this implies that $|V_{j^*}| \geq \frac{\epsilon^{1/2}}{2c^{1/2}} N$. For each $i \in [\ell]$, we define a subset $B_{j^*,i}$ of $V_{j^*}$ as follows.

$$B_{j^*,i} = \left\{ v \in V_{j^*} \,:\, |V_{j^*} \setminus (\{v\} \cup \Gamma(v))| \geq \frac{N}{2^i} \right\}, \tag{126}$$

where $B_{j^*,0} = \emptyset$. By Eq. (125) and since the contribution of vertices outside $B_{j^*,\ell}$ is at most $N \cdot 2^{-\ell} N = \epsilon N^2 / 8c^2$, we have

$$\sum_{i=1}^{\ell} |B_{j^*,i} \setminus B_{j^*,i-1}| \cdot \frac{N}{2^{i-1}} \;>\; \frac{\epsilon}{8c} N^2 \tag{127}$$

and thus there exists $i^* \in [\ell]$ (i.e., a set $B_{j^*,i^*}$) such that

$$|B_{j^*,i^*}| \;>\; \frac{2^{i^*-1}\epsilon}{8c\ell} N \;>\; 2^{i^*} \epsilon' N . \tag{128}$$

By the definition of $B_{j^*,i}$ if $B_{j^*,i} \neq \emptyset$, then $|V_{j^*}| \geq N/2^{i^*}$. Since $B_{j^*,i^*} \neq \emptyset$, it holds that $|V_{j^*}| \geq \alpha N$ where $\alpha = \max\{1/2^{i^*}, \frac{\epsilon^{1/2}}{2c^{1/2}}\}$. We shall show that, with high probability, the following three events occur: (1) $S_{i^*}$ contains at least one vertex $w$ from $C_{j^*}$; (2) $T_{i^*}^1$ contains at least one vertex $v$ from $B_{j^*,i^*} \subseteq V_{j^*}$; and (3) $T_{i^*}^2$ contains at least one vertex $u$ from $V_{j^*} \setminus \Gamma(v)$. If the three events occur then the algorithm rejects since it obtains evidence that the graph is not in $\mathcal{CC}^{\leq c}$ (in the form of $(w,v),(w,u) \in E$ and $(v,u) \notin E$). (Indeed, $v \in \Gamma(w)$ since $w \in C_{j^*}$ and $v \in V_{j^*}$, and $u \in \Gamma(w) \setminus \Gamma(v)$ since $u \in V_{j^*} \setminus \Gamma(v)$. Also note that the algorithm queries all pairs in $(S_{i^*} \times (T_{i^*}^1 \cup T_{i^*}^2)) \cup (T_{i^*}^1 \times T_{i^*}^2)$.)

Let $\alpha$ be as defined in the foregoing discussion. Since $|V_{j^*}| \geq \alpha N$ and so $|C_{j^*}|/|S| \geq |V_{j^*}|/2cN$, the probability that the first event does not occur is at most $(1-(\alpha/2c))^{|S_{i^*}|}$ which is a small constant (due to our choice of $|S_{i^*}| = \Theta(1/\alpha)$). Similarly (by our choice of $|T_{i^*}^1| = \Theta(\log(1/\epsilon)/(\epsilon 2^{i^*})) = \Theta(\ell/(\epsilon 2^{i^*})) = \Omega(1/(\epsilon' 2^{i^*})))$, the probability that $T_{i^*}^1$ does not contain any vertex from $B_{j^*,i^*}$ is a small constant (due to the lower bound on the density of $B_{j^*,i^*}$ given in Eq. (128)). Finally, assuming that $T_{i^*}^1$ contains a vertex $v \in B_{j^*,i^*}$, the probability that $T_{i^*}^2$ (which has size $\Theta(2^{i^*})$) does not contain any vertex from $V_{j^*} \setminus \Gamma(v)$ is a small constant as well (since, by definition of $B_{j^*,i^*}$, the set $V_{j^*} \setminus \Gamma(v)$ has density at least $2^{-i^*}$).

Case 2: $\sum_{j=1}^{c'-1} \sum_{k=j+1}^{c'} |(V_j \times V_k) \cap E| \geq \frac{\epsilon}{4} N^2$. In this case there exists at least one pair of subsets, $V_{j^*}$ and $V_{k^*}$ (where $j^* \neq k^*$), such that $|(V_{j^*} \times V_{k^*}) \cap E| \geq \frac{\epsilon}{4c^2} N^2$. Assume, without loss of generality, that $|V_{j^*}| \geq |V_{k^*}|$, so that in particular $|V_{j^*}| \geq \frac{\epsilon^{1/2}}{2c} N$. Similarly to Case 1, it follows that there exists a index $i^* \in \{1, \ldots, \ell\}$ and a subset $B_{j^*,i^*} \subseteq V_{j^*}$ such that $|B_{j^*,i^*}| \geq \epsilon' 2^{i^*} N$ (recall that $\epsilon' = \epsilon/(8c^2\ell))$ and for every $v \in B_{j^*,i^*}$ it holds that $|V_{k^*} \cap \Gamma(v)| \geq N/2^{i^*}$. Analogously to Case 1, here we can show that, with high probability, the following three events occur: (1) $S_{i^*}$ contains at least one vertex $w$ from $C_{j^*}$, (2) $T^1_{i^*}$ contains at least one vertex $v$ from $B_{j^*,i^*}$, and (3) $T^2_{i^*}$ contains at least one vertex $u$ from $V_{k^*} \cap \Gamma(v)$. If these three events occur then the algorithm rejects since it obtains evidence that the graph is not in $\mathcal{CC}^{\leq c}$ (in the form of $(w, v) \in E$, $(w, u) \notin E$ and $(v, u) \in E$). The probability that these three events occur is lower-bounded as in Case 1.

Case 3: $\sum_{j=0}^{c'} \sum_{v \in R \cap \widetilde{V}_j} x(v) \geq \frac{\epsilon}{4} N^2$. For each $v \in R$, let $x(v)$ be as defined in Eq. (123) & (124), and let $R' \stackrel{\text{def}}{=} \left\{v \in R : x(v) \geq \frac{\epsilon^{1/2}}{4} N\right\}$. Since $|R| \leq \frac{1}{2}\epsilon^{1/2}N$, we have that $\sum_{v \in (R \setminus R')} x(v) < |R| \cdot \frac{\epsilon^{1/2}}{4} N \leq \frac{\epsilon}{8} N^2$. Therefore,

$$\sum_{v \in R'} x(v) \geq \frac{\epsilon}{8} N^2. \tag{129}$$

By the definition of $R'$, for every $v \in R'$, we have that $x(v) \geq N/2^i$ for some $i \leq \log_2(4/\epsilon^{1/2})$. Therefore, if we define $B_i = \{v \ : \ x(v) \geq N/2^i\}$ for $i \in [\log_2(4/\epsilon^{1/2})]$, then there is an index $i^* \in [\log_2(4/\epsilon^{1/2})]$ such that

$$|B_{i^*}| \geq \frac{\epsilon}{8 \log_2(1/\epsilon)} \cdot 2^{i^*} N > \epsilon' 2^{i^*} N. \tag{130}$$

Similarly to the previous cases, *with high probability, the sample $T^1_{i^*}$ contains at least one vertex $v$ in $B_{i^*}$*. We next show that, for each fixed choice of such a vertex $v \in B_{i^*}$, with high probability over the choice of the samples $S_{i^*}$ and $T^2_{i^*}$, we obtain evidence containing $v$ that $G$ is not in $\mathcal{CC}^{\leq c}$ (i.e., a set of vertices that induces a subgraph not in $\mathcal{CC}^{\leq c}$, while having at most one vertex in each $T^b_{i^*}$).

Let $j^* \in \{0, 1, \ldots, c'\}$ be such that $v \in \widetilde{V}_{j^*}$, and define $\bar{e}_0(v) = e_0(v) = 0$. Observe that since $v \in \widetilde{V}_{j^*}$ we must have that

$$\bar{e}_{j^*}(v) - e_{j^*}(v) \leq \bar{e}_k(v) - e_k(v) \qquad (\forall k \neq j^*)\,, \tag{131}$$

where if $c' = c$ then $1 \leq k \leq c'$, while if $c' < c$ then $0 \leq k \leq c'$. (Note that Eq. (131) holds since otherwise $v$ would be placed in $\widetilde{V}_k$.) Eq. (131) will be useful when we consider the following subcases (which refer to $v \in \widetilde{V}_{j^*}$).

- We first consider the subcase in which $j^* = 0$ (which may occur only when $c' < c$). In this subcase, since $\bar{e}_{j^*}(v) - e_{j^*}(v) = 0 - 0 = 0$, for every $k \in [c']$ we have that $\bar{e}_k(v) \geq e_k(v)$. On the other hand, since $x(v) = \sum_{k=1}^{c'} e_k(v) \geq N/2^{i^*}$, there exists at least one index $k^* \in [c']$ such that $e_{k^*}(v) \geq N/(c2^{i^*})$. Since $\bar{e}_{k^*}(v) \geq e_{k^*}(v)$, we have that $\bar{e}_{k^*}(v) \geq N/(c2^{i^*})$ as well. This also implies that $|V_{k^*}|/N \geq (c2^{i^*})^{-1}$, and so $|C_{k^*}|/|S| \geq |V_{k^*}|/2cN$, we have that $|C_{k^*}|/|S| \geq (2c^2 2^{i^*})^{-1}$. Recall that $|T^2_{i^*}| = \Theta(2^{i^*})$, and that $|S_{i^*}| = \Theta(\min\{2^{i^*}, 1/(\epsilon 2^{i^*})\}) = \Theta(2^{i^*})$, since $i^* \leq \log_2(4/\epsilon^{1/2})$.

  Now, if $|C_{k^*} \cap \Gamma(v)| \geq |C_{k^*}|/2$, then, with high constant probability, the sample $S_{i^*}$ contains a vertex $w$ in $C_{k^*} \cap \Gamma(v)$ (since $|C_{k^*}| = \Omega(|S|/2^{i^*})$), and $T^2_{i^*}$ contains a vertex $u$ in $V_{k^*} \setminus \Gamma(v)$ (since $\bar{e}_{k^*}(v) = \Omega(N/2^{i^*})$). Otherwise (i.e., $|C_{k^*} \setminus \Gamma(v)| \geq |C_{k^*}|/2$), with high probability, $S_{i^*}$ contains a vertex $w$ in $C_{k^*} \setminus \Gamma(v)$, and $T^2_{i^*}$ contains a vertex $u$ in $V_{k^*} \cap \Gamma(v)$ (since $e_{k^*}(v) = \Omega(N/2^{i^*})$). In either case, $w \in C_{k^*}$ and $u \in V_{k^*}$, which implies $(u, w) \in E$, and $w \in \Gamma(v)$ iff $u \notin \Gamma(v)$, which implies that $|\{(u, w), (w, v), (u, v)\} \cap E| = 2$.

  In the subsequent subcases we assume that $j^* > 0$. Using Eq. (123) and the hypothesis $v \in B_{i^*}$, we have $\bar{e}_j(v) + \sum_{k \in [c'] \setminus \{j\}} e_k(v) \ \geq \ N/2^{i^*}$.

- We next consider the subcase in which both $\bar{e}_{j^*}(v) \geq N/2^{i^*+1}$ and $e_{j^*}(v) \geq N/2^{i^*+2}$ hold. Setting $k^* \leftarrow j^*$, we reach a situation as in the first subcase (since $\bar{e}_{k^*}(v) = \Omega(N/2^{i^*})$ and $e_{k^*}(v) = \Omega(N/2^{i^*})$), and we are done as in the first subcase (while noting that the first subcase does not rely on $j^* \neq k^*$).

- The next subcase refers to $\bar{e}_{j^*}(v) \geq N/2^{i^*+1}$ and $e_{j^*}(v) < N/2^{i^*+2}$. In this subcase $\bar{e}_{j^*}(v) - e_{j^*}(v) > 0$ and so it can occur only when $c' = c$ (since otherwise $v$ would be placed in $\widetilde{V}_0$, whereas here $j^* \neq 0$). The fact that $\bar{e}_{j^*}(v) - e_{j^*}(v) \geq N/2^{i+2}$ implies that, for every $k \in [c'] \setminus \{j^*\}$, it holds that $\bar{e}_k(v) \geq e_k(v) + \bar{e}_{j^*}(v) - e_{j^*}(v) \geq N/2^{i^*+2}$. It follows that, for each $k \in [c']$, it holds that $|C_k|/|S| \geq 1/2^{i^*+3}$ (since $|V_k|/N \geq 1/2^{i^*+2}$). Recall that $|S_{i^*}| = \Theta(2^{i^*})$ (and $|T^2_{i^*}| = \Theta(2^{i^*})$).

  If there exists $k^* \in [c']$ such that $|C_{k^*} \cap \Gamma(v)| \geq |C_{k^*}|/2$, then with high probability, $S_{i^*}$ contains a vertex in $C_{k^*} \cap \Gamma(v)$, and $T^2_{i^*}$ contains a vertex in $V_{k^*} \setminus \Gamma(v)$. Otherwise (i.e., $|C_k \setminus \Gamma(v)| \geq |C_k|/2$ for every $k \in [c']$), with high probability, for every $k \in [c']$, the sample $S_{i^*}$ contains a vertex in $C_k \setminus \Gamma(v)$, and recalling that $c' = c$ we obtain evidence (in the form of an independent set of size $c+1$) that $G$ is not in $\mathcal{CC}^{\leq c}$.

- Lastly, we consider the subcase in which $\bar{e}_{j^*}(v) \leq N/2^{i^*+1}$. Since $\bar{e}_{j^*}(v) + \sum_{k \in [c'] \setminus \{j^*\}} e_k(v) = x(v) > N/2^{i^*}$, we obtain $\sum_{k \in [c'] \setminus \{j^*\}} e_k(v) \geq N/2^{i^*+1}$. In such a case, there exists a $k^* \in [c'] \setminus \{j^*\}$ for which $e_{k^*}(v) \geq N/(c2^{i^*+1})$. If $e_{j^*}(v) \geq N/(c2^{i^*+2})$, then with high probability, $T^2_{i^*}$ contains one vertex $u$ in $V_{k^*} \cap \Gamma(v)$ and one vertex $u'$ in $V_{j^*} \cap \Gamma(v)$, while $S_{i^*}$ contains one vertex $w$ in $C_{k^*}$ and one vertex $w'$ in $C_{j^*}$, and we have evidence that $G$ is not a union of cliques (since $(v,u),(v,u'),(u,w),(u',w') \in E$ whereas $(w,w') \notin E$, and all five vertex pairs are inspected by the algorithm).[17] Otherwise (i.e., $e_{j^*}(v) < N/(c2^{i^*+2})$), by Eq. (131), we have that $\bar{e}_{k^*}(v) \geq e_{k^*}(v) + \bar{e}_{j^*}(v) - e_{j^*}(v) \geq N/(c2^{i^*+2})$, and we are in essentially the same situation as the first subcase (since we have $e_{k^*}(v) = \Omega(N/2^{i^*})$ and $\bar{e}_{k^*}(v) = \Omega(N/2^{i^*})$).

This completes the handling of all possible subcases of Case 3, and the lemma follows. ■

# 7 Conclusions

We presented various results regarding the complexity of testing graph properties in the adjacency matrix model. All the properties we considered are easily testable in poly$(1/\epsilon)$-time, and their testing requires at least $\Omega(1/\epsilon)$ queries. Our focus was on a finer study of their query complexity, which distinguishes $O(1/\epsilon)$ queries from poly$(1/\epsilon)$ queries. While the particular properties considered are of natural appeal, our interest in them was as demonstrations of various phenomena and/or perspectives. We conclude this paper by explicitly presenting three perspectives on our results.

**The role of algorithmic design in this model.** Indeed, this is the perspective promoted by the paper's title, and it is delivered most eloquently by Theorems 1.2 and 1.4. In particular, Theorem 1.2 provides the strongest separation know between the query complexity of adaptive testers and non-adaptive ones, whereas Theorem 1.4 (along with Proposition 6.2) provides the strongest separation

[17] Actually, note that it also holds that $(u',w) \notin E$, and thus we obtain evidence in the form of the four vertex pairs $(v,u),(v,u'),(u,w),(u',w)$. Note that we can obtain evidence in the form of three vertex pairs by considering either $(v,u'),(u',w),(v,w)$ or $(v,u),(u,w),(v,w)$.

possible between the query complexity of carefully designed non-adaptive testers and canonical testers.

Indeed, with respect to this perspective, Theorem 1.2 supersedes Theorem 1.1, while Conjecture 1.3 if true would supersede both. Theorem 5.7 provides evidence that Conjecture 1.3 may be true.

**Initiating a study of the general relation of adaptive versus non-adaptive testers (in this model).** Theorems 1.1 and 1.4 are the only results that establish a *tight* relation between the query complexity of adaptive and non-adaptive testers. Furthermore, the upper bounds are demonstrated by efficient one-sided error testers, whereas the lower bounds refer to the query complexity of general (two-sided error) testers. These results assert that the exponent of the relation may be $4/3$ and 1, respectively. Theorem 1.2 does not supersede Theorem 1.1, because Theorem 1.2 just partially establishes another relation exponent (i.e., it asserts that the exponent may be at least $3/2$).

With respect to this perspective, even if Conjecture 1.3 is true for any $t > 4$, this will not supersede any of the above, but rather extrapolate them to all exponents of the form $2 - (2/t)$. (Again, Theorem 5.7 provides evidence that Conjecture 1.3 may be true.)

We mention that Alon [A02] presented non-trivial graph properties that can be tested by $O(1/\epsilon)$ non-adaptive queries, but these testers had two-sided error probability.[18] We also mention that approximating the edge density of a graph (or testing whether it is within some fixed interval) can be performed by $O(1/\epsilon^2)$ non-adaptive queries and does require $\Omega(1/\epsilon^2)$ queries (even if adaptivity is allowed, cf. [CEG]).

**Advancing the study of the properties that are testable in small complexity (i.e., poly$(1/\epsilon)$ queries).** Indeed, Alon *et al.* [AFNS] provided a characterization of graph properties that are testable in complexity that is only related to the proximity parameter $\epsilon$, but we believe that further study of the lower complexity classes is begging, where the lowest complexity classes are firstly $\widetilde{O}(1/\epsilon)$ and secondly poly$(1/\epsilon)$. This paper makes a small contribution to this direction, while focusing on the first class and actually decoupling it to two classes: the class of properties that are testable in $\widetilde{O}(1/\epsilon)$ non-adaptive queries, and the rest of the class of properties that are testable by $\widetilde{O}(1/\epsilon)$ (adaptive) queries. Theorems 1.1, 1.2, and 1.4 all have something to say about it.

## Acknowledgments

We thank Lidor Avigad for comments regarding a previous version of this work and Michael Krivelevich for discussions regarding related issues. We are also grateful to the reviewers for their helpful comments and suggestions.

[18] Specifically, testing $H$-freeness, for any fixed bipartite graph $H$, can be performed by inspecting $O(1/\epsilon)$ uniformly chosen vertex pairs and accepting if and only if no edge is seen (see Remark at the end of [A02, Sec. 2]). Alon [A02, Thm. 1(i)] also shows that $H$-freeness can be tested by one-sided error testers of query complexity poly$(1/\epsilon)$, where the polynomial depends on $H$.

## References


[A81] N. Alon. On the number of subgraphs of prescribed type of graphs with a given number of edges. *Israel J. Math. 38*, pages 116–130, 1981.

[A02] N. Alon. Testing subgraphs of large graphs. *Random Structures and Algorithms*, Vol. 21, pages 359–370, 2002.

[AFKS] N. Alon, E. Fischer, M. Krivelevich and M. Szegedy. Efficient Testing of Large Graphs. *Combinatorica*, Vol. 20, pages 451–476, 2000.

[AFN] N. Alon, E. Fischer, and I. Newman. Testing of bipartite graph properties. *SIAM Journal on Computing*, Vol. 37, pages 959–976, 2007.

[AFNS] N. Alon, E. Fischer, I. Newman, and A. Shapira. A Combinatorial Characterization of the Testable Graph Properties: It's All About Regularity. In *38th STOC*, pages 251–260, 2006.

[AS] N. Alon and A. Shapira. A Characterization of Easily Testable Induced Subgraphs. *Combinatorics Probability and Computing*, 15:791–805, 2006.

[A09] L. Avigad. On the Lowest Level of Query Complexity in Testing Graph Properties. Master Thesis, Weizmann Institute of Science, December 2009.

[AG] L. Avigad and O. Goldreich. Testing Graph Blow-Up. Unpublished manuscript, March 2010. Available from `http://www.wisdom.weizmann.ac.il/~oded/p_lidor.html`

[BHR] E. Ben-Sasson, P. Harsha, and S. Raskhodnikova. 3CNF properties are hard to test. *SIAM Journal on Computing*, Vol. 35 (1), pages 1–21, 2005.

[BT] A. Bogdanov and L. Trevisan. Lower Bounds for Testing Bipartiteness in Dense Graphs. In *IEEE Conference on Computational Complexity*, pages 75–81, 2004.

[CEG] R. Canetti, G. Even and O. Goldreich. Lower Bounds for Sampling Algorithms for Estimating the Average. *IPL*, Vol. 53, pages 17–25, 1995.

[ESY] S. Even, A.L. Selman, and Y. Yacobi. The Complexity of Promise Problems with Applications to Public-Key Cryptography. *Inform. and Control*, Vol. 61, pages 159–173, 1984.

[F04] E. Fischer. On the strength of comparisons in property testing. *Inform. and Comput.*, Vol. 189 (1), pages 107-116, 2004.

[GGR] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *Journal of the ACM*, pages 653–750, July 1998.

[GR02] O. Goldreich and D. Ron. Property Testing in Bounded Degree Graphs. *Algorithmica*, Vol. 32 (2), pages 302–343, 2002.

[GR08] O. Goldreich and D. Ron. On Proximity Oblivious Testing. *ECCC*, TR08-041, 2008. Extended abstract in the proceedings of the *41st STOC*, 2009.

[GT] O. Goldreich and L. Trevisan. Three theorems regarding testing graph properties. *Random Structures and Algorithms*, Vol. 23 (1), pages 23–57, August 2003.

[GR07] M. Gonen and D. Ron. On the Benefit of Adaptivity in Property Testing of Dense Graphs. In *Proc. of RANDOM'07*, LNCS Vol. 4627, pages 525–539, 2007. To appear in *Algorithmica* (special issue of RANDOM and APPROX 2007).

[KKR] T. Kaufman, M. Krivelevich, and D. Ron. Tight Bounds for Testing Bipartiteness in General Graphs. *SIAM Journal on Computing*, 33(6):1441–1483, 2004.

[PR] M. Parnas and D. Ron. Testing the diameter of graphs. *Random Structures and Algorithms*, Vol. 20 (2), pages 165–183, 2002.

[R1] D. Ron. Property Testing: A Learning Theory Perspective. *Foundations and Trends in Machine Learning*, Vol. 1 (3), pages 307–402, 2008.

[R2] D. Ron. Algorithmic and Analysis Techniques in Property Testing. *Foundations and Trends in TCS*, to appear.

[RS06] S. Raskhodnikova and A. Smith. A note on adaptivity in testing properties of bounded-degree graphs. *ECCC*, TR06-089, 2006.

[RS96] R. Rubinfeld and M. Sudan. Robust characterization of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2), pages 252–271, 1996.

[Y77] A.C. Yao. Probabilistic Computation, Towards a Unified Measure of Complexity. In *Proceedings of the Eighteenth Annual Symposium on Foundations of Computer Science*, pages 222–227, 1977.