

Two-Sided Error Proximity Oblivious Testing

Oded Goldreich* Igor Shinkar†

June 11, 2012

Abstract

Loosely speaking, a proximity-oblivious (property) tester is a randomized algorithm that makes a constant number of queries to a tested object and distinguishes objects that have a predetermined property from those that lack it. Specifically, for some threshold probability c , objects having the property are accepted with probability at least c , whereas objects that are ϵ -far from having the property are accepted with probability at most $c - F(\epsilon)$, where $F : (0, 1] \rightarrow (0, 1]$ is some fixed monotone function. (We stress that, in contrast to standard testers, a proximity-oblivious tester is not given the proximity parameter.)

The foregoing notion, introduced by Goldreich and Ron (STOC 2009), was originally defined with respect to $c = 1$, which corresponds to one-sided error (proximity-oblivious) testing. Here we study the two-sided error version of proximity-oblivious testers; that is, the (general) case of arbitrary $c \in (0, 1]$. We show that, in many natural cases, two-sided error proximity-oblivious testers are more powerful than one-sided error proximity-oblivious testers; that is, many natural properties that have no one-sided error proximity-oblivious testers do have a two-sided error proximity-oblivious tester.

Keywords: Property testing, proximity-oblivious testers, one-sided vs two-sided error probability, graph properties, testing properties of distributions.

*Department of Computer Science, Weizmann Institute of Science, Rehovot, ISRAEL.
oded.goldreich@weizmann.ac.il Partially supported by the Israel Science Foundation (grant No. 1041/08).

†Department of Computer Science, Weizmann Institute of Science, Rehovot, ISRAEL.
igor.shinkar@weizmann.ac.il

Contents

1	Introduction	1
I	A First Set of Results	3
2	Classes of Boolean Functions	3
2.1	A generic tester and its analysis	4
2.2	Generalization of Theorem 2.2	6
2.3	POTs can test only intervals	7
3	Graph Properties (in the Adjacency Representation Model)	7
3.1	The class of k -regular graphs	8
3.2	Other regular graph properties	9
3.3	The class of regular graphs	11
3.4	Bounded density of induced copies	14
3.5	Towards a characterization	15
3.6	Impossibility results	19
4	In the Bounded-Degree Graph Model	19
II	Additional Results	22
	A brief summary of Part II	22
5	Classes of Distributions	23
5.1	Characterizing the classes of distributions that have a POT	23
5.2	Closure under disjoint union	25
5.3	Positive corollaries	26
5.4	Negative corollaries	28
6	More on Graph Properties in the Adjacency Rep. Model	32
6.1	Testing $\mathcal{CC}(D)$ for finite sets $D \subseteq \Delta^{(t)}$	33
6.2	Testing $\mathcal{CC}(D)$ for unbounded sets $D(N) \subseteq \Delta^{(2)}$	36
6.3	Testing $\mathcal{CC}(D)$ when $D(N) \subseteq \Delta^{(2)}$ is almost everything	38
6.4	Impossibility results for subclasses of $\mathcal{CC}^{\leq 2}$	42
	Acknowledgments	46
	Bibliography	47
	Appendices: Proofs of Technical Claims	48
A.1	Proof of Claim 3.1.1	48
A.2	Proofs of Propositions 3.4 and 3.5	49
A.3	Proof of Proposition 5.5	52
A.4	Strengthening Corollary 5.8	54
A.5	Testable classes of distributions are not closed under taking complements	55
A.6	Proof of Claim 6.1	57
A.7	Proof of Lemma 6.3.2	57

1 Introduction

In the last fifteen years, 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 local views of the object by performing queries; that is, the object is seen as a function and the testers get oracle access to this function (and thus may be expected to work in time that is sub-linear in the length of the object).

The foregoing description refers to the notion of “far away” objects, which in turn presumes a notion of distance between objects as well as a parameter determining when two objects are considered to be far from one another. The latter parameter is called the **proximity parameter**, and is often denoted ϵ ; that is, one typically requires the tester to reject with high probability any object that is ϵ -far from the property.

Needless to say, in order to satisfy the aforementioned requirement, any tester (of a reasonable property) must obtain the proximity parameter as auxiliary input (and determine its actions accordingly). A natural question, first addressed systematically by Goldreich and Ron [GR09b], is what does the tester do with this parameter (or how does the parameter affect the actions of the tester). A very minimal effect is exhibited by testers that, based on the value of the proximity parameter, determine the number of times that a basic test is invoked, *where the basic test is oblivious of the proximity parameter*. Such basic tests, called *proximity-oblivious testers*, are indeed at the focus of the study initiated in [GR09b].

Loosely speaking, a proximity-oblivious tester (POT) makes a number of queries that does not depend on the proximity parameter, but the quality of its ruling does depend on the actual distance of the tested object to the property.¹ (A standard tester of constant error probability can be obtained by repeatedly invoking a POT for a number of times that depends on the proximity parameter.)

The original presentation (in [GR09b]) focused on POTs that always accept objects having the property. Indeed, the setting of one-sided error probability is the most appealing and natural setting for the study of POT. Still, one can also define a meaningful notion of two-sided error probability proximity-oblivious testers (POTs) by generalizing the definition (i.e., [GR09b, Def. 2.2]) as follows:²

Definition 1.1 (POT, generalized): *Let $\Pi = \bigcup_{n \in \mathbb{N}} \Pi_n$, where Π_n contains functions defined over the domain $[n] \stackrel{\text{def}}{=} \{1, \dots, n\}$, and let $\varrho : (0, 1] \rightarrow (0, 1]$ be monotone. A two-sided error POT with detection probability ϱ for Π is a probabilistic oracle machine T that makes a constant number of queries and satisfies the following two conditions, with respect to some constant $c \in (0, 1]$:*

1. *For every $n \in \mathbb{N}$ and $f \in \Pi_n$, it holds that $\Pr[T^f(n) = 1] \geq c$.*
2. *For every $n \in \mathbb{N}$ and $f : [n] \rightarrow \{0, 1\}^*$ not in Π_n , it holds that $\Pr[T^f(n) = 1] \leq c - \varrho(\delta_{\Pi_n}(f))$, where $\delta_{\Pi_n}(f) = \min_{g \in \Pi_n} \{\delta(f, g)\}$ and $\delta(f, g) \stackrel{\text{def}}{=} |\{x \in [n] : f(x) \neq g(x)\}|/n$.*

The constant c is called the threshold probability (of T).

¹A formal definition is presented below (cf. Definition 1.1).

²For simplicity, we define POTs as making a constant number of queries, and this definition is used throughout the current work. However, as in [GR09b], the definition may be extended to allow the query complexity to depend on n .

Indeed, one-sided error POTs (i.e., [GR09b, Def. 2.2]) are obtained as a special case by letting $c = 1$. Furthermore, for every $c \in (0, 1]$, every property having a one-sided error POT also has a two-sided error POT of threshold probability c (e.g., consider a generalized POT that activates the one-sided error POT with probability c and rejects otherwise). Likewise, every property having a (two-sided error) POT, has a two-sided error POT of threshold probability $1/2$. Lastly, a standard property tester is obtained by repeatedly invoking such a POT for $O(1/\varrho(\epsilon)^2)$, where ϵ is the value of the proximity parameter given to the tester. (Indeed, in case of one-sided error POT, we obtain a one-sided error property tester by $O(\varrho(\epsilon))$ invocations.)

Motivation. Property testing can be thought of as relating local views to global properties, where the local view is provided by the queries and the global property is the distance to a predetermined set. Proximity-oblivious testing takes this relation to an extreme by making the local view independent of the distance. In other words, it refers to the smallest local view that may provide information about the global property (i.e., the distance to a predetermined set). Hence, POTs are a natural context for the study of the relation between local views and global properties of various objects. In addition, a major concrete motivation for the study of POTs is that understanding a natural subclass of testers (i.e., those obtained via POTs) may shed light on property testing at large. This motivation was advocated in [GR09b], while referring to one-sided error POTs, but it extends to the generalized notion defined above.

The first question. The first question that arises is whether the latter generalization (i.e., from one-sided to two-sided error POTs) is a generalization at all (i.e., does it increase the power of POTs). This is not obvious, and for some time the first author implicitly assumed that the answer is negative. However, considering the issue seriously, one may realize that two-sided error POTs exist also for properties that have no one-sided error POT. A straightforward example is the property of Boolean functions that have at least a τ fraction of 1-values, for any constant $\tau \in (0, 1)$. But this example is quite artificial and contrived, and the real question is whether there exist more natural examples. In this paper we provide a host of such examples.

Our results. The current work reports of several natural properties that have two-sided error POTs, although they have no one-sided error POTs. A partial list of such examples includes:

1. Properties of Boolean functions that refer to the fraction of 1-values (i.e., the density of the preimage of 1). Each such property is specified by a constant number of subintervals of $[0, 1]$, and a function satisfies such a property if the fraction of 1-values (of the function) resides in one of these subintervals. See Theorems 2.2 and 2.3.
2. Testing graph properties in the adjacency representation model. One class of properties refers to regular graphs of a prescribed degree and to subclasses of such regular graphs (e.g., regular graphs that consists of a collection of bicliques). Another class refers to graphs in which some fixed graph occurs for a bounded number of times (e.g., at most 1% of the vertex triplets form triangles). See Theorems 3.1, 3.2 and 3.9.
3. Testing graph properties in the bounded-degree model. One class of properties refers to graphs that contain a fraction of isolated vertices that falls in a predetermined set of densities (as in the foregoing Item 1). See Theorem 4.2.

It is evident that none of the foregoing properties has a one-sided error POTs.³ The point is showing that they all have two-sided error POTs.

Unfortunately, the current work does not provide conclusive answers regarding the scope of two-sided error POTs (although some of our results aim at that direction).⁴ It rather provides a first exploration of this notion.

Organization. In order to facilitate developing an intuition regarding the power of two-sided error POTs, we partitioned the exposition into two parts. The first part is organized in three sections, which correspond to the three aforementioned domains: Section 2 deals with properties of Boolean functions, Section 3 deals with testing graph properties in the adjacency representation model (of [GGR]), and Section 4 deals with testing graph properties in the bounded-degree model (of [GR97]). In the second part, we revisit the study of classes of distributions, which underlies the study presented in Part I, and apply the results to further study of graph properties in the adjacency representation model.

The second part is far more technical than the first part, and we chose to present it later in order to allow the reader to go from simple examples to more complex ones. Also, for sake of readability, the proofs of many technical claims (especially in the second part) were moved to the appendix.

Part I

A First Set of Results

2 Classes of Boolean Functions

As mentioned above, a simple example of a property of Boolean functions that has a (two-sided error) POT is provided by the set of all functions that have at least a τ fraction of 1-values, for any constant $\tau \in (0, 1)$. In this case, the POT may query the function at a single uniformly chosen preimage and return the function's value. Indeed, every function in the foregoing set is accepted with probability at least τ , whereas every function that is ϵ -far from the set is accepted with probability at most $\tau - \epsilon$.

A more telling example refers to the set of Boolean function having a fraction of 1-values that is at least τ_1 but at most τ_2 , for any $0 < \tau_1 < \tau_2 < 1$. This property has a two-sided error POT that selects uniformly two samples in the function's domain, obtains the function values on them, and accept with probability α_i if the sum of the answers equals i , where $(\alpha_0, \alpha_1, \alpha_2) = (0, 1, \frac{2(\tau_1 + \tau_2 - 1)}{\tau_1 + \tau_2})$ if $\tau_1 + \tau_2 \geq 1$, and $(\alpha_0, \alpha_1, \alpha_2) = (\frac{2(1 - \tau_1 - \tau_2)}{2 - \tau_1 - \tau_2}, 1, 0)$ otherwise.

In general, we consider properties that are each specified by a sequence of t density thresholds, denoted $\vec{\tau} = (\tau_1, \dots, \tau_t)$, such that $0 < \tau_1 < \tau_2 < \dots < \tau_t < 1$. The corresponding property, denoted

³Consider, for example, the task of testing the set of Boolean functions that have at least a τ fraction of 1-values, for any constant $\tau \in (0, 1)$. A hypothetical one-sided error POT for this property is required to accept each function that has exactly a τ fraction of 1-values, with probability 1, which implies that it must accept regardless of the answers it obtains (since each sequence of answers is consistent with such a function). But, then, this POT accepts each Boolean function with probability 1, which means that it is a POT for the trivial property (rather than for the aforementioned one).

⁴In particular, Theorem 2.4 provides a characterization of binary distributions having a POT, Theorem 5.1 provides a less effective characterization w.r.t multi-valued distributions, and Theorem 3.10 may be viewed as a programmatic step in the context of graph properties.

$\mathcal{B}_{\bar{\tau}}$, consists of all Boolean functions $f : [n] \rightarrow \{0, 1\}$ such that for some $i \leq \lceil t/2 \rceil$ it holds that $\tau_{2i-1} \leq \Pr_{r \in [n]}[f(r)=1] \leq \tau_{2i}$, where $\tau_{t+1} \stackrel{\text{def}}{=} 1$ for odd t .

We observe that the foregoing testing task, which refers to Boolean functions, can be reduced to testing 0-1 distributions when the tester is given several samples of the tested distribution (i.e., these samples are independently and identically distributed according to the tested distribution).⁵ Specifically, the corresponding class of distributions, denoted $\mathcal{D}_{\bar{\tau}}$, consists of all 0-1 random variables X such that for some $i \leq \lceil t/2 \rceil$ it holds that $\tau_{2i-1} \leq \Pr[X=1] \leq \tau_{2i}$. Indeed, (uniformly selected) queries made to a Boolean function (when testing $\mathcal{B}_{\bar{\tau}}$) correspond to samples obtained from the tested distribution.

2.1 A generic tester and its analysis

A generic tester for $\mathcal{D}_{\bar{\tau}}$ obtains k samples from the tested distribution, where k may (but need not) equal t , and *outputs 1 with probability α_i if exactly i of the samples have value 1*. That is, this generic tester is parameterized by the sequence $\bar{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_k)$. The question, of course, is how many samples do we need (i.e., how is k related to t and/or to other parameters); in other words, whether it is possible to select a $(k+1)$ -long sequence $\bar{\alpha}$ such that the resulting tester, denoted $T_{\bar{\alpha}}$, is a POT for $\mathcal{D}_{\bar{\tau}}$. (We shall show that $k = t$ is sufficient and necessary.) The key quantity to analyze is the probability that this tester (i.e., $T_{\bar{\alpha}}$) accepts a distribution that is 1 with probability q . This accepting probability, denoted $P_{\bar{\alpha}}(q)$, satisfies

$$P_{\bar{\alpha}}(q) = \sum_{i=0}^k \binom{k}{i} \cdot q^i (1-q)^{k-i} \cdot \alpha_i. \quad (1)$$

Indeed, the function $P_{\bar{\alpha}}$ is a degree k polynomial. Noting that 0-1 distributions are determined by the probability that they assume the value 1, we associate these distributions with the corresponding probabilities (e.g., we may say that q is in $\mathcal{D}_{\bar{\tau}}$ and mean that the distribution that is 1 with probability q is in $\mathcal{D}_{\bar{\tau}}$). Thus, $T_{\bar{\alpha}}$ is a POT for $\mathcal{D}_{\bar{\tau}}$ if every distribution that is ϵ -far from $\mathcal{D}_{\bar{\tau}}$ is accepted with probability at most $c - \varrho(\epsilon)$, where $c \stackrel{\text{def}}{=} \min_{q \in \mathcal{D}_{\bar{\tau}}} \{P_{\bar{\alpha}}(q)\}$ and $\varrho : (0, 1] \rightarrow (0, 1]$ is some monotone function.

One necessary condition for the foregoing condition to hold is that for every $i \in [t]$ it holds that $P_{\bar{\alpha}}(\tau_i) = c$, because otherwise a tiny shift from some τ_i to outside $\mathcal{D}_{\bar{\tau}}$ will *not* reduce the value of $P_{\bar{\alpha}}(\cdot)$ below c . Another necessary condition is that $P_{\bar{\alpha}}(\cdot)$ is not a constant function. We first show that there exists a setting of $\bar{\alpha}$ for which both conditions hold (and, in particular, for $k = t$).

Proposition 2.1 (on the existence of $\bar{\tau}$ such that $P_{\bar{\alpha}}$ is “good”): *For every sequence $\bar{\tau} = (\tau_1, \dots, \tau_t)$ such that $0 < \tau_1 < \tau_2 < \dots < \tau_t < 1$, there exists a sequence $\bar{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_t) \in [0, 1]^{t+1}$ such that the following two conditions hold*

1. *For every $i \in [t]$, it holds that $P_{\bar{\alpha}}(\tau_i) = P_{\bar{\alpha}}(\tau_1)$.*
2. *The function $P_{\bar{\alpha}}$ is not a constant function.*

Proof: Fixing any q , we view Eq. (1) as a linear expression in the α_i ’s. Thus, Condition 1 yields a system of $t - 1$ linear equations in the $t + 1$ variables $\alpha_0, \alpha_1, \dots, \alpha_t$. This system is not contradictory, since the uniform vector, denoted \bar{u} , is a solution (i.e., $\bar{\alpha} = ((t+1)^{-1}, \dots, (t+1)^{-1})$ satisfies $P_{\bar{\alpha}}(\tau_i) = (t+1)^{-1}$). Thus, this $(t - 1)$ dimensional system has also a solution that is

⁵In this case, the distance between distributions is merely the standard notion of statistical distance.

linearly independent of \bar{u} . Denoting such a solution by \bar{s} , consider arbitrary $\beta \neq 0$ and γ such that $\beta\bar{s} + \gamma\bar{u} \in [0, 1]^{t+1} \setminus \{0^{t+1}\}$. Note that $\bar{\alpha} \stackrel{\text{def}}{=} \beta\bar{s} + \gamma\bar{u}$ satisfies the linear system and is not spanned by \bar{u} . To establish Condition 2, we show that only vectors $\bar{\alpha}$ that are spanned by \bar{u} yield a constant function $P_{\bar{\alpha}}$. To see this fact, write $P_{\bar{\alpha}}(q)$ as a polynomial in q , obtaining:

$$P_{\bar{\alpha}}(q) = \sum_{d=0}^t (-1)^d \binom{t}{d} \cdot \left(\sum_{i=0}^d (-1)^i \binom{d}{i} \cdot \alpha_i \right) \cdot q^d. \quad (2)$$

Hence, if $P_{\bar{\alpha}}$ is a constant function, then for every $d \in [t]$ it holds that $\sum_{i=0}^d (-1)^i \binom{d}{i} \cdot \alpha_i = 0$, which yields a system of t linearly independent equations in $t+1$ unknowns. Thus, the only solutions to this system are vectors that are spanned by \bar{u} , and the claim follows. \blacksquare

We next prove that the sequence $\bar{\alpha}$ guaranteed by Proposition 2.1 yields a POT for $D_{\bar{\tau}}$.

Theorem 2.2 (analysis of $T_{\bar{\alpha}}$): *For every sequence $\bar{\tau} = (\tau_1, \dots, \tau_t)$ such that $0 < \tau_1 < \tau_2 < \dots < \tau_t < 1$, there exists a sequence $\bar{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_t) \in [0, 1]^{t+1}$ such that $T_{\bar{\alpha}}$ is a POT with linear detection probability for $D_{\bar{\tau}}$.*

Proof: Let $\bar{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_t) \in [0, 1]^{t+1}$ be as guaranteed by Proposition 2.1. Then, the (degree t) polynomial $P_{\bar{\alpha}}$ “oscillates” in $[0, 1]$, while obtaining the value $P_{\bar{\alpha}}(\tau_1)$ on the t points $\tau_1, \tau_2, \dots, \tau_t$ (and only on these points). Thus, for every $i \in [t]$ and all sufficiently small $\epsilon > 0$, exactly one of the values $P_{\bar{\alpha}}(\tau_i - \epsilon)$ and $P_{\bar{\alpha}}(\tau_i + \epsilon)$ is larger than $P_{\bar{\alpha}}(\tau_1)$ (and the other is smaller than it). Without loss of generality, it holds that $P_{\bar{\alpha}}(q) \geq P_{\bar{\alpha}}(\tau_1)$ for every q in $\mathcal{D}_{\bar{\tau}}$ and $P_{\bar{\alpha}}(q) < P_{\bar{\alpha}}(\tau_1)$ otherwise.⁶ Furthermore, we claim that there exists a constant γ such that, for any probability q that is ϵ -far from $\mathcal{D}_{\bar{\tau}}$, it holds that $P_{\bar{\alpha}}(q) \leq P_{\bar{\alpha}}(\tau_1) - \gamma \cdot \epsilon$. This claim can be proved by considering the Taylor expansion of $P_{\bar{\alpha}}$; specifically, expanding $P_{\bar{\alpha}}(q)$ based on the value at τ_i yields

$$P_{\bar{\alpha}}(q) = P_{\bar{\alpha}}(\tau_i) + P'_{\bar{\alpha}}(\tau_i) \cdot (q - \tau_i) + \sum_{j=2}^t \frac{P_{\bar{\alpha}}^{(j)}(\tau_i)}{j!} \cdot (q - \tau_i)^j,$$

where $P'_{\bar{\alpha}}$ is the derivative of $P_{\bar{\alpha}}$ and $P_{\bar{\alpha}}^{(j)}$ is the j^{th} derivative of $P_{\bar{\alpha}}$. By the above, $P'_{\bar{\alpha}}(\tau_i) \neq 0$ (for all $i \in [t]$). Let $v \stackrel{\text{def}}{=} \min_{i \in [t]} \{|P'_{\bar{\alpha}}(\tau_i)|\} > 0$ and $w \stackrel{\text{def}}{=} \max_{i \in [t], j \geq 2} \{|P_{\bar{\alpha}}^{(j)}(\tau_i)|/j!\}$. Then, for all sufficiently small $\epsilon > 0$ (say for $\epsilon \leq \min(1, v)/3w$), it holds that $\sum_{j=2}^t \frac{P_{\bar{\alpha}}^{(j)}(\tau_i)}{j!} \cdot (q - \tau_i)^j$ is upper bounded by $\sum_{j \geq 2} w \cdot \epsilon(v/3w) \cdot (1/3)^{j-2} = v \cdot \epsilon/2$; and so, for every $i \leq \lceil t/2 \rceil$, it holds that $P_{\bar{\alpha}}(\tau_{2i-1} - \epsilon) < P_{\bar{\alpha}}(\tau_{2i-1}) - v \cdot \epsilon/2$ and $P_{\bar{\alpha}}(\tau_{2i} + \epsilon) < P_{\bar{\alpha}}(\tau_{2i}) - v \cdot \epsilon/2$. Using $\gamma = \min(1, v)/3tw$, the claim holds for all $\epsilon \leq 1$. \blacksquare

Sample optimality: We have analyzed a generic tester that uses $k = t$ samples for testing a property parameterized by t thresholds (i.e., $\bar{\tau} = (\tau_1, \dots, \tau_t)$). The proof of Theorem 2.2 implies that using t samples (i.e., $k \geq t$) is necessary, because for $\bar{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_k)$ we need the (non-constant) degree k polynomial $P_{\bar{\alpha}}$ to attain the same value on t points (i.e., the τ_i 's).

The case of $t = 2$: The considerations underlying the proof of Theorem 2.2 imply that in this case (i.e., $t = 2$) the polynomial $P_{\bar{\alpha}}$ is quadratic and equals $P_{\bar{\alpha}}(q) = \alpha_0 - 2(\alpha_0 - \alpha_1) \cdot q + (\alpha_0 - 2\alpha_1 + \alpha_2) \cdot q^2$

⁶Otherwise, use $1 - P_{\bar{\alpha}}$.

(cf. Eq. (2)). Thus, $P_{\bar{\alpha}}$ obtains its maximum at the point $\tau \stackrel{\text{def}}{=} (\tau_1 + \tau_2)/2$, which in turn equals $\frac{2(\alpha_0 - \alpha_1)}{2(\alpha_0 - 2\alpha_1 + \alpha_2)}$. The derivative of $P_{\bar{\alpha}}$ at τ_2 (and likewise $-P'_{\bar{\alpha}}(\tau_1)$) equals

$$\begin{aligned} P'_{\bar{\alpha}}(\tau_2) &= -2(\alpha_0 - \alpha_1) + 2 \cdot (\alpha_0 - 2\alpha_1 + \alpha_2) \cdot \tau_2 \\ &= -2(\alpha_0 - \alpha_1) + 2 \cdot \frac{\alpha_0 - \alpha_1}{\tau} \cdot \tau_2 \\ &= \frac{2}{\tau} \cdot (\tau_2 - \tau) \cdot (\alpha_0 - \alpha_1) \\ &= -\frac{\tau_2 - \tau_1}{\tau} \cdot (\alpha_1 - \alpha_0), \end{aligned}$$

where the second equality is due to $\alpha_0 - \alpha_1 = \tau \cdot (\alpha_0 - 2\alpha_1 + \alpha_2)$ (and the last to $\tau = (\tau_1 + \tau_2)/2$). Thus, we wish to maximize $\alpha_1 - \alpha_0$ subject to $\alpha_0, \alpha_1, \alpha_2 \in [0, 1]$. Using $\alpha_0 - \alpha_1 = \tau \cdot (\alpha_0 - 2\alpha_1 + \alpha_2)$ again, we obtain $\alpha_2 = \frac{(1-\tau)\alpha_0 + (2\tau-1)\alpha_1}{\tau}$. Hence, if $\tau \geq 1/2$, then we may just set $\alpha_0 = 0$ and $\alpha_1 = 1$ (and $\alpha_2 = \frac{2\tau-1}{\tau} \in [0, 1]$). On the other hand, if $\tau \leq 1/2$, then the maximum of $\alpha_1 - \alpha_0$ subject to $\alpha_0, \alpha_1, \alpha_2 \in [0, 1]$ is obtained at $\alpha_2 = 0$, which implies $(1 - \tau)\alpha_0 = (1 - 2\tau)\alpha_1$ (i.e., setting $\alpha_1 = 1$ and $\alpha_0 = \frac{1-2\tau}{1-\tau} \in [0, 1]$). In both cases, letting $\gamma = \max(\tau, 1 - \tau) \in [0, 5, 1]$, we obtain

$$\begin{aligned} -P'_{\bar{\alpha}}(\tau_2) &= \frac{\tau_2 - \tau_1}{\tau} \cdot (\alpha_1 - \alpha_0) \\ &= \frac{\tau_2 - \tau_1}{\gamma} \geq 2(\tau_2 - \tau_1), \end{aligned}$$

which means that distributions that are ϵ -far from $\mathcal{D}_{\bar{\tau}}$ are rejected with probability at least $2(\tau_2 - \tau_1) \cdot \epsilon$.

2.2 Generalization of Theorem 2.2

So far we considered distribution classes $\mathcal{D}_{\bar{\tau}}$ such that $\bar{\tau} = (\tau_1, \dots, \tau_t)$ and $0 < \tau_1 < \tau_2 < \dots < \tau_t < 1$. Recall that this class contains the distribution q if and only if $\tau_{2i-1} \leq q \leq \tau_{2i}$ for some $i \leq t/2$. Here we also allow $\tau_1 = 0$, which corresponds to including in $\mathcal{D}_{\bar{\tau}}$ all distributions X such that $\Pr[X = 1] \leq \tau_2$. In such case we define the sequence $\bar{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_{t-1})$ as guaranteed by Proposition 2.1 such that the polynomial $P_{\bar{\alpha}}$ of degree $t - 1$ is non-constant and satisfies $P_{\bar{\alpha}}(\tau_i) = P_{\bar{\alpha}}(\tau_2)$ for all $i \geq 2$. Analogously we treat the case of t being even and $\tau_t = 1$, which corresponds to including in $\mathcal{D}_{\bar{\tau}}$ all distributions X such that $\Pr[X = 1] \geq \tau_{t-1}$. In both cases the induced tester $T_{\bar{\alpha}}$ is a POT for $\mathcal{D}_{\bar{\tau}}$.

We consider also the case that $\tau_{2i-1} = \tau_{2i}$ (for some i 's); that is, some of the allowed intervals can be collapsed to single points. Consider, for example, the distribution classes $\mathcal{D}_{\tau, \tau}$, for some $\tau \in (0, 1)$. The foregoing design of a POT for $\mathcal{D}_{\tau_1, \tau_2}$ can be easily adapted for the case of $\mathcal{D}_{\tau, \tau}$. Specifically, rather than ensuring that $P_{\bar{\alpha}}(\tau_1) = P_{\bar{\alpha}}(\tau_2)$, we ensure that $P_{\bar{\alpha}}$ obtain a maximum at τ (equiv., $P'_{\bar{\alpha}}(\tau) = 0$), which is actually what we did in the case of $t = 2$ in Section 2.1 for $\tau \stackrel{\text{def}}{=} (\tau_1 + \tau_2)/2$. Thus, we again get $\tau = \frac{\alpha_0 - \alpha_1}{\alpha_0 - 2\alpha_1 + \alpha_2}$, which implies $(\alpha_0, \alpha_1, \alpha_2) = (0, 1, (2\tau - 1)/\tau)$ if $\tau \geq 1/2$ and $(\alpha_0, \alpha_1, \alpha_2) = (\frac{1-2\tau}{1-\tau}, 1, 0)$ otherwise. Next, we approximate $P_{\bar{\alpha}}$ at $\tau + \epsilon$ by $P_{\bar{\alpha}}(\tau) + P_{\bar{\alpha}}^{(2)}(\tau) \cdot \epsilon^2/2$, where $P_{\bar{\alpha}}^{(2)}(\tau) = 2(\alpha_0 - 2\alpha_1 + \alpha_2) = 2(\alpha_0 - \alpha_1)/\tau$. (Note that a much simpler test and analysis is begging in the case that $t = 2$ and $\tau_1 = \tau_2 = 1/2$.)⁷ More generally, we get

⁷In this case, a POT may just select two random samples and accept if and only if exactly one of them assumed the value 1. The probability that this test accepts the distribution q equals $2q(1 - q) = \frac{1}{2} - 2(q - 0.5)^2$.

Theorem 2.3 (Theorem 2.2, generalized): *For every sequence $\bar{\tau} = (\tau_1, \dots, \tau_t) \in [0, 1]^t$ such that $\tau_1 \leq \tau_2 < \tau_3 \leq \tau_4 < \dots < \tau_{t-1} \leq \tau_t$, there exists a sequence $\bar{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_t) \in [0, 1]^{t+1}$ such that $T_{\bar{\alpha}}$ is a POT with quadratic detection probability for $D_{\bar{\tau}}$. Furthermore, if $\tau_{2i-1} = \tau_{2i}$ for every $i \in \lceil [t/2] \rceil$, then $P_{\bar{\alpha}}(q) = P_{\bar{\alpha}}(\tau_1)$ for every q in $D_{\bar{\tau}}$.*

Proof: Let $J = \{j : \tau_{2j-1} = \tau_{2j}\}$. Then, the system of equations regarding the α_i 's contains $t - |J| - 1$ equations that arise from the equalities imposed on the values of $P_{\bar{\alpha}}$ at $t - |J|$ different points and $|J|$ additional equalities that arise from equalities imposed on the values of $P_{\bar{\alpha}}$ at $|J|$ different points. The same considerations (as in the proof of Theorem 2.2) imply the existence of a solution $\bar{\tau}$ such that $P_{\bar{\alpha}}$ is not a constant function, but here the analysis of $P_{\bar{\alpha}}(\tau_j \pm \epsilon)$ depends on whether or not $\lceil j/2 \rceil \in J$: The case of $\lceil j/2 \rceil \notin J$ is handled as in the proof of Theorem 2.2, but the case of $\lceil j/2 \rceil \in J$ relies on the fact that $P_{\bar{\alpha}}^{(2)}(\tau_{2j}) < 0$. ■

2.3 POTs can test only intervals

In this section we show that the only testable classes of Boolean distributions are those defined by a finite collection of intervals in $[0, 1]$, where intervals of length zero (i.e., points) are allowed. This means that the only properties of Boolean distribution that have a POT are those covered in Theorem 2.3.

Theorem 2.4 (characterization of Boolean distributions having a POT): *Let \mathcal{D}_S be a property of Boolean distributions associated with a set $S \subseteq [0, 1]$ such that distribution X is in \mathcal{D}_S if and only if $\Pr[X = 1] \in S$. Then, the property \mathcal{D}_S has a POT if and only if S consists of a finite subset of subintervals of $[0, 1]$.*

Proof: The “if” direction follows from Theorem 2.3. For the other direction, assume that \mathcal{T} is POT for \mathcal{D}_S that makes k queries. Then, for a view $\bar{b} = (b_1, \dots, b_k) \in \{0, 1\}^k$, the tester \mathcal{T} accepts this view with some probability, denoted $\alpha_{\bar{b}} \in [0, 1]$. Note that when testing a distribution X such that $\Pr[X = 1] = p$, the probability of seeing this view is $p^{w(\bar{b})}(1-p)^{k-w(\bar{b})}$, where $w(\bar{b}) = \sum_j b_j$ denotes the number of 1's in \bar{b} . Hence, when given a distribution X such that $\Pr[X = 1] = p$, the acceptance probability of \mathcal{T} on X is

$$\Pr[\mathcal{T} \text{ accepts } X] = \sum_{i=0}^k \left(\sum_{\bar{b} \in \{0,1\}^k : w(\bar{b})=i} \alpha_{\bar{b}} \right) p^i (1-p)^{k-i},$$

which is a polynomial of degree k (in p). Thus, for every $r \in \mathbb{R}$, the set of points $p \in [0, 1]$ on which the value of this polynomial is at least r equals a union of up to $\lceil (k+1)/2 \rceil$ intervals. In particular, this holds for $r = c$, where c denotes the threshold probability of \mathcal{T} , in which case this set of points equals the set S (because \mathcal{T} is POT for \mathcal{D}_S). The theorem follows. ■

3 Graph Properties (in the Adjacency Representation Model)

Symmetric properties of Boolean functions induce graph properties (in the adjacency representation model of [GGR]), and so the statistical properties of the previous section yield analogous properties that refer to the edge densities of graphs. The question addressed in this section is whether the study of two-sided error POT can be extended to “genuine” graph properties. The first property that we consider is degree regularity.

Recall that, in the adjacency matrix model, an N -vertex graph $G = ([N], E)$ is represented by the Boolean function $g : [N] \times [N] \rightarrow \{0, 1\}$ such that $g(u, v) = 1$ if and only if u and v are adjacent in G (i.e., $\{u, v\} \in E$). Distance between graphs is measured in terms of their aforementioned representation (i.e., as the fraction of (the number of) different matrix entries (over N^2)), but occasionally we shall use the more intuitive notion of the fraction of (the number of) edges over $\binom{N}{2}$.

3.1 The class of k -regular graphs

For every function $k : \mathbb{N} \rightarrow \mathbb{N}$, we consider the set $\mathcal{R}^{(k)} = \cup_{N \in \mathbb{N}} \mathcal{R}_N^{(k)}$ such that $\mathcal{R}_N^{(k)}$ is the set of all $k(N)$ -regular N -vertex graphs. That is, $G \in \mathcal{R}_N^{(k)}$ if and only if G is a simple N -vertex graph in which each vertex has degree $k(N)$. Clearly, $\mathcal{R}^{(k)}$ has no one-sided error POT, provided that $0 < k(N) < N - 1$ (cf. [GR09b, Sec. 4]).⁸ In contrast, we show that it has a two-sided error POT.

Theorem 3.1 (a POT for $\mathcal{R}^{(k)}$): *For every function $k : \mathbb{N} \rightarrow \mathbb{N}$ such that $k(N) = \kappa N$ for some fixed constant $\kappa \in (0, 1)$, the property $\mathcal{R}^{(k)}$ has a two-sided error POT. Furthermore, all graphs in $\mathcal{R}^{(k)}$ are accepted with equal probability.*

Proof: We may assume that $N \cdot k(N)$ is an even integer (since otherwise the test may reject without making any queries). On input N and oracle access to an N -vertex graph $G = ([N], E)$, the tester sets $\tau = k(N)/N = \kappa$ and proceeds as follows.

1. Selects uniformly a vertex $s \in [N]$ and consider the Boolean function $f_s : [N] \rightarrow \{0, 1\}$ such that $f_s(v) = 1$ if and only if $\{s, v\} \in E$.
2. Invokes the POT of Theorem 2.3 to test whether the function f_s has density τ ; that is, it tests whether the random variable X_s defined uniformly over $[N]$ such that $X_s(v) = f_s(v)$ is in the class $\mathcal{D}_{\tau, \tau}$.

Recall that this POT takes two samples of X_s and accepts with probability α_i when seeing i values of 1. (The values of $(\alpha_0, \alpha_1, \alpha_2)$ are set based on τ .)⁹

The implementation of Step 2 calls for taking two samples of X_s , which amounts to selecting uniformly two vertices and checking whether or not each of them neighbors s . Thus, we make two queries to the graph G .

Turning to the analysis of the foregoing test, let $P(q)$ denote the probability that the POT invoked in Step 2 accepts a random variable X such that $\Pr[X = x] = q$. Then, the probability that our graph tester accepts the graph G equals

$$\frac{1}{N} \cdot \sum_{s \in [N]} P(d_G(s)/N), \tag{3}$$

where $d_G(v)$ denotes the degree of vertex v in G . Thus, every $k(N)$ -regular N -vertex graph G is accepted with probability $P(\tau)$. As we shall show, the following claim (which improves over a similar claim in [GGR, Apdx D]) implies that every graph that is ϵ -far from $\mathcal{R}_N^{(k)}$ is accepted with probability $P(\tau) - \Omega(\epsilon^2)$.

⁸Specifically, the characterization in [GR09b, Thm. 4.7] implies that it suffices to show that $\mathcal{R}^{(k)}$ is not a subgraph freeness property. Assume, without loss of generality that $k(N) > N/2$. Then, the subgraphs disallowed in $\mathcal{R}^{(k)}$ cannot contain a clique, and it follows that the N -vertex clique is in $\mathcal{R}^{(k)}$, which contradicts $k(N) < N - 1$.

⁹Recall (cf. Section 2.2) that we may use the setting outlined at the end of Section 2.1: That is, we set $\alpha_1 = 1$, and if $\tau \leq 1/2$, then $\alpha_2 = 0$ and $\alpha_0 = \frac{1-2\tau}{1-\tau}$, and otherwise $\alpha_0 = 1$ and $\alpha_2 = \frac{2\tau-1}{\tau}$.

Claim 3.1.1 *If $\sum_{v \in [N]} |d_G(v) - k(N)| \leq \epsilon' \cdot N^2$, then G is $6\epsilon'$ -close to $\mathcal{R}_N^{(k)}$.*

The proof of Claim 3.1.1 is presented in Appendix A.1. Note that the claim is non-trivial, since it asserts that small local discrepancies (in the vertex degrees) imply small distance to regularity. The converse is indeed trivial.

Using Claim 3.1.1, we infer that if G is ϵ -far from $\mathcal{R}_N^{(k)}$, then $\sum_{v \in [N]} |d_G(v) - k(N)| > \epsilon \cdot N^2/6$. On the other hand, by Theorem 2.3 (or the analysis of the case $t = 2$ that precedes it), we have, for some $\gamma > 0$,

$$\begin{aligned} \frac{1}{N} \cdot \sum_{s \in [N]} \mathbb{P}(d_G(s)/N) &\leq \frac{1}{N} \cdot \sum_{s \in [N]} (\mathbb{P}(\tau) - \gamma \cdot ((d_G(s) - k(N))/N)^2) \\ &= \mathbb{P}(\tau) - \frac{\gamma}{N^2} \cdot \frac{1}{N} \cdot \sum_{s \in [N]} (d_G(s) - k(N))^2 \\ &\leq \mathbb{P}(\tau) - \frac{\gamma}{N^2} \cdot \left(\frac{\sum_{s \in [N]} |d_G(s) - k(N)|}{N} \right)^2 \end{aligned}$$

where the last inequality follows by the Cauchy-Schwarz inequality. Now, using $\sum_{v \in [N]} |d_G(v) - k(N)| > \epsilon \cdot N^2/6$, we conclude that G is accepted with probability at most $\mathbb{P}(\tau) - \gamma \cdot (\epsilon/6)^2$. The theorem follows. \blacksquare

3.2 Other regular graph properties

The two-sided error POT guaranteed by Theorem 3.1 can be combined with one-sided POT for other graph properties to yield two-sided error POTs for the intersection. This combination is possible whenever the two properties behave nicely with respect to intersection in the sense that being close to both properties (i.e., to both $\mathcal{R}^{(k)}$ and Π) implies being close to their intersection (i.e., to $\mathcal{R}^{(k)} \cap \Pi$). Recall that, as pointed out in [GGR], in general it may not be that case that objects that are close to two properties are also close to their intersection.

Theorem 3.2 (a POT for $\mathcal{R}^{(k)} \cap \Pi$): *Let Π be a graph property that has a one-sided error and $k(N) = \kappa N$ for some fixed constant $\kappa \in (0, 1)$. Suppose that there exists a monotone function $F : (0, 1] \rightarrow (0, 1]$ such that if G is δ -close to both Π and $\mathcal{R}^{(k)}$ then G is $F(\delta)$ -close to $\Pi \cap \mathcal{R}^{(k)}$. Then, $\Pi \cap \mathcal{R}^{(k)}$ has a two-sided error POT.*

Note that the condition made in Theorem 3.2 may not hold in general. For example, consider Π that consists of all bicliques as well as all graphs that each consist of two isolated cliques. Then, for $k(N) = N/2$, it holds that $\mathcal{R}^{(k)} \cap \Pi$ consists of N -vertex bicliques with $N/2$ vertices on each side, and so the graph G consisting of two $N/2$ -vertex cliques is 0.49-far from $\mathcal{R}^{(k)} \cap \Pi$. On the other hand, G is in Π and is $1/N$ -close to $\mathcal{R}^{(k)}$ (by virtue of adding a perfect matching between the two cliques). In contrast, it can be shown that the condition in Theorem 3.2 holds with respect to $k(N) = 2N/3$ and the set Π consisting of all complete tripartite graphs (see either Proposition 3.3 or Proposition 3.4).

Proof: On input N and oracle access to an N -vertex graph $G = ([N], E)$, the tester proceeds as follows (while assuming that κN is an integer and κN^2 is even).¹⁰

¹⁰Otherwise, the tester rejects upfront, since no N -vertex graph can be κN -regular.

1. Invokes the POT for $\mathcal{R}^{(k)}$ and reject if it halts while rejecting. Otherwise, proceeds to the next step.
2. Invokes the POT for Π and halts with its verdict.

The analysis relies crucially on the fact that the (two-sided error) POT for $\mathcal{R}^{(k)}$ accepts any graph in $\mathcal{R}^{(k)}$ with the same probability, denoted c . It follows that any N -vertex graph in $\Pi \cap \mathcal{R}^{(k)}$ is accepted with probability $c \cdot 1 = c$. Next, we show that graphs that are far from $\Pi \cap \mathcal{R}^{(k)}$ are accepted with probability that is significantly smaller than c .

Let G be a graph that is δ -far from $\Pi \cap \mathcal{R}^{(k)}$. Then, by the hypothesis regarding Π and $\mathcal{R}^{(k)}$, either G is $F^{-1}(\delta)$ -far from Π or G is $F^{-1}(\delta)$ -far from $\mathcal{R}^{(k)}$. In the first case, G is accepted with probability at most $c \cdot (1 - \varrho_1(F^{-1}(\delta)))$, where ϱ_1 is the detection probability function of the one-sided error POT for Π . Note that we rely on the fact that the (two-sided error) POT for $\mathcal{R}^{(k)}$ accepts any graph with probability at most c . In the second case (i.e., G is far from $\mathcal{R}^{(k)}$), it holds that G is accepted with probability at most $c - \varrho_2(F^{-1}(\delta))$, where ϱ_2 is the detection probability function of the two-sided error POT for $\mathcal{R}^{(k)}$. The claim follows. ■

Corollaries. One natural question is which properties Π satisfy the condition of Theorem 3.2 and what properties arise from their intersection with $\mathcal{R}^{(k)}$. Recall that by the characterization result of [GR09b], the property Π must be defined in terms of subgraph freeness (since only such properties have a one-sided error POT). However, the intersection $\Pi \cap \mathcal{R}^{(k)}$ may not be easy to characterize in general. Furthermore, as indicated above, some subgraph freeness properties satisfy the condition of Theorem 3.2 while others do not. We consider this issue in the context of two specific classes of properties, studied in [GR09a]. The first class consists of all complete t -partite graphs, where a graph is called **complete t -partite** if its vertex set can be partitioned into t (independent) sets such that two vertices are connected by an edge if and only if they belong to different sets.

Proposition 3.3 (on regular complete t -partite graphs): *Let $t \geq 2$ be an integer and $k(N) = (t - 1)N/t$.*

1. *The set of k -regular complete t -partite graphs equals to the set of complete t -partite graphs in which each part (i.e., independent set) has density $1/t$.*
2. *If a graph $G = ([N], E)$ is δ -close to both the set of complete t -partite graphs and to $\mathcal{R}^{(k)}$, then G is $O(\sqrt{\delta})$ -close to some k -regular complete t -partite graph.*

Thus, the set of k -regular complete t -partite graph has a two-sided error POT.

Proof: Let Π denote the set of t -partite graphs. First, we show that $\mathcal{R}^{(k)} \cap \Pi$ equals the set of all k -regular complete t -partite graphs, which we denote by Π' . This follows by considering the t -partition (V_1, \dots, V_t) of an arbitrary N -vertex graph in Π , and observing that the degree condition implies that for every $i \in [t]$ such that $V_i \neq \emptyset$ it holds that $\sum_{j \neq i} |V_j| = k(N)$. Thus, for every such $i \in [t]$ it holds that $|V_i| = N/t$, and Item 1 follows.

Turning to the proof of Part 2, we note that if G is δ -close to both Π and $\mathcal{R}^{(k)}$, then there exists $G' \in \Pi$ that is 2δ -close to $\mathcal{R}^{(k)}$. Let I_1, \dots, I_t be the partition of G' to t independent sets such that there is a complete bipartite graph between each two I_j 's. Then, we have $\sum_{i \neq j} |I_i| \cdot |I_j| \geq k(N)N - 4\delta N^2$, which implies $\sum_{i \in [t]} x_i^2 \leq (1/t) + 4\delta$, where $x_i = |I_i|/N$. It follows that $\sum_{i \in [t]} (x_i - (1/t))^2 \leq 4\delta$, and thus G' is $O(\sqrt{\delta})$ -close to Π' . ■

The second class, studied in [GR09a], is the class of super-cycle collections, where a super-cycle (of length ℓ) is a graph consisting of a sequence of disjoint sets of vertices, called clouds, such that two vertices are connected if and only if they reside in neighboring clouds (i.e., denoting the ℓ clouds by $S_0, \dots, S_{\ell-1}$, vertices $u, v \in \bigcup_{i \in \{0, 1, \dots, \ell-1\}} S_i$ are connected by an edge if and only if for some $i \in \{0, 1, \dots, \ell-1\}$ and $j \in \{i-1 \bmod \ell, i+1 \bmod \ell\}$ it holds that $u \in S_i$ and $v \in S_j$). 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 ℓ by $\mathcal{SC}_\ell\mathcal{C}$. As is shown in the next two propositions, for every $\ell \leq 3$, there is a dichotomy in the behavior of the set $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$: For some integers t and $k(N) = 2N/t\ell$, the sets $\mathcal{SC}_\ell\mathcal{C}$ and $\mathcal{R}^{(k)}$ satisfy the conditions of Theorem 3.2, whereas for the remaining values of t the conditions of Theorem 3.2 are not satisfied.

Proposition 3.4 (on $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$ and testing it, for some values of ℓ, t and $k(N) = 2N/t\ell$): *Let $\ell \geq 3$ be an integer and $t \in \mathbb{T}_\ell$, where*

$$\mathbb{T}_\ell = \begin{cases} \{1, 2, 3\} & \text{if } \ell \equiv 1 \pmod{2} \\ \{1\} & \text{if } \ell \equiv 2 \pmod{4} \\ \mathbb{N} & \text{if } \ell \equiv 0 \pmod{4} \end{cases} \quad (4)$$

and $k(N) = 2N/t\ell$. Then:

1. *The set $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$ equals the set of graphs that consists of t super-cycles of length ℓ , each containing N/t vertices, such that clouds that are at distance four apart have equal size. Furthermore, if $\ell \not\equiv 0 \pmod{4}$, then each cloud has size $N/t\ell$.*
2. *If a graph $G = ([N], E)$ is δ -close to both $\mathcal{SC}_\ell\mathcal{C}$ and $\mathcal{R}^{(k)}$, then G is $O(\sqrt{\delta})$ -close to $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$, where the hidden constant depends polynomially on $t\ell$.*

Thus, $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$ has a two-sided error POT.

The proof of Proposition 3.4 is quite tedious and is deferred to Appendix A.2.

Proposition 3.5 (on $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$ in other cases): *Let $\ell \geq 3$ and $t \in \mathbb{N} \setminus \mathbb{T}_\ell$, where \mathbb{T}_ℓ is as in Eq. (4). Then, for any integer $k(N) = 2N/t\ell$, there exists an N -vertex graph in $\mathcal{R}^{(k)}$ that is $O(1/N)$ -close to $\mathcal{SC}_\ell\mathcal{C}$ but $\Omega(1)$ -far from $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$.*

Indeed, Proposition 3.5 is non-trivial only in the case that $\ell \not\equiv 0 \pmod{4}$. We stress that Proposition 3.5 does not assert that in certain cases $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$ has no POT, but rather that the conditions stated in Theorem 3.2 are not satisfied (and so the approach suggested by its proof will not work). The proof of Proposition 3.5 refers to some elements of the proof of Proposition 3.4, and thus is also deferred to Appendix A.2.

3.3 The class of regular graphs

We consider the class \mathcal{REG} of regular graphs. Note that this class strictly contains the classes considered in Section 3.1, since we make no restriction on the degrees. (Still, this does not mean that testing \mathcal{REG} is either easier or harder than testing these subclasses.) Clearly, \mathcal{REG} has no one-sided error POT. In contrast, we show that it has a two-sided error POT.

Theorem 3.6 (a POT for \mathcal{REG}): *The class \mathcal{REG} has a two-sided error POT. Furthermore, all graphs in \mathcal{REG} are accepted with equal probability.*

Recall that a standard tester of regularity can be obtained by estimating the degrees of random vertices (cf. [GGR, Prop. 10.2.1.3]), where these estimations are related to the proximity parameter. However, such good approximations are not possible in the context of proximity oblivious testing. Still, as in Section 3.1, crude approximations (which are obtained by a constant number of queries) turn out to be sufficiently good. Specifically, we construct a POT that picks two random vertices in the given graph, and checks that these two vertices have the same degree in a *proximity oblivious* manner. This checking is reduced to the problem of testing equality between two Boolean distributions, where in the reduction the distributions correspond to the density of the neighbors of each of the chosen vertices.¹¹ We show first that the task of testing that two Boolean distributions are equal can be tested in a proximity oblivious manner (and will return to Theorem 3.6 later).

Proposition 3.7 (a POT for \mathcal{EQ}): *Let $\mathcal{EQ} = \{(P, Q) : \Pr[P=1] = \Pr[Q=1]\}$ be the class that consists of pairs of equal Boolean distributions, and let the distance of a pair (P, Q) from \mathcal{EQ} be defined as*

$$\text{dist}((P, Q), \mathcal{EQ}) = |\Pr[P=1] - \Pr[Q=1]|.$$

Then, the property \mathcal{EQ} has a two-sided error POT. Given two distributions, the tester makes two queries to each of them, and has quadratic detection probability. Moreover, all pairs of equal distributions are accepted with the same probability.

As shown below (see Proposition 3.8), the property \mathcal{EQ} has no POT that (always) makes less than two queries to one of the two distributions.

Proof: Following the receipt of Section 2.1, the design of the desired POT calls for choosing a sequence of $\alpha_{(i,j)}$'s, where $\alpha_{(i,j)}$ represents the acceptance probability when seeing i ones in the (2-element) sample of P and j ones in the (2-element) sample of Q . The corresponding acceptance probability is a polynomial of individual degree 2 in p and q , where $p = \Pr[P=1]$ and $q = \Pr[Q=1]$. The goal is thus to choose the $\alpha_{(i,j)}$'s such that this polynomial evaluates to c for every (p, q) such that $p = q$, and evaluates to $c - \Omega((p - q)^2)$ otherwise.

Let $c, \delta \in (0, 1)$ be two parameters such that $c - 2\delta, c + \delta \in [0, 1]$. Indeed, we may chose $c = 0.5$ and $\delta = 0.25$. For every $(i, j) \in \{0, 1, 2\}^2$ define

$$\alpha_{(i,j)} = \begin{cases} c - 2\delta & \text{if } (i, j) \in \{(0, 2), (2, 0)\} \\ c + \delta & \text{if } (i, j) = (1, 1) \\ c & \text{otherwise} \end{cases} \quad (5)$$

Given a pair of distributions (P, Q) , the tester, denoted T , proceeds as following:

1. Make two queries to each distribution. Denote by i the number of ones obtained from P , and denote by j the number of ones obtained from Q .
2. Accept with probability $\alpha_{(i,j)}$.

Letting $p = \Pr[P=1]$ and $q = \Pr[Q=1]$, the acceptance probability of the tester is

$$\Pr[T \text{ accepts } (P, Q)] = \sum_{i,j \in \{0,1,2\}} \alpha_{(i,j)} \cdot \binom{2}{i} p^i (1-p)^{2-i} \cdot \binom{2}{j} q^j (1-q)^{2-j}. \quad (6)$$

¹¹This reduction is analogous to the proof of Theorem 3.1 in which we check that the degree of a vertex equals to some $k \in \mathbb{N}$ that is fixed and is known to the tester.

Note that almost all the $\alpha_{(i,j)}$'s in Eq. (5) are equal to c , with the exception of $\alpha_{(2,0)} = \alpha_{(0,2)}$ and $\alpha_{(1,1)}$. Thus, the tester ‘‘penalizes’’ a highly unbalanced view (i.e., $(2, 0)$ or $(0, 2)$) and ‘‘awards’’ a balanced view. Indeed, plugging in the parameters in Eq. (6), we get

$$\begin{aligned} \Pr[T \text{ accepts } (P, Q)] &= c - (p^2(1-q)^2 + q^2(1-p)^2) \cdot 2\delta + 4p(1-p)q(1-q) \cdot \delta \\ &= c - 2\delta \cdot (p^2(1-q)^2 + q^2(1-p)^2 - 2p(1-p)q(1-q)) \\ &= c - 2\delta \cdot (p-q)^2 \end{aligned}$$

The proposition follows. \blacksquare

Proof of Theorem 3.6. Given an N -vertex graph $G = ([N], E)$, the tester proceeds as follows.

1. Select uniformly two vertices $v_1, v_2 \in [N]$ and consider the Boolean functions $f_{v_1} : [N] \rightarrow \{0, 1\}$ and $f_{v_2} : [N] \rightarrow \{0, 1\}$ such that $f_{v_i}(w) = 1$ if and only if $\{v_i, w\} \in E$.
2. Invoke the POT of Proposition 3.7 to test whether the function f_{v_1} and f_{v_2} have the same density, and act according to its answer. That is, test whether the random variables X_{v_1} and X_{v_2} defined over $[N]$ such that $X_{v_i}(w) = f_{v_i}(w)$ are equal (i.e., whether the pair (X_{v_1}, X_{v_2}) is in \mathcal{EQ}).

Recall that this POT accepts all equal pairs of distributions with the same probability c (which is an absolute constant independent of $d_G(v_1)$ and $d_G(v_2)$).

In the implementation of Step 2 the tester takes two samples of X_{v_1} and two sample of X_{v_2} , which amounts to selecting uniformly four vertices, and checking whether that the first two are adjacent to v_1 and the last two are adjacent to v_2 . Thus, we make four queries to the graph G .

Clearly, if G is a regular graph, then for every pair of vertices (v_1, v_2) , chosen in Step 1, the POT of Step 2 will accept with the same probability c (which is also independent of the degree of G). Suppose now that a graph $G = ([N], E)$ is accepted with probability $c - \epsilon$. Then, by an averaging argument, there are $(1 - \epsilon)N^2$ pairs of vertices (v_1, v_2) such that when these vertex-pairs are chosen in Step 1, then Step 2 accepts with probability at least $c - \epsilon$. Thus, there exists a vertex $v_1 \in [N]$ and a subset $V_2 \subseteq [N]$ of size at least $(1 - \epsilon)N$ such that for every vertex $v_2 \in V_2$ the acceptance probability of Step 2 when applied on the pair (v_1, v_2) is at least $c - \epsilon$. Proposition 3.7 implies that for every such vertex v_2 it holds that $|\Pr[X_{v_1} = 1] - \Pr[X_{v_2} = 1]| = O(\sqrt{\epsilon})$, and therefore $|d_G(v_2) - d_G(v_1)| < O(\sqrt{\epsilon} \cdot N)$.

Let $K = d_G(v_1)$ be the degree of v_1 . We show that the graph is close to being K -regular. Indeed, using the fact that $|V_2| \geq (1 - \epsilon)N$ and for every $v_2 \in V_2$ it holds that $|d_G(v_2) - K| < O(\sqrt{\epsilon}N)$, we have

$$\sum_{v \in [N]} |d_G(v) - K| = \sum_{v \in V_2} |d_G(v) - K| + \sum_{v \in [N] \setminus V_2} |d_G(v) - K| < O(\sqrt{\epsilon} \cdot N^2).$$

Therefore, by applying Claim 3.1.1 we get that the graph G is $O(\sqrt{\epsilon})$ -close to being K -regular. This completes the proof of Theorem 3.6. \blacksquare

Comment. A result analogous to Theorem 3.2 can be proved in the current context. That is, $\Pi \cap \mathcal{RE}\mathcal{G}$ has a two-sided error POT if Π has a one-sided error POT and there exists a monotone function $F : (0, 1] \rightarrow (0, 1]$ such that every graph that is δ -close to both Π and $\mathcal{RE}\mathcal{G}$ is $F(\delta)$ -close to $\Pi \cap \mathcal{RE}\mathcal{G}$.

Proposition 3.8 *The property \mathcal{EQ} has no two-sided error POT that always makes at most one query to the first (resp., second) distribution.*

Indeed, Proposition 3.8 does not rule out the possibility that a POT for \mathcal{EQ} can *sometimes* (i.e., depending on its coin tosses) make less than two queries to the first (resp., second) distribution. In fact, there exists a POT for \mathcal{EQ} that always makes two queries (in total) such that with probability $\frac{1}{4}$ it makes two queries to the first (resp., second) distribution, and otherwise it makes a single query to each distribution.¹²

Proof: Suppose, without loss of generality, that there exists a POT, denoted T , that always makes t queries to distribution P and a single query to distribution Q . Let us denote the threshold probability of T by c . Note that, in general, without loss of generality, the activities of any POT for \mathcal{EQ} depend on the number of ones that it sees among the P samples and among the Q samples; that is, when it sees i ones in the sample of P and j ones in the sample of Q it accepts with probability $\alpha_{(i,j)}$ (e.g., let $\alpha_{(i,j)}$ be the average acceptance probability taken over all the corresponding cases). Now, denoting (again) $p = \Pr[P=1]$ and $q = \Pr[Q=1]$, the acceptance probability of T equals

$$A(p, q) \stackrel{\text{def}}{=} \Pr[T \text{ accepts } (P, Q)] = \sum_{i \in \{0, 1, \dots, t\}} \binom{t}{i} p^i (1-p)^{t-i} \cdot (q\alpha_{(i,1)} + (1-q)\alpha_{(i,0)}). \quad (7)$$

Letting $\delta = p - q$, we have $A(p, q) = B(p) + \delta \cdot D(p)$, where

$$B(p) \stackrel{\text{def}}{=} \sum_{i \in \{0, 1, \dots, t\}} \binom{t}{i} p^i (1-p)^{t-i} \cdot (p\alpha_{(i,1)} + (1-p)\alpha_{(i,0)})$$

and $D(p) \stackrel{\text{def}}{=} \sum_{i \in \{0, 1, \dots, t\}} \binom{t}{i} p^i (1-p)^{t-i} \cdot (\alpha_{(i,1)} - \alpha_{(i,0)})$.

The following observations only rely on the fact that the acceptance probability $A(p, p + \delta)$ is linear in δ .

1. For every p it holds that $B(p) \geq c$, because (P, P) must be accepted with probability at least c (i.e., because $A(p, p) \geq c$).
2. For every p it holds that $B(p) \leq c$, because otherwise for some δ such that $|\delta| > 0$ the no-instance (P, Q) is accepted with probability at least $B(p) - |\delta| > c$ (using $|D(p)| \leq 1$), which violates $A(p, q) < c$.
3. For every p it holds that $D(p) \neq 0$, because otherwise for some δ such that $|\delta| > 0$ the no-instance (P, Q) is accepted with probability $B(p) = c$.

Using (3), we note that for every $p \in (0, 1)$ and $\epsilon > 0$ such that $p - \epsilon, p + \epsilon \in [0, 1]$, either $(p, p - \epsilon)$ or $(p, p + \epsilon)$ is accepted with probability greater than $B(p) = c$ (since either $\epsilon D(p) > 0$ or $-\epsilon D(p) > 0$). This violates the requirements from a POT for \mathcal{EQ} , and the proposition follows. \blacksquare

3.4 Bounded density of induced copies

We now turn to a different type of graph properties; specifically, to sets of graphs in which a fixed graph appears as an induced subgraph for a bounded number of times.

¹²This tester can be obtained by a reduction to testing a property of an auxiliary 4-valued distribution R such that $R = (1, P)$ with probability $\frac{1}{2}$ and $R = (2, Q)$ otherwise. Testing whether P equals Q reduces to testing whether $\Pr[R = (1, 1)] = \Pr[R = (2, 1)]$, whereas (by Corollary 5.9) this property has a POT that uses two samples. We comment that a similar POT can be obtained for testing the equality of $t \geq 2$ distributions.

Fixing any n -vertex graph H , denote by $\rho_H(G)$ the density of H as a subgraph in G ; that is, $\rho_H(G)$ is the probability that a random sample of n vertices in G induces the subgraph H . For any graph H and $\tau \in [0, 1]$, we consider the graph property $\Pi_{H,\tau} \stackrel{\text{def}}{=} \{G : \rho_H(G) \leq \tau\}$; in particular, $\Pi_{H,0}$ is the class of H -free graphs. Alon *et al.* [AFKS] showed that, for some monotone function $F_n : (0, 1] \rightarrow (0, 1]$ if G is δ -far from the class of H -free graphs, then $\rho_H(G) > F_n(\delta)$. Here we provide a much sharper bound for the case of $\tau > 0$ (while using an elementary proof).¹³

Theorem 3.9 (distance from $\Pi_{H,\tau}$ yields $\rho_H > \tau$): *For every n -vertex graph H and $\tau > 0$, if $G = ([N], E)$ is δ -far from $\Pi_{H,\tau}$, then $\rho_H(G) > (1 + (\delta n/3)) \cdot \tau$, provided that $\delta > 6/N$.*

It follows that $\Pi_{H,\tau}$ has a two-sided error POT, which just inspects a random sample of n vertices and checks whether the induced subgraph is isomorphic to H . This POT accepts a graph in $\Pi_{H,\tau}$ with probability at least $1 - \tau$, whereas it accepts any graph that is δ -far from $\Pi_{H,\tau}$ with probability at most $1 - \tau - (\tau n/3) \cdot \delta$ (if $\delta > 6/N$, and with probability at most $1 - \tau - (\delta/6)^n$ otherwise).

Proof: Let us consider first the case that H contains no isolated vertices. Setting $G_0 = G$, we proceed in iterations while preserving the invariant that G_i is $(\delta - 2i/N)$ -far from $\Pi_{H,\tau}$. In particular, we enter the i^{th} iteration with a graph G_{i-1} not in $\Pi_{H,\tau}$, and infer that G_{i-1} contains a vertex, denoted v_i , that participates in at least $M \stackrel{\text{def}}{=} \tau \cdot \binom{N-1}{n-1}$ copies of H . We obtain a graph G_i that is $(N-1)/\binom{N}{2}$ -close to G_{i-1} by omitting from G_{i-1} all edges incident at v_i . We stress that the M copies of H counted in the i^{th} iterations are different from the copies counted in the prior $i-1$ iterations, because all copies counted in the i^{th} iteration touch the vertex v_i and do not touch the vertices v_1, \dots, v_{i-1} , since the latter vertices are isolated in G_{i-1} (whereas H contains no isolated vertices). Also note that the copies of H counted in the i^{th} iteration also occur in G , since they contain no vertex pair on which G_{i-1} differs from G . Thus, after $t \stackrel{\text{def}}{=} \lfloor \delta N/2 \rfloor$ iterations, we obtain a graph $G_t \notin \Pi_{H,\tau}$, which contain $\tau \cdot \binom{N}{n}$ copies of H that are disjoint from the $t \cdot M$ copies of H counted in the t iterations. It follows that

$$\begin{aligned} \rho_H(G) &\geq \tau + t \cdot \frac{M}{\binom{N}{n}} \\ &= \tau + \lfloor \delta N/2 \rfloor \cdot \frac{n \cdot \tau}{N} \\ &> \tau + \left(\frac{\delta n}{2} - \frac{n}{N} \right) \cdot \tau \end{aligned}$$

and the claim follows (using $\delta > 6/N$). Recall, however, that the foregoing relies on the hypothesis that H has no isolated vertices. If this hypothesis does not hold, then the complement graph of H has no isolated vertices, and we can proceed analogously. In other words, if H has an isolated vertex, then no vertex in H is connected to all the other vertices. In this case, we consider the graph G_i obtained from G_{i-1} by connecting the vertex v_i to all other vertices in the graph. Also in this case, H -copies in G_i cannot touch v_1, \dots, v_{i-1} (this time because each vertex in v_1, \dots, v_{i-1} is connected to all vertices in G_{i-1}), and we can proceed as before. ■

3.5 Towards a characterization

The foregoing results beg the question of characterizing the class of graph properties that have a two-sided error POT and also suggest that such a characterization may be related to the densities

¹³In contrast, the proof of Alon *et al.* [AFKS] relies on Szemeredy's Regularity Lemma.

in which various fixed-size graphs appear as induced subgraphs in the graph. In the current section we pursue these ideas.

Recall that, for an n -vertex graph H , we denote by $\rho_H(G)$ the density of H as a subgraph in G (i.e., $\rho_H(G)$ equals the probability that a random sample of n vertices in G induces the subgraph H). We consider graph properties that are each parameterized by a sequence of weights $\bar{w} = (w_H)$ and by $b \in [0, 1]$, where $w_H \in [0, 1]$ for each n -vertex (unlabeled) graph H . The corresponding graph property is denoted $\Pi_{\bar{w}, b}$, and a graph G is in $\Pi_{\bar{w}, b}$ if and only if $\sum_H w_H \cdot \rho_H(G) \leq b$.

Note that the case of $b = 0$ corresponds to \mathcal{F} -freeness for $\mathcal{F} = \{H : w_H > 0\}$. More generally, if for every H it holds that $w_H \geq b$, then $\Pi_{\bar{w}, b}$ equals the set of \mathcal{F} -free graphs, where $\mathcal{F} = \{H : w_H > b\}$. Another interesting case is where $w_{H_0} = 1$ for a unique graph H_0 and $w_H = 0$ otherwise (i.e., for every $H \neq H_0$): In this case the property $\Pi_{\bar{w}, b}$ corresponds to having an H_0 -density that does not exceed b (i.e., in this case $G \in \Pi_{\bar{w}, b}$ if and only if $\rho_{H_0}(G) \leq b$, which is the case studied in Section 3.4).

In the rest of this section, we shall discard the case of a uniform sequence \bar{w} (i.e., $w_H = w$ for some w and all H 's), since in this case the property is trivial. We conjecture that, for any $b > 0$ and \bar{w} , the property $\Pi_{\bar{w}, b}$ has a two-sided error POT, but we are only able to establish it in special cases (see Theorem 3.11). On the other hand, we show that any graph property having a two-sided error POT is essentially of the foregoing type. The latter statement requires some clarification.

Recall that it was shown in [GR09b, Thm. 4.7] that a graph property has a one-sided error POTs if and only if it is a subgraph freeness property. However, the equivalence is not to \mathcal{F} -freeness where \mathcal{F} is a fixed set of forbidden subgraphs, but rather to an infinite sequence of subgraph freeness properties that correspond to different graph sizes. Specifically, it was shown that $\Pi = \cup_N \Pi_N$ has a one-sided error POT if there exists a constant n and an infinite sequence $(\mathcal{F}_N)_{N \in \mathbb{N}}$ such that for every N it holds that (1) all graphs in \mathcal{F}_N are of size n , and (2) Π_N equals the set of all N -vertex \mathcal{F}_N -free graphs.

Note that in the latter context there are only finitely many possible sets \mathcal{F}_N , whereas in our context there are infinitely many possible sequences $\bar{w} = (w_H)$ (and ditto b 's). In other words, for every fixed N , the number of possible properties of N -vertex graphs that arises from such $2^{\binom{n}{2}+1}$ -long sequences depends on N (and is not upper bounded by a function of n). For example, for every $m(N) \in \{0, 1, \dots, \binom{N}{2}\}$, we may consider the property of N -vertex graphs having at most $m(N)$ edges.¹⁴

Another difficulty that arises regarding the foregoing properties is that, in general, it is not clear how the following notions of violating the property $\Pi_{\bar{w}, b}$ are related:

1. The graph G is δ -far from $\Pi_{\bar{w}, b}$; that is, for every $G' \in \Pi_{\bar{w}, b}$ it holds that G and G' differ on at least δ fraction of vertex pairs.
2. The graph G satisfies $\sum_H w_H \cdot \rho_H(G) \geq b + \epsilon$.

Indeed, $\delta > 0$ if and only if $\epsilon > 0$ (since $G \notin \Pi_{\bar{w}, b}$ if and only if $\sum_H w_H \cdot \rho_H(G) > b$). It also holds that $\epsilon \leq \binom{n}{2} \cdot \delta$ (since the probability that a random sample of n vertices hits a pair of vertices that differs in two graph can be upper bounded in term of the distance between the graphs). But what is missing is a general bound in the opposite direction, whereas we do have such bounds in special cases (e.g., either $b = 0$ or $|\{H : w_H > 0\}| = 1$, see Section 3.4).¹⁵ In light of this state of affairs, a first step towards a characterization is provided by the following result (where $V(H)$ denotes the vertex set of H).

¹⁴This property can be represented by setting $b = m(N)/\binom{N}{2}$, $n = 2$, and $w_H = 1$ if H is a connected 2-vertex graph (i.e., an edge) and $w_H = 0$ if H consists of two isolated vertices.

¹⁵A more general result is presented in Theorem 3.11.

Theorem 3.10 (a kind of characterization): *Let $\Pi = \cup_N \Pi_N$ be a graph property. Then, Π has a two-sided error POT if and only if there exists an integer n , a number $b \in [0, 1]$, and a function F such that for every N there exists a sequence $\bar{w} = (w_H)_{H:|V(H)|=n}$ that satisfies the following two conditions:*

1. Π_N equals the set of N -vertex graphs in $\Pi_{\bar{w},b}$.
2. If G is δ -far from Π_N , then $\sum_H w_H \cdot \rho_H(G) \geq b + F(\delta)$.

Indeed, the second condition drastically limits the usefulness of the current characterization; still, Theorem 3.11 (which generalizes Theorem 3.9) presents cases in which this condition holds. Note that while one direction of Theorem 3.10 is quite obvious (i.e., that properties that correspond to such sequences of $\Pi_{\bar{w},b}$'s have a POT), the opposite direction requires a proof (i.e., that having a POT implies a correspondence to such sequences of $\Pi_{\bar{w},b}$'s).

Proof: The proof follows the outline of the proof of [GR09b, Thm. 4.7]. Suppose that Π has a constant-query (two-sided error) POT. Then, by following the proof of [GT03, Thm. 4.5] (see also [GT05]), we can obtain a POT that inspects the subgraph induced by a random set of $n = O(1)$ vertices and accepts with probability α_H if the induced subgraph seen is isomorphic to H . Note that n equals twice the query complexity of the original POT, and that the resulting POT maintains the acceptance probability of the original POT (on any random isomorphic copy of any fixed graph G).¹⁶ Let c be the acceptance threshold of the original POT (i.e., $c = \min_{G \in \Pi} \{\Pr[\text{Test}^G(|V(G)|) = 1]\}$). Then, $\Pi_N = \{G : |V(G)| = N \wedge \sum_H \rho_H(G) \cdot \alpha_H \geq c\}$, which equals the set of N -vertex graphs in $\Pi_{\bar{w},b}$ for $w_H = 1 - \alpha_H$ and $b = 1 - c$. That is, this \bar{w} satisfies the first condition. Furthermore (by the POT guarantee), if the N -vertex graph G' is ϵ' -far from Π , then $\sum_H \rho_H(G') \cdot \alpha_H \leq c - \varrho(\epsilon')$, where ϱ is the guaranteed detection probability function. That is, this \bar{w} satisfies the second condition, with respect to $F \stackrel{\text{def}}{=} \varrho$, since $\sum_H \rho_H(G') \cdot w_H \geq b + F(\epsilon')$. Thus, we obtained n, b and F such that for every N there exists a sequence of w_H 's that satisfies both conditions.

Suppose, on the other hand, that for some n, b and F , it holds that for every N there exists a sequence $\bar{w} = (w_H)_{H:|V(H)|=n}$ that satisfies the two conditions (i.e., (i) Π_N equals the set of N -vertex graphs in $\Pi_{\bar{w},b}$, and (ii) if G is δ -far from Π_N then $\sum_H w_H \cdot \rho_H(G) \geq b + F(\delta)$). Our goal is to derive a constant-query two-sided error POT for Π , which we achieve using the following natural test: The test selects a random set of n vertices, inspects the induced subgraph, and accepts with probability $1 - w_H$ when seeing a subgraph isomorphic to H . Clearly, every graph in Π_N is accepted with probability at least $c \stackrel{\text{def}}{=} 1 - b$, whereas if G is δ -far from Π_N then it is accepted with probability at most $\sum_H (1 - w_H) \cdot \rho_H(G) \leq c - F(\delta)$. Thus, this test is a two-sided error POT with $\varrho = F$. ■

Discussion. As admitted upfront, Theorem 3.10 leaves open the question of which graph properties can be captured by sequences of w_H 's that satisfy the second condition (i.e., that being δ -far from $\Pi_{\bar{w},b}$, means satisfying $\sum_H w_H \cdot \rho_H(\cdot) \geq b + F(\delta)$). Generalizing Theorem 3.9, it is easy to prove the following.

Theorem 3.11 (Theorem 3.9, generalized): *Let $b \in [0, 1]$ and $\bar{w} = (w_H)_{H:|V(H)|=n}$. If the set $\{H : w_H \geq b\}$ contains only graphs with no isolated vertices, then for every graph G that is δ -far from $\Pi_{\bar{w},b}$ it holds that $\sum_H w_H \cdot \rho_H(G) \geq b + (dn/3) \cdot \delta$, where $d \stackrel{\text{def}}{=} b - \max_{H:w_H < b} \{w_H\}$ and*

¹⁶We avoid the final step in [GT03, Sec. 4] (and [GR09b]), where each $\alpha_H > 0$ is replaced by $\alpha_H = 1$, yielding a deterministic decision (which in turn corresponds to \mathcal{F} -freeness).

provided that $\delta > 6/|V(G)|$. The same holds if $\{H : w_H \geq b\}$ contains only graphs in which no vertex neighbors all other vertices.

When applying the argument used in proving Theorem 3.9, each iteration reduces the value of $\sum_H w_H \cdot \rho_H(\cdot)$ by at least $d \cdot \binom{N-1}{n-1}$ units. Thus, we obtain $\sum_H w_H \cdot \rho_H(G) \geq b + (\delta n/3) \cdot d$. In contrast to Theorem 3.11, we observe that not every $\Pi_{\bar{w},b}$ satisfies the second condition of Theorem 3.10. Specifically, we show the following

Proposition 3.12 (violating the second condition of Theorem 3.10): *There exists $b \in (0, 1)$ and $\bar{w} = (w_H)_{H:|V(H)|=O(1)}$ such that for every N there exists an N -vertex graph G that is $\Omega(1)$ -far from $\Pi_{\bar{w},b}$ and yet $\sum_H w_H \cdot \rho_H(G) = b + O(1/N)$.*

Note that this does not say that $\Pi_{\bar{w},b}$ does not have a POT, since such a possible POT may use an alternative characterization of the same property (i.e., $\Pi_{\bar{w},b}$ may equal $\Pi_{\bar{w}',b'}$ such that the former violates the second condition of Theorem 3.10 whereas the latter satisfies this very condition). In Proposition 6.12 (see also Corollary 6.11) we provide an example of a property $\Pi = \cup_{N \in \mathbb{N}} \Pi_N$ that does not have a POT, but for some $b \in (0, 1)$ and every $N \in \mathbb{N}$ there are $\bar{w} = (w_H)_{H:|V(H)|=O(1)}$ such that Π_N is exactly the set of all N -vertex graphs in $\Pi_{\bar{w},b}$. Thus, Proposition 6.12 asserts that a property $\Pi_{\bar{w},b}$ that satisfies the first condition of Theorem 3.10 does not necessarily have a POT.¹⁷

Proof: Consider any of the properties $\mathcal{SC}_\ell \mathcal{C} \cap \mathcal{R}^{(k)}$ asserted in Proposition 3.5, where $k(N) = 2N/\ell t$ for some $t \in \mathbb{N} \setminus T_\ell$. Next, consider the tester for $\mathcal{SC}_\ell \mathcal{C} \cap \mathcal{R}^{(k)}$ described in the proof of Theorem 3.2: This tester selects a random sample of $n = O(1)$ vertices, and inspect the corresponding induced subgraph, denoted H . Specifically, let H' be the subgraph induced by the first three vertices and H'' be the subgraph induced by the other $n-3$ vertices. Then, this tester accepts with probability $\alpha'_{H'}$ if $H'' \in \text{AS}$ and rejects otherwise, where $(\alpha'_{H'})_{H'}$ denotes the sequence of probabilities used by the POT of $\mathcal{R}^{(k)}$ and AS denotes the set of all possible induced subgraphs of graphs in $\mathcal{SC}_\ell \mathcal{C}$. Recall that, for some $c' \in (0, 1)$ and every graph $G \in \mathcal{R}^{(k)}$, it holds that $\sum_{H'} \rho_H(G) \cdot \alpha'_{H'} = c'$ (since all graphs in $\mathcal{R}^{(k)}$ are accepted by the corresponding POT with exactly the same probability). Denoting by $\mathbb{H}(H', H'')$ the set of all n -vertex graphs H that induce the graphs H' and H'' as above, we observe that for every graph G it holds that $\sum_{H \in \mathbb{H}(H', H'')} \rho_H(G) = \rho_{H'}(G) \cdot \rho_{H''}(G) \pm O(n/N)$. Note that for every $H \in \mathbb{H}(H', H'')$, the test accepts a graph when seeing the induced subgraph H with probability $\alpha_H = \alpha'_{H'}$ if $H'' \in \text{AS}$ and $\alpha_H = 0$ otherwise. Then, for the N -vertex graph G , we have

$$\begin{aligned} \sum_{H:|V(H)|=n} \alpha_H \cdot \rho_H(G) &= \sum_{H':|V(H')|=3} \sum_{H'':|V(H'')|=n-3} \sum_{H \in \mathbb{H}(H', H'')} \alpha_H \cdot \rho_H(G) \\ &= \sum_{H':|V(H')|=3} \alpha'_{H'} \cdot \sum_{H'' \in \text{AS}} \sum_{H \in \mathbb{H}(H', H'')} \rho_H(G) \\ &= \sum_{H':|V(H')|=3} \alpha'_{H'} \cdot \sum_{H'' \in \text{AS}} [\rho_{H'}(G) \rho_{H''}(G) \pm O(n/N)] \\ &= \left(\sum_{H'} \alpha'_{H'} \rho_{H'}(G) \right) \cdot \left(\sum_{H'' \in \text{AS}} \rho_{H''}(G) \right) \pm O(2^{n^2}/N) \end{aligned}$$

Recall that $n = O(1)$ and thus $O(2^{n^2}/N) = O(1/N)$. Then, for the N -vertex graph G asserted in Proposition 3.5, we have $\sum_{H:|V(H)|=n} \alpha_H \cdot \rho_H(G) = c' - O(1/N)$, because $G \in \mathcal{R}^{(k)}$ implies

¹⁷We note that Proposition 3.12 is not subsumed by Proposition 6.12, since the parameters \bar{w} used in the former are independent of N whereas in the latter \bar{w} depend on N .

$\sum_{H'} \alpha'_{H'} \rho_{H'}(G) = c'$ whereas the fact that G is $O(1/N)$ -close to $\mathcal{SC}_\ell \mathcal{C}$ implies that the density (in G) of subgraphs not in \mathcal{AS} is at most $O(n^2/N)$. Finally, using the same translation as in the proof of Theorem 3.10 (i.e., $b = 1 - c'$ and $w_H = 1 - \alpha_H$), we conclude that although G is $\Omega(1)$ -far from $\Pi_{\overline{w}, b}$ it holds that $\sum_H w_H \cdot \rho_H(G) = b + O(1/N)$. The claim follows. ■

3.6 Impossibility results

It is easy to derive impossibility results regarding general POTs by considering two distributions on N -vertex such that the following two conditions hold: (1) the two distributions cannot be distinguished by a constant number of queries, and (2) the first distribution is concentrated on graphs that have the property whereas the second distribution is concentrated on graphs that do not have the property.¹⁸

For example, wishing to prove that bipartiteness has no constant-query POTs, we consider for each constant q , the following two distributions that refer to $\ell = 2\lceil q/2 \rceil + 1$: The first distribution consists of random isomorphic copies of an N -vertex graph that is obtained by a balanced blow-up of a single 2ℓ -cycle, and the second distribution is analogously obtained by a balanced blow-up of two ℓ -cycles. Thus, each graph in each of the two distributions consists of 2ℓ clouds such that each cloud consists of an independent set of size $N/2\ell$, and the clouds are arranged either in a single 2ℓ -cycle or in two disjoint ℓ -cycles. Clearly, these distributions cannot be distinguished by an algorithm that makes less than ℓ queries, but graphs in the first distribution are bipartite whereas graphs in the second distribution are far from being bipartite. Thus, we get:

Theorem 3.13 *Bipartiteness has no two-sided error POT.*

4 In the Bounded-Degree Graph Model

The bounded-degree graph model refers to a fixed degree bound, denoted $d \geq 2$. An N -vertex graph $G = ([N], E)$ (of maximum degree d) is represented in this model by a function $g : [N] \times [d] \rightarrow \{0, 1, \dots, N\}$ such that $g(v, i) = u \in [N]$ if u is the i^{th} neighbor of v and $g(v, i) = 0$ if v has less than i neighbors. Distance between graphs is measured in terms of their aforementioned representation (i.e., as the fraction of (the number of) different array entries (over dN)).

The straightforward method for showing impossibility results (outlined in Section 3.6), is applicable also in the current (bounded-degree) model. To demonstrate this, we show that (for any constant q) the connectivity property has no q -query (two-sided error) POT in this model. The two distributions that we consider are: (1) a random isomorphic copy of the graph consisting of a single N -vertex Hamiltonian cycle, and (2) a random isomorphic copy of the graph consisting of $N/(q+1)$ isolated $(q+1)$ -vertex cycles. Thus, we get:

Theorem 4.1 *Connectivity has no two-sided error POT (in the bounded-degree graph model, for any $d \geq 2$).*

Turning to positive results, we note that the properties of distributions studied in Section 2 give rise to graph properties that have POTs in the bounded-degree model. The first type of such graph properties refer to the edge densities of graphs, where in the current section densities are measured

¹⁸The foregoing method directly establishes the non-existence of a two-sided error POT. Alternatively, one may use this method to show that the first condition in Theorem 3.10 is not satisfied. Indeed, using Theorem 3.10 merely allows to replace (1) by (1') the two distributions have the same densities of various induced subgraphs of constant size.

as a fraction of $dN/2$. (Note that a Boolean function $f : [N] \times [d] \rightarrow \{0, 1\}$ can be defined such that $f(v, i) = 1$ if and only if $g(v, i) \in [N]$.)¹⁹ As in Section 3, we are more interested in “genuine” graph properties, and the first type of properties that we consider refer to the density of isolated vertices in the graph.

Recall that for any sequence of t density thresholds, denoted $\bar{\tau} = (\tau_1, \dots, \tau_t) \in [0, 1]^t$, such that $\tau_1 \leq \tau_2 < \tau_3 \leq \tau_4 < \dots \leq \tau_t$, we considered (in Section 2) the class of distributions, denoted $\mathcal{D}_{\bar{\tau}}$, consists of all 0-1 random variables X such that for some $i \leq \lceil t/2 \rceil$ it holds that $\tau_{2i-1} \leq \Pr[X = 1] \leq \tau_{2i}$. The corresponding class of bounded-degree graphs will consist of graphs that contain a fraction of isolated vertices that corresponds to a distribution in $\mathcal{D}_{\bar{\tau}}$. That is, $\mathcal{G}_{\bar{\tau}}$ contains the N -vertex graph G if and only if G contains M isolated vertices such that the fraction M/N (viewed as a probability) is in $\mathcal{D}_{\bar{\tau}}$.

Theorem 4.2 (POT for $\mathcal{G}_{\bar{\tau}}$): *For every $\bar{\tau} = (\tau_1, \dots, \tau_t)$, the property $\mathcal{G}_{\bar{\tau}}$ has a two-sided error POT.*

Proof: On input N and oracle access to an N -vertex graph $G = ([N], E)$, of degree bound d , the tester proceeds as follows.

1. Selects uniformly and independently t vertices, denoted u_1, \dots, u_t , and explore their immediate neighborhood. That is, for each i determine whether or not u_i is isolated in G .
2. Let $j \in \{0, 1, \dots, t\}$ denote the number of isolated vertices seen in Step 1. Then, the tester accepts with probability α_j , where $(\alpha_0, \alpha_1, \dots, \alpha_t)$ is the sequence of probabilities used by the POT that is guaranteed by Theorem 2.3 (i.e., the tester for $\mathcal{D}_{\bar{\tau}}$).

Let c be the threshold probability associated with the tester of Theorem 2.3. Then, each graph $G \in \mathcal{G}_{\bar{\tau}}$ is accepted with probability at least c . On the other hand, we shall show that if G is ϵ -far from being in $\mathcal{G}_{\bar{\tau}}$, then the fraction of isolated vertices in G is $(\epsilon/4)$ -far from $\mathcal{D}_{\bar{\tau}}$, and the theorem follows. Actually, the validity of this claim presupposes that all the thresholds in $\bar{\tau}$ are multiples of $1/N$, and we shall defer this issue to the end of the proof.

We shall prove the counterpositive (i.e., if the fraction of isolated vertices in G is ϵ -close to $\mathcal{D}_{\bar{\tau}}$, then G is 4ϵ -close to $\mathcal{G}_{\bar{\tau}}$). Suppose that G is an N -vertex graph with M isolated vertices such that there exists $p \in \mathcal{D}_{\bar{\tau}}$ that satisfies $|p - (M/N)| \leq \epsilon$. If $M > pN$ (and $M < N$)²⁰, then we may decrement the number of isolated vertices by connecting any isolated vertex v to some non-isolated vertex. (If some non-isolated vertex u has degree smaller than d , then we connect v to u , else we connect v to an arbitrary vertex of degree d while omitting one of its current edges.) The case $M < pN$ is slightly more complex, since we wish to turn some non-isolated vertex v into an isolated vertex. If each of the neighbors of v has degree at least two, then these is no problem. Otherwise, we may need to connect these neighbors among themselves so to prevent them from becoming isolated. The details are omitted.

Note that the foregoing argument presupposes that pN is an integer, which is indeed the case when all the thresholds in $\bar{\tau}$ are multiples of $1/N$. Thus, our argument needs to be augmented to deal with the general case, in which the latter presumption does not hold. We distinguish between dealing threshold pairs of the form $\tau_{2i-1} < \tau_{2i}$ and pairs of the form $\tau_{2i-1} = \tau_{2i}$. In the first case, ignoring finitely many N 's, we may replace $p \in [\tau_{2i-1}, \tau_{2i}]$ by $p' \in [\tau_{2i-1}, \tau_{2i}] \cap \{j/N : j = 0, 1, \dots, N\}$ (while increasing ϵ by at most $1/N$, which is fine since it suffices to establish an upper bound of $4(\epsilon + (1/N))$). In the second case, we should actually modify the algorithm and omit the pair

¹⁹Thus, the fraction of 1-values in f equals the fraction of edges in the graph represented by g .

²⁰If $M = N$, then it must be that $pN \leq N - 2$, and thus connecting a pair of isolated vertices is fine.

(τ_{2i-1}, τ_{2i}) from $\bar{\tau}$. That is, the algorithm will refer to a modified $\bar{\tau}$ that contains a pair of the form $\tau_{2i-1} = \tau_{2i}$ if and only if such a pair is a multiple of $1/N$ (for the current N).²¹ ■

Generalization. The foregoing treatment can be extended to properties that refer to the density of certain isolated patterns in the graph. Specifically, for any fixed family of graphs \mathcal{H} , we denote by $\#\mathcal{H}(G)$ the number of connected components in G that are isomorphic to some graph in \mathcal{H} . Next, for any $\bar{\tau}$ as above, we may consider the property $\mathcal{G}_{\mathcal{H}, \bar{\tau}}$ that consist of N -vertex graphs G such that the fraction $\#\mathcal{H}(G)/N$ is in $\mathcal{D}_{\bar{\tau}}$. (Indeed, $\mathcal{G}_{\bar{\tau}}$ is a special case obtained when letting \mathcal{H} be a singleton consisting of the 1-vertex graph.) The integrality issue (i.e., the τ_i 's not necessarily being multiples of $1/N$) dealt with at the end of the proof of Theorem 4.2 takes a more acute form in the current setting, since if \mathcal{H} consists only of n -vertex graphs then $\#\mathcal{H}(G)/N$ resides in the interval $[0, 1/n]$ (rather than in $[0, 1]$). Therefore, letting $s(\mathcal{H})$ denote the (number of vertices in the) smallest graph in \mathcal{H} , we may restrict our attention to the interval $[0, 1/s(\mathcal{H})]$.

Theorem 4.3 (POT for $\mathcal{G}_{\bar{\tau}}$): *For every \mathcal{H} and every $\bar{\tau} = (\tau_1, \dots, \tau_t)$ such that $\tau_t \leq 1/s(\mathcal{H})$ the property $\mathcal{G}_{\mathcal{H}, \bar{\tau}}$ has a two-sided error POT.*

Proof: We build on the proof of Theorem 4.2, while somewhat adapting both the tester and its analysis. For starters, the tester should look for isolated copies of graphs in \mathcal{H} (rather than isolated vertices), and count them in proportion to their size (which reflects the probability that a uniformly selected vertex hits such a copy). Let $n = n(\mathcal{H})$ denote the (number of vertices in the) largest graph in \mathcal{H} . Then, on input N and oracle access to an N -vertex graph $G = ([N], E)$, of degree bound d , the modified tester proceeds as follows.

1. Selects uniformly and independently t vertices, denoted u_1, \dots, u_t , and explore the neighborhood of each vertex u_i till discovering at most $n+1$ vertices. For each $i \in [t]$, let $p_i = 1/|V(H)|$ if u_i resides in a connected component of G that is isomorphic to $H \in \mathcal{H}$, and $p_i = 0$ otherwise.
2. For each $i \in [t]$, let $c_i = 1$ with probability p_i and $c_i = 0$ otherwise, and let $j = \sum_{i=1}^t c_i$. Then, the tester accepts with probability α_j , where $(\alpha_0, \alpha_1, \dots, \alpha_t)$ is the sequence of probabilities used by the POT that is guaranteed by Theorem 2.3 (i.e., the tester for $\mathcal{D}_{\bar{\tau}}$).

Let c be the threshold probability associated with the tester of Theorem 2.3. Then, each graph $G \in \mathcal{G}_{\mathcal{H}, \bar{\tau}}$ is accepted with probability at least c , since for each i it holds that $\Pr[c_i = 1] = \#\mathcal{H}(G)/N$. On the other hand, we shall show that if G is ϵ -far from being in $\mathcal{G}_{\mathcal{H}, \bar{\tau}}$, then $\#\mathcal{H}(G)/N$ is $\Omega(\epsilon)$ -far from $\mathcal{D}_{\bar{\tau}}$, and the theorem follows.

Following the proof of Theorem 4.2, we show how to decrement and increment the number of good connected components in a graph, where a component is called **good** if it is isomorphic to some $H \in \mathcal{H}$ (and is **bad** otherwise). We consider two cases that refer to whether or not the single vertex is in \mathcal{H} (i.e., whether or not $s(\mathcal{H}) = 1$).

We start with the case that $s(\mathcal{H}) > 1$ (i.e., an isolated vertex is a bad component). In this case, we can decrement the number of good components by omitting all edges that appear in an arbitrary good component, turning this component to a collection of isolated vertices (which are bad components in this case). To increment the number of good components, we may combine s vertices that are taken from bad components, while keeping each of these components bad by either maintaining its connectivity (by adding edges, if it contains more than n vertices) or replacing

²¹Indeed, this means that the algorithm may use up to $2^{t/2}$ different sequences $\bar{\tau}$, each having its own corresponding POT. This requires scaling the threshold probabilities of all these POTs so that they are all equal, and it is indeed crucial that we are dealing with a finite number of algorithms (or threshold probabilities).

it by isolated vertices (if this component contains at most n vertices). Thus, each decrement or increment operation is charged with $O(n^2)$ edge modifications. This completes the treatment of the case $s(\mathcal{H}) > 1$.

We now turn to the case that $s(\mathcal{H}) = 1$ (i.e., an isolated vertex is a good component). If we wish to decrement number of good components, then we pick a (largest) good component, and connect it to any bad connected component (or to another good component if all components are good). (This connection is made via a pair of vertices of degree less than d , and if no such vertex exists in the relevant component then we create it by omitting an arbitrary edge.) This operation either decreases the number of good components or increases the size of the largest good component, and so we can decrease the number of good components by $O(n)$ edge modifications. (Note that in case we connect two good components, the number of good component may decrease by two units.)

If we wish to increment the number of good components, then we select a vertex that belong to any bad connected component (or from a non-singleton good component if all components are good), and disconnect it from its current neighborhood, thus creating a new isolated vertex (which is a good component). When disconnecting this vertex from its neighbors, we may add edges so to maintain the connectivity of this component. Note that when modifying the said component, we may turn a bad component to a good one (or turn a good one to a bad one). Thus, either the number of good components increases (by either one or two units) or a bad component is created and can be used in our next attempt.

The forgoing description suffices for getting the number of good components to either equal the desired number or be one unit below the desired number. To close this final gap, we make two observations.

1. Suppose that the graph contains at least $n + 2$ vertices in bad components. Then, by picking at most $(n + 2)/2$ bad components that contain $m \geq n + 2$ vertices, we can form a new collection of connected components with exactly one good component (by creating a single isolated vertex and a single bad component that contains all the other $m - 1 \geq n + 1$ vertices).
2. Suppose that the graph contains at least $n + 2$ good non-singleton components. Then, by picking $n + 2$ such components, we can form a new collection of connected components with exactly $n + 3$ good component (by creating $n + 3$ isolated vertices and a single bad component that contains all the other vertices, the number of which is at least $2(n + 2) - (n + 3) > n$).

In both cases, $O(n^2)$ edge modifications are used. The only case where we cannot apply either of these observations is when the number of isolated vertices is $N - O(n^2)$. Fortunately, we can ignore this case, because it may occur only if $1 \in \mathcal{D}_{\neq}$ and in such a case we may just increase the number of isolated vertices to N in the trivial manner. This completes the treatment of the case $s(\mathcal{H}) = 1$.

■

Part II

Additional Results

Theorems 3.9, 3.10 and 4.3 refer to the distribution of certain (fixed size) subgraphs inside a big graph. This motivates the study of POTs for properties of distributions over an arbitrary fixed-size domain. This study is initiated in Section 5 and its results are applied to graph properties in Section 6.

It turns out that POTs for properties of multi-valued distributions are more exceptional than their binary-valued analogues. Recall that in Section 2 we showed that properties (of binary distributions) that correspond to intervals (bounding the probability that the outcome is 1) have POTs. It is tempting to hope that properties of ternary distributions that correspond to rectangles (bounding the probabilities of the outcomes 1 and 2 respectively) also have POTs; however, as shown in Section 5, this is *typically* not the case! In contrast, properties of multi-valued distributions that corresponds to regions that are ellipsoids do have POTs.

Given these limitations of POTs for properties of multi-valued distributions, we focus in Section 6 on very simple properties of graphs (in the adjacency matrix representation). Specifically, for every sequence of intervals $(I_N)_{N \in \mathbb{N}}$ such that $I_N \subseteq [0, 0.5]$, we consider the set of graphs consisting of two isolated cliques such that the density of the smaller clique resides in I_N , where N denotes the number of vertices in the graph. We show that if for all $N \in \mathbb{N}$ the interval I_N is some non-trivial interval I of $[0, 0.5]$ (i.e., I is neither a singleton nor contain all of $[0, 0.5]$), then the corresponding property has no POT. In contrast, if the length of I_N is either smaller than $N^{-1/2}$ or larger than $0.5 - N^{-1/2}$, then the corresponding property has a POT.

5 Classes of Distributions

In this section we generalize the results from Section 2 to distributions over larger (finite) domains. We give a characterization for the classes of distributions that have a two-sided error POT. For $r \in \mathbb{N}$ we shall identify a distribution $\bar{q} = (q_1, \dots, q_r)$ on $[r]$ with a point in $\Delta^{(r)}$, where

$$\Delta^{(r)} = \{(q_1, \dots, q_r) \in [0, 1]^r : \sum_{i \in [r]} q_i = 1\}. \quad (8)$$

Similarly, a class of distributions with domain $[r]$ will be identified with a subset of $\Delta^{(r)}$ in a natural way. The special case of boolean distributions discussed in Section 2 corresponds to $r = 2$, for which $\Delta^{(2)} = \{(p, 1 - p) : p \in [0, 1]\}$.

5.1 Characterizing the classes of distributions that have a POT

The following result asserts that a class of distributions has a POT if and only if there exists a polynomial that is non-negative exactly on the points that correspond to distributions in that class. Thus, the question of whether or not there exists a POT for $\Pi \subseteq \Delta^{(r)}$ reduces to whether or not some polynomial can be non-negative on Π and negative on $\Delta^{(r)} \setminus \Pi$.

Theorem 5.1 (POT and polynomials in the context of distribution testing): *Let Π be an arbitrary class of distributions $\bar{q} = (q_1, \dots, q_r)$ with domain $[r]$; that is, $\Pi \subseteq \Delta^{(r)}$. Then, Π has a two-sided error POT if and only if there is a polynomial $P : \Delta^{(r)} \rightarrow \mathbb{R}$ such that for every distribution $\bar{q} = (q_1, \dots, q_r) \in \Delta^{(r)}$ it holds*

$$P(q_1, \dots, q_r) \geq 0 \iff \bar{q} \in \Pi. \quad (9)$$

If the total degree of P is t , then Π has a two-sided error POT \mathcal{T}_Π that makes t queries and has polynomial detection probability $\varrho(\epsilon) = \Omega(\epsilon^C)$, where $C < t^{O(r)}$.²² Moreover, the acceptance probability of \mathcal{T}_Π when testing $\bar{q} \in \Delta^{(r)}$ can be written as

$$\Pr[\mathcal{T}_\Pi \text{ accepts } \bar{q}] = \frac{1}{2} + \delta \cdot P(q_1, \dots, q_r) \quad (10)$$

²²The constant in the $\Omega()$ notation depends on P , while the $O()$ notation hides some absolute constant.

for some constant $\delta > 0$ that depends only on the degree of \mathbf{P} and on an upper bound of the absolute value of all coefficients of \mathbf{P} .

Proof: The “only if” direction is proved by using the independence of samples of the given distribution. Consider a POT \mathcal{T}_Π for Π , that makes t sampling queries and accepts each distribution in Π with probability at least c .

When testing $\bar{q} = (q_1, \dots, q_r)$, for every view $\bar{v} = (v_1, \dots, v_t) \in [r]^t$, the probability of seeing this view is $\prod_{i=1}^t q_{v_i}$. Denoting by $\alpha_{\bar{v}}$ the probability that the tester accepts the view $\bar{v} = (v_1, \dots, v_t)$, we have

$$\Pr[\mathcal{T}_\Pi \text{ accepts } \bar{q}] = \sum_{\bar{v}=(v_1, \dots, v_t) \in [r]^t} \left(\prod_{i=1}^t q_{v_i} \right) \cdot \alpha_{\bar{v}}.$$

Define a polynomial \mathbf{P} to be

$$\mathbf{P}(q_1, \dots, q_r) = \left(\sum_{\bar{v} \in [r]^t} \alpha_{\bar{v}} \prod_{i=1}^t q_{v_i} \right) - c.$$

Then, by definition of the tester, \mathbf{P} satisfies Eq. (9).

For the other direction, let $\mathbf{P} : \Delta^{(r)} \rightarrow \mathbb{R}$ be a polynomial of degree t . We show that the class

$$\Pi = \{(q_1, \dots, q_r) \in \Delta^{(r)} : \mathbf{P}(q_1, \dots, q_r) \geq 0\} \quad (11)$$

has a POT, that makes t queries, and has polynomial detection probability.

In order to simplify the proof, we shall slightly modify \mathbf{P} , while making sure that the modifications of \mathbf{P} does not affect Π in Eq. (11). Specifically, we multiply each monomial of degree $d < t$ (of \mathbf{P}) by $(\sum_{i \in [r]} q_i)^{t-d}$. This does not change the value of \mathbf{P} in $\Delta^{(r)}$, and hence does not affect Π .²³ Henceforth we shall assume that \mathbf{P} is a homogeneous polynomial of degree t , and therefore can be written as

$$\mathbf{P}(q_1, \dots, q_r) = \sum_{\bar{v} \in [r]^t} \alpha_{\bar{v}} \prod_{i=1}^t q_{v_i} \quad (12)$$

for some coefficients $\alpha_{\bar{v}} \in \mathbb{R}$.

Assume that Π is non trivial. This implies that not all coefficients $\alpha_{\bar{v}}$ are zeros. Given Eq. (12), we define a POT \mathcal{T}_Π for Π as follows. The tester makes t queries to a given distribution, gets t samples, denoted by $\bar{v} = (v_1, \dots, v_t)$, and accepts with probability

$$\beta_{\bar{v}} = \frac{1}{2} + \delta \cdot \alpha_{\bar{v}},$$

where we choose $\delta = \frac{1}{2 \cdot \max\{\alpha_{\bar{v}} : \bar{v} \in [r]^t\}} > 0$, in order to assure that $\beta_{\bar{v}} \in [0, 1]$ for all \bar{v} . Therefore, when testing $\bar{q} = (q_1, \dots, q_r)$ the acceptance probability of the test is

$$\Pr[\mathcal{T}_\Pi \text{ accepts } \bar{q}] = \sum_{\bar{v} \in [r]^t} \beta_{\bar{v}} \prod_{i=1}^t q_{v_i} = \frac{1}{2} + \delta \cdot \left(\sum_{\bar{v} \in [r]^t} \alpha_{\bar{v}} \prod_{i=1}^t q_{v_i} \right),$$

²³This grouping of monomials to homogeneous monomials maps at most 2^t monomials to a single homogeneous monomials, and thus the coefficients in the \mathbf{P} may grow by a factor of at most 2^t .

and hence, by Eq. (12), the equality above becomes

$$\Pr[\mathcal{T}_\Pi \text{ accepts } \bar{q}] = \frac{1}{2} + \delta \cdot \mathbb{P}(q_1, \dots, q_r). \quad (13)$$

Next, we analyze the acceptance probability in Eq. (13). If $\bar{q} \in \Pi$, then, by Eq. (11), we have $\mathbb{P}(q_1, \dots, q_r) \geq 0$, and therefore

$$\Pr[\mathcal{T}_\Pi \text{ accepts } \bar{q}] \geq \frac{1}{2}.$$

Assume \bar{q} is ϵ -far from Π . Then, in particular $\bar{q} \notin \Pi$, and hence $\mathbb{P}(q_1, \dots, q_r) < 0$. Thus, using Eq. (13), we have $\Pr[\mathcal{T}_\Pi \text{ accepts } \bar{q}] < \frac{1}{2}$. In order to prove that \mathcal{T}_Π is a POT, we need to show that $\Pr[\mathcal{T}_\Pi \text{ accepts } \bar{q}]$ is bounded below $\frac{1}{2}$ by some function that depends on ϵ . This type of result is known in real algebraic geometry as the Łojasiewicz inequality (see [BCR, Chapter 2.6]). Specifically, we use the following theorem of Solernó [Sol].

Theorem 5.2 (Effective Łojasiewicz inequality): *Let $\mathbb{P} : \Delta^{(r)} \rightarrow \mathbb{R}$ be a polynomial, and let*

$$\Pi = \{(p_1, \dots, p_r) \in \Delta^{(r)} : \mathbb{P}(p_1, \dots, p_r) \geq 0\}.$$

Assume that for $\bar{q} = (q_1, \dots, q_r) \in \Delta^{(r)}$ it holds

$$\text{dist}(\bar{q}, \Pi) = \inf\left\{\frac{1}{2} \sum_{i \in [r]} |q_i - p_i| : (p_1, \dots, p_r) \in \Pi\right\} > \epsilon.$$

Then, $\mathbb{P}(q_1, \dots, q_r) < -\Omega(\epsilon^C)$ for some constant $C < \deg(\mathbb{P})^{O(r)}$, where the constant in the $\Omega()$ notation depends on \mathbb{P} , and the $O()$ notation hides some absolute constant.

By applying Theorem 5.2 on Eq. (13), we conclude that if $\bar{q} \in \Delta^{(r)}$ is ϵ -far from Π , then $\Pr[\mathcal{T}_\Pi \text{ accepts } \bar{q}] < \frac{1}{2} - \Omega(\epsilon^C)$, where $C < \deg(\mathbb{P})^{O(r)}$. This completes the proof of Theorem 5.1. ■

Corollaries: As hinted upfront, Theorem 5.1 provides a tool towards proving both positive and negative results regarding the existence of POTs for various properties. First, we use this tool to show that classes of distributions that have a POT are closed under taking *disjoint* unions. Next, we use it to present POTs for some concrete properties of interest, and lastly we use it to derive impossibility results about other concrete properties of interest.

5.2 Closure under disjoint union

Recall that in the standard property testing model, as well as in one-sided error POT model, testable properties are closed under union. However, for properties of distributions with two-sided error POT, the closure under union does not hold: Indeed, in Proposition 5.13 (see Section 5.4), we show two properties that have two-sided error POTs, but their union does not have a POT. Nevertheless, we prove next that if two *disjoint* classes of distributions have two-sided error POTs, then so does their union.

Corollary 5.3 (closure under disjoint union): *Let Π and Π' be two disjoint classes of distributions with domain $[r]$, and suppose that both Π and Π' have a two-sided error POT. Then, their union $\Pi \cup \Pi'$ also has a two-sided error POT.*

Proof: By Theorem 5.1 if Π has a POT, then there is a polynomial $P : \Delta^{(r)} \rightarrow \mathbb{R}$, such that $\Pi = \{\bar{q} \in \Delta^{(r)} : P(\bar{q}) \geq 0\}$. Similarly, there is a polynomial $P' : \Delta^{(r)} \rightarrow \mathbb{R}$, such that $\Pi' = \{\bar{q} \in \Delta^{(r)} : P'(\bar{q}) \geq 0\}$. Define a polynomial $P_{\text{union}} : \Delta^{(r)} \rightarrow \mathbb{R}$ to be

$$P_{\text{union}}(\bar{q}) = -P(\bar{q}) \cdot P'(\bar{q}).$$

Since Π and Π' are disjoint subsets of $\Delta^{(r)}$, it holds that $P_{\text{union}}(\bar{q}) \geq 0$ if and only if $\bar{q} \in \Pi \cup \Pi'$: Indeed, if $\bar{q} \in \Pi$, then $\bar{q} \notin \Pi'$ (since the classes are disjoint), and therefore $P_{\text{union}}(\bar{q}) = -P(\bar{q}) \cdot P'(\bar{q}) \geq 0$. Similarly $P_{\text{union}}(\bar{q}) \geq 0$ for $\bar{q} \in \Pi'$. On the other hand, if $\bar{q} \notin \Pi \cup \Pi'$, then $P(\bar{q}) < 0$ and $P'(\bar{q}) < 0$, and hence $P_{\text{union}}(\bar{q}) = -P(\bar{q}) \cdot P'(\bar{q}) < 0$. By Theorem 5.1 the class $\Pi \cup \Pi'$ has a two-sided error POT. ■

By applying Corollary 5.3 repeatedly, we get

Corollary 5.4 *Let Π_1, \dots, Π_k be disjoint classes of distributions with domain $[r]$, and suppose that each of the classes Π_i has a two-sided error POT. Then, their union $\Pi = \cup_{i=1}^k \Pi_i$ also has a two-sided error POT.*

In Proposition 5.5 we strengthen the foregoing corollary. Specifically, we prove a result specifying (and improving) the query complexity and the detection probability of a POT for disjoint union of properties. We defer the proof of the proposition to Appendix A.3.

Proposition 5.5 (Corollary 5.4, revisited): *Let Π_1, \dots, Π_k be disjoint classes of distributions with domain $[r]$. Suppose that for each $i \in [k]$ the class Π_i has a two-sided error POT that makes t_i queries and has detection probability ρ_i . Then, their union $\Pi = \cup_{i \in [k]} \Pi_i$ has a two-sided error POT that makes $\sum_{i \in [k]} t_i$ queries and has detection probability $\Omega(\min\{\rho_i : i \in [k]\})$.*

Closure to complement: It is natural to ask whether properties having POTs are closed under taking the complement. Note, however, that if Π has a POT, then $\Pi = \{\bar{q} \in \Delta^{(r)} : P(\bar{q}) \geq 0\}$ for some polynomial $P : \Delta^{(r)} \rightarrow \mathbb{R}$, and thus is a closed²⁴ subset of $\Delta^{(r)}$. Hence, its complement is an open set, and cannot have a POT. Still, it could be natural to conjecture that the closure²⁵ of the complement, denoted by $\text{cl}(\Delta^{(r)} \setminus \Pi)$, has a POT. In Appendix A.5 we show that this is not true in general, by presenting a class of distributions $\Pi \subseteq \Delta^{(3)}$, that has a POT, such that $\text{cl}(\Delta^{(r)} \setminus \Pi)$ does not have one.

5.3 Positive corollaries

In this section we give several concrete examples of properties that have two-sided error POTs. We first note that (a weaker quantitative version of) Theorem 2.3 can be derived by observing that for every segment there exists a quadratic polynomial that is non-negative on this segment and negative outside it. Indeed, this claim appears explicitly in Section 2 and underlies all results presented there. Still, the quadratic detection probability in Theorem 2.3 doesn't follow from Theorem 5.1, and requires explicit calculations as done in Section 2. For similar reasons, the proof of the following corollary requires an explicit calculation beyond the application Theorem 5.1.

²⁴A set $A \subseteq \Delta^{(r)}$ is a closed subset of $\Delta^{(r)}$ if the complement set $\Delta^{(r)} \setminus A$ is open in $\Delta^{(r)}$, where a B is open in $\Delta^{(r)}$ if each point in B has a small neighborhood that is contained in B ; that is, for every $\bar{q} \in B$ there exists an $\epsilon > 0$ such that every $\bar{q}' \in \Delta^{(r)}$ that is at distance at most ϵ from \bar{q} is actually in B .

²⁵For a set $A \subseteq \Delta^{(r)}$, the closure of A , is the set of all $\bar{q} \in \Delta^{(r)}$ that are arbitrarily close to A ; that is, \bar{q} is in the closure of A if for every $\epsilon > 0$ there is $\bar{q}' \in A$ such that $\text{dist}(\bar{q}, \bar{q}') < \epsilon$.

Corollary 5.6 (classes containing a single distribution have POTs): *For a fixed $r \geq 2$, and a distribution $\bar{p} = (p_1, \dots, p_r)$ with domain $[r]$, let $\Pi(\bar{p})$ be the class consisting of the single distribution \bar{p} . Then, the property $\Pi(\bar{p})$ has a two-sided error POT that makes two queries, and has quadratic detection probability.*

Proof: Define a quadratic polynomial P in r variables that is negative for all $(q_1, \dots, q_r) \in \Delta^{(r)} \setminus \{(p_1, \dots, p_r)\}$ and equals zero in (p_1, \dots, p_r) . Specifically, let

$$P(q_1, \dots, q_r) = - \sum_{i \in [r]} (q_i - p_i)^2.$$

It is immediate that P satisfies the required conditions. Therefore, by Theorem 5.1, we get a POT for $\Pi(\bar{p})$, which makes 2 queries.

In order to show that the POT has quadratic detection probability, by Theorem 5.1, it is enough to show that for any $\bar{q} \in \Delta^{(r)}$ that satisfies $\text{dist}(\bar{q}, \bar{p}) > \epsilon$ it holds $P(q_1, \dots, q_r) < -\Omega(\epsilon^2)$. Indeed, by Cauchy-Schwarz inequality we have

$$P(q_1, \dots, q_r) = - \sum_{i \in [r]} (q_i - p_i)^2 \leq -\frac{1}{r} \cdot \left(\sum_{i \in [r]} |q_i - p_i| \right)^2 < -\frac{1}{r} \cdot \epsilon^2,$$

which completes the proof. \blacksquare

By combining Corollary 5.6 with Proposition 5.5 we can test any property that consists of finitely many distributions.

Corollary 5.7 (finite classes of distributions have POTs): *Fix $r \geq 2$ and $k \geq 2$, and let Π be a property that contains exactly k distributions with domain $[r]$. Then, Π has a POT that makes $2k$ queries and has quadratic detection probability.*

POTs for infinite classes of distributions. By slightly generalizing the proof of Corollary 5.6, we show POTs for classes that contain infinitely many distributions. Specifically, we first shall show that any class of distributions that correspond to an ellipsoid has a POT. Let $\bar{p} = (p_1, \dots, p_r)$ be a distribution, and let $\bar{B} = (B_0; B_1, \dots, B_r) \in \mathbb{R}^{r+1}$ such that $B_0 \geq 0$ and $B_i > 0$ for all $i \in [r]$. Let $\Pi(\bar{p}, \bar{B})$ be a class of distributions that lie within an ellipsoid centered at $\bar{p} = (p_1, \dots, p_r)$ with radii $(\sqrt{\frac{B_0}{B_1}}, \dots, \sqrt{\frac{B_0}{B_r}})$. That is

$$\Pi(\bar{p}, \bar{B}) = \left\{ \bar{q} = (q_1, \dots, q_r) \in \Delta^{(r)} : \sum_{i \in [r]} B_i (q_i - p_i)^2 \leq B_0 \right\}. \quad (14)$$

In the special case of $B_0 = 0$, the property $\Pi(\bar{p}, \bar{B})$ contains exactly one distributions; that is, it corresponds to the properties discussed in Corollary 5.6.

Corollary 5.8 (classes that correspond to ellipsoids have POTs): *Fix $r \geq 2$, and let $\bar{p} = (p_1, \dots, p_r)$ and $\bar{B} = (B_0; B_1 \dots B_r)$ as above. Then, the property $\Pi(\bar{p}, \bar{B})$ has a two-sided error POT that makes two queries and has polynomial detection probability.*

Proof: As in the proof of Corollary 5.6, define a polynomial P in r variables, that is non-negative for all points (q_1, \dots, q_r) in the ellipsoid, and negative outside the ellipsoid. Specifically, let

$$P(q_1, \dots, q_r) = B_0 - \sum_{i \in [r]} B_i (q_i - p_i)^2.$$

By Eq. (14), it is immediate that P satisfies the required conditions. The corollary follows by applying Theorem 5.1. ■

In Appendix A.4 we strengthen Corollary 5.8; specifically, we show that for $B_0 > 0$ the POT given in the foregoing proof has linear detection probability (rather than some polynomial probability, as guaranteed by Theorem 5.1).

Corollary 5.9 (POT for classes of distributions that satisfy polynomial equalities): *Let $P_1, \dots, P_t : \mathbb{R}^r \rightarrow \mathbb{R}$ be polynomials of total degree at most d . Then, the class of distributions*

$$\left\{ \bar{q} = (q_1, \dots, q_r) \in \Delta^{(r)} : (\forall i \in [t]) P_i(q_1, \dots, q_r) = 0 \right\} \quad (15)$$

has a two-sided error POT. Furthermore, this POT makes $2d$ queries and has polynomial detection probability.

An appealing special case of Corollary 5.9 refers to the case that each P_i asserts the equality of two arguments (i.e., $P_i(q_1, \dots, q_r) = q_{i_1} - q_{i_2}$). This case corresponds to the class of distributions in which some outcomes occur with the same probability (i.e., the probability that the outcome is i_1 equals the probability that the outcome is i_2).

Proof: Consider the polynomial $P(q_1, \dots, q_r) = -\sum_{i \in [t]} P_i(q_1, \dots, q_r)^2$. Clearly, $P(q_1, \dots, q_r) \geq 0$ if and only if for every $i \in [t]$ it holds that $P_i(q_1, \dots, q_r) = 0$. The corollary follows by applying Theorem 5.1. ■

5.4 Negative corollaries

In this section we give examples of classes of distributions that do not have a POT. We start with a simple claim.

Claim 5.10 (a polynomial cannot be non-negative only in one quarter): *Let $B \subseteq \mathbb{R}^2$ be a neighborhood²⁶ of $(0, 0) \in \mathbb{R}^2$. Then, there is no real polynomial $P : \mathbb{R}^2 \rightarrow \mathbb{R}$ such that for all $(x, y) \in B$ it holds*

1. $P(x, y) \geq 0$, if $x < 0$ and $y < 0$
2. $P(x, y) \leq 0$, if either $x > 0$ or $y > 0$

Proof: Assume towards contradiction that there exists a polynomial satisfying the conditions of the claim, and let P be such polynomial of minimal degree. The conditions on P , together with the continuity of P , implies that $P(x, 0) = 0$ for all sufficiently large $x < 0$ (since for all sufficiently small $\delta > 0$ it holds that $(x, \delta), (x, -\delta) \in B$, which implies $P(x, \delta) \leq 0$ and $P(x, -\delta) \geq 0$). By considering $P(x, 0)$ as a univariate polynomial in x , we conclude that for some small enough $\delta > 0$, it holds $P(x, 0) = 0$ for all $x \in [-\delta, 0]$. Thus, this univariate polynomial must be identically zero (i.e., $P(x, 0) = 0$ for all $x \in \mathbb{R}$), which means that $P(x, y)$ can be written as $P(x, y) = y \cdot P'(x, y)$ for some polynomial $P' : \mathbb{R} \rightarrow \mathbb{R}$ of degree smaller than $\deg(P)$. Then, for all $(x, y) \in B$, we have

²⁶A set $B \subseteq \mathbb{R}^n$ is a neighborhood of $\bar{x} \in \mathbb{R}^n$ if for some $\delta > 0$ it holds that $\{\bar{y} \in \mathbb{R}^n : \text{dist}(\bar{x}, \bar{y}) < \delta\} \subseteq B$, where $\text{dist}(\bar{x}, \bar{y}) = \frac{1}{2} \sum_{i \in [n]} |x_i - y_i|$

1. $P'(x, y) \leq 0$, if $x < 0$ and $y < 0$
2. $P'(x, y) \leq 0$, if $x < 0$ and $y > 0$
3. $P'(x, y) \geq 0$, if $x > 0$ and $y < 0$
4. $P'(x, y) \leq 0$, if $x > 0$ and $y > 0$

Thus, by considering the polynomial $P''(x, y) = P'(-x, y)$, we are getting a polynomial that satisfies the conditions of the claim, whose degree is strictly smaller than $\deg(P)$. This contradicts the choice of P . ■

By “shifting”, we obtain the following corollary to Claim 5.10.

Corollary 5.11 *Let $(x_0, y_0) \in \mathbb{R}^2$ and let $B \subseteq \mathbb{R}^2$ be a neighborhood of (x_0, y_0) . Then, there is no real polynomial $P : \mathbb{R}^2 \rightarrow \mathbb{R}$ such that for all $(x, y) \in B$ it holds*

1. $P(x, y) \geq 0$, if $x < x_0$ and $y < y_0$
2. $P(x, y) \leq 0$, if either $x > x_0$ or $y > y_0$

Proof: Assume towards contradiction that there exists a polynomial $P : \mathbb{R}^2 \rightarrow \mathbb{R}$ satisfying the foregoing conditions. Define a polynomial $P'(x, y) = P(x - x_0, y - y_0)$, and let $B' = \{(x, y) : (x + x_0, y + y_0) \in B\}$ be a neighborhood of $(0, 0)$. The set B' satisfies the conditions of Claim 5.10, whereas the polynomial P' violates its conclusion. This contradicts the hypothesis that the polynomial P exists. ■

Now, using Theorem 5.1 and Corollary 5.11, we show a negative result for a large family of classes of distributions that correspond to polytopes in $\Delta^{(r)}$. For sake of simplicity, we restrict ourselves to ternary distributions (i.e., $r = 3$).²⁷ A more general result will appear in Proposition 5.15.

Proposition 5.12 (a simple impossibility result): *For $\beta, \gamma \in (0, 1)$ let $\Pi_{\beta, \gamma}$ the class of distributions $\bar{q} = (q_1, q_2, q_3)$ with domain $\{1, 2, 3\}$, such that $q_1 \leq \beta$ and $q_2 \leq \gamma$; that is*

$$\Pi_{\beta, \gamma} = \{\bar{q} = (q_1, q_2, q_3) : q_1 \leq \beta, q_2 \leq \gamma\}.$$

If $\beta + \gamma < 1$, then the class $\Pi_{\beta, \gamma}$ does not have a POT.

Proof: Let $\beta, \gamma \in (0, 1)$ such that $\beta + \gamma < 1$, and assume, towards a contradiction, that $\Pi_{\beta, \gamma}$ has a POT. Then, according to Theorem 5.1, there is a polynomial $P(q_1, q_2, q_3) : \Delta^{(3)} \rightarrow \mathbb{R}$, such that

1. $P(q_1, q_2, q_3) \geq 0$, if $q_1 \leq \beta$ and $q_2 \leq \gamma$.
2. $P(q_1, q_2, q_3) < 0$, otherwise.

By substituting $q_3 = 1 - q_1 - q_2$, the polynomial P induces a bi-variate polynomial $P' : B \rightarrow \mathbb{R}$, where $B = \{(q_1, q_2) \in \mathbb{R}^2 : q_1, q_2 \geq 0 \text{ and } q_1 + q_2 \leq 1\}$, such that

1. $P'(q_1, q_2) \geq 0$, if $q_1 \leq \beta$ and $q_2 \leq \gamma$.
2. $P'(q_1, q_2) < 0$, otherwise.

Now, if $\beta + \gamma < 1$, then B is a neighborhood of $(\beta, \gamma) \in \mathbb{R}^2$. Hence, the existence of such P' contradicts Corollary 5.11. Therefore the property $\Pi_{\beta, \gamma}$ does not have a POT. ■

²⁷Note that by Theorem 2.2 every class of binary distributions that corresponds to a polytope in $\Delta^{(2)}$ has a POT (since this polytope must be a segment in \mathbb{R}^2). Thus, $r \geq 3$ is necessary for a result of the kind of Proposition 5.12.

Remark: The restriction $\beta + \gamma < 1$ in Proposition 5.12 is necessary. Indeed, if $\beta + \gamma \geq 1$, then the class $\Pi_{\beta,\gamma} = \{\bar{q} = (q_1, q_2, q_3) : q_1 \leq \beta, q_2 \leq \gamma\}$ has a two-sided error POT, as can be seen by considering the polynomial $P_{\beta,\gamma}(q_1, q_2, q_3) = (\beta - q_1)(\gamma - q_2)$. The reader can easily verify that for all $(q_1, q_2, q_3) \in \Delta^{(3)}$ it holds that $P_{\beta,\gamma}(q_1, q_2, q_3) \geq 0$ if and only if $q_1 \leq \beta$ and $q_2 \leq \gamma$. Thus, by Theorem 5.1, for $\beta + \gamma \geq 1$, the property $\Pi_{\beta,\gamma}$ has a POT.

Next, we use a similar argument in presenting two classes of distributions that have two-sided error POTs, whose union does not have one.

Proposition 5.13 (failure of closure under non-disjoint union): *Fix $\beta, \gamma > 0$ that satisfy $\beta + \gamma < 1$. Let $\Pi_1 = \{\bar{q} = (q_1, q_2, q_3) : q_1 \geq \beta\}$ and $\Pi_2 = \{\bar{q} = (q_1, q_2, q_3) : q_2 \geq \gamma\}$ be classes of distributions with domain $\{1, 2, 3\}$. Then, both Π_1 and Π_2 have a two-sided error POT, while their union does not have one.*

Proof: Clearly Π_1 has a two-sided error POT that makes one query and has linear detection probability:²⁸ Specifically, let the test make one query to the given distribution and accept if and only if the outcome is 1. If $\bar{q} \in \Pi_1$, then the test accepts with probability at least β , while any distribution $\bar{q} \in \Delta^{(3)}$ that is ϵ -far from Π_1 is accepted with probability at most $\beta - \epsilon$. Similarly Π_2 has a two-sided error POT.

We prove that $\Pi_1 \cup \Pi_2 = \{\bar{q} = (q_1, q_2, q_3) : q_1 \geq \beta \text{ or } q_2 \geq \gamma\}$ does not have a POT by, essentially, repeating the argument in the proof of Proposition 5.12. Assume towards contradiction that $\Pi_1 \cup \Pi_2$ has a POT. Then, there is a polynomial $P : \Delta^{(3)} \rightarrow \mathbb{R}$, such that $P(q_1, q_2, q_3) \geq 0$ if and only if either $q_1 \geq \beta$ or $q_2 \geq \gamma$. By substituting $q_3 = 1 - q_1 - q_2$, the polynomial P induces a bi-variate polynomial $P' : B \rightarrow \mathbb{R}$, where $B = \{(q_1, q_2) \in \mathbb{R}^2 : q_1, q_2 \geq 0 \text{ and } q_1 + q_2 \leq 1\}$, such that

1. $P'(q_1, q_2) \geq 0$, if either $q_1 \geq \beta$ or $q_2 \geq \gamma$.
2. $P'(q_1, q_2) < 0$, if $q_1 < \beta$ and $q_2 < \gamma$.

By letting $P''(q_1, q_2) = -P'(q_1, q_2)$ we get a polynomial that violates the conclusion of Corollary 5.11. Hence, we conclude that the polynomials P' and P can not exist, thus contradicting the hypothesis that $\Pi_1 \cup \Pi_2$ has a POT. ■

Finally, in the following Proposition 5.15 we generalize Proposition 5.12. Specifically, Proposition 5.15 makes assertions regarding the boundaries of properties having a POT, which in turn may lead to the impossibility results (regarding POTs). We will need the following definition of boundary of a class of distributions.

Definition 5.14 (boundary of subsets to $\Delta^{(r)}$): *Let $\Pi \subseteq \Delta^{(r)}$ be a class of distributions with domain $[r]$. The boundary of Π , denoted by $\partial\Pi$, is the set of all distributions $\bar{q} = (q_1, \dots, q_r) \in \Delta^{(r)}$ that are arbitrary close both to Π and to $\Delta^{(r)} \setminus \Pi$. That is, $\bar{q} \in \partial\Pi$ if, for every $\epsilon > 0$, there is a distribution $\bar{q}' \in \Pi$ such that $\text{dist}(\bar{q}, \bar{q}') < \epsilon$, and there is a distribution $\bar{q}'' \in \Delta^{(r)} \setminus \Pi$ such that $\text{dist}(\bar{q}, \bar{q}'') < \epsilon$. In particular, the boundary of $\Delta^{(r)}$ is the empty set.*

Proposition 5.15 (on the boundaries of properties that have a POT): *Let $\Pi \subseteq \Delta^{(r)}$ be a property of distributions and suppose that Π has a POT.*

1. If $P : \Delta^{(r)} \rightarrow \mathbb{R}$ is a polynomial that satisfies $\Pi = \{\bar{q} \in \Delta^{(r)} : P(\bar{q}) \geq 0\}$, then $P(\bar{q}) = 0$ for all $\bar{q} \in \partial\Pi$. In particular, $\partial\Pi \subseteq \Pi$.

²⁸Indeed, this test is analogous to the first example that was presented in the introduction (and labeled “straight-forward”).

2. Let $S \subseteq \Delta^{(r)}$ be a non-trivial segment.²⁹ If $S \subseteq \partial\Pi$, then Π contains the entire line that goes through S (restricted to $\Delta^{(r)}$).
3. More generally, let $S \subseteq \Delta^{(r)}$ be a convex set³⁰ containing more than a single point. If $S \subseteq \partial\Pi$, then Π contains the entire affine hull³¹ of S (restricted to $\Delta^{(r)}$).

Proposition 5.12 can be derived by using a special case of the second item: Indeed, for $\beta + \gamma < 1$, the segment $S = \{(\beta, t, 1 - \beta - t) : t \in [0, \gamma]\}$ is contained in the boundary of $\Pi_{\beta, \gamma}$ (which consists of $S \cup \{(t, \gamma, 1 - t - \gamma) : t \in [0, \beta]\}$). On the other hand, the line $\{\bar{q} = (\beta, t, 1 - \beta - t) : t \in [0, 1 - \beta]\}$ that contains S is not contained in $\Pi_{\beta, \gamma}$ (since $(\beta, 1 - \beta, 0) \notin \Pi_{\beta, \gamma}$, as $1 - \beta > \gamma$). Thus, $\Pi_{\beta, \gamma}$ violates the assertion of this item, and so it can not have a POT. More generally, Proposition 5.15 implies, for example, that a large family of classes of distributions $\Pi \subseteq \Delta^{(r)}$ that correspond to polytopes do not have POTs (see Corollary 5.16 below).

Proof: Let $P : \Delta^{(r)} \rightarrow \mathbb{R}$ be a polynomial such that $\Pi = \{\bar{q} \in \Delta^{(r)} : P(\bar{q}) \geq 0\}$, and let $\bar{q} \in \partial\Pi$. Then, for every $\epsilon > 0$, there is some $\bar{q}' \in \Delta^{(r)}$ such that $\text{dist}(\bar{q}, \bar{q}') < \epsilon$ and $\bar{q}' \in \Pi$, and hence $P(\bar{q}') \geq 0$. Therefore, by continuity of P , we have $P(\bar{q}) \geq 0$. Similarly, for every $\epsilon > 0$, there is some $\bar{q}'' \in \Delta^{(r)}$ such that $\text{dist}(\bar{q}, \bar{q}'') < \epsilon$ and $\bar{q}'' \notin \Pi$, and hence $P(\bar{q}'') < 0$. Therefore, again, by continuity of P , we have $P(\bar{q}) \leq 0$. We conclude that $P(\bar{q}) = 0$ for all $\bar{q} \in \partial\Pi$, and in particular $\partial\Pi \subseteq \Pi$. This completes the proof of the first part.

For the second part let $S \subseteq \partial\Pi$ be a segment, and let $P : \Delta^{(r)} \rightarrow \mathbb{R}$ be a polynomial such that $\Pi = \{\bar{q} \in \Delta^{(r)} : P(\bar{q}) \geq 0\}$ (as guaranteed by Theorem 5.1). By the first part we have $P(\bar{q}) = 0$ for every $\bar{q} \in S$. Consider the restriction of P to the line containing S . The univariate polynomial describing this line is zero on the (infinitely many points residing on the) segment S , and thus it must be zero on the entire line containing S . That is, P is non-negative on the entire line containing S , and so it must be the case that the entire line (restricted to $\Delta^{(r)}$) is contained in Π .

For the third part denote by $\text{aff}(S)$ the affine hull of S . For any $\bar{p} \in \text{aff}(S)$ there are two distinct points $\bar{q}, \bar{q}' \in S$ such that \bar{p} belongs to the line containing both \bar{q} and \bar{q}' .³² Since S is convex, the segment defined by \bar{q} and \bar{q}' is contained in S . Thus, by the second part, Π contains the entire line that goes through S , and in particular contains the points \bar{p} itself. Therefore $\text{aff}(S) \subseteq \Pi$ as required. ■

Corollary 5.16 (in general, polytopes have no POT): *Let $r \geq 3$. Let $\Pi \subset \Delta^{(r)}$ be a non-trivial polytope³³ that has a vertex \bar{v} that is internal to $\Delta^{(r)}$ (i.e., \bar{v} is not a convex combination of $\Pi \setminus \{\bar{v}\}$ and all coordinates of \bar{v} are positive). Then, Π does not have a POT.*

Proof: We shall prove that there exists a non-trivial segment S that satisfies the conditions (1) S is contained in $\partial\Pi$, and (2) the entire line that passes through S is not contained in Π . Thus, by the contrapositive of the second item of Proposition 5.15, we will conclude that Π does not have a POT. Actually, as we shall shortly see, such a segment can be found under more general conditions.

²⁹A non-trivial segment S is defined by two distinct points $\bar{p}, \bar{q} \in \Delta^{(r)}$, and is the set of all convex combinations of these points; that is, $S = \{\lambda\bar{p} + (1 - \lambda)\bar{q} : \lambda \in [0, 1]\}$. The line that goes through S is the set $\{\lambda\bar{p} + (1 - \lambda)\bar{q} : \lambda \in \mathbb{R}\}$.

³⁰A set $S \subseteq \mathbb{R}^r$ is said to be convex if for every $\bar{p}, \bar{q} \in S$ and every $\lambda \in [0, 1]$, it holds $\lambda\bar{p} + (1 - \lambda)\bar{q} \in S$. In particular, every segment is a convex set.

³¹For a set $S \subseteq \mathbb{R}^r$, the affine hull of S is the smallest affine subspace of \mathbb{R}^r containing S . For example, if S contains just two points (or the segment between them), then the affine hull of S is the line that goes through these points.

³²This fact follows from the convexity of S and the minimality of $\text{aff}(S)$.

³³A non-trivial polytope Π is a set in \mathbb{R}^r of more than a single point (i.e., $|\Pi| > 1$) that satisfy a system of linear inequalities; that is, each point in Π satisfies all inequalities.

We first dispose of the case that the polytope is a line segment. In that case, the polytope itself may serve as the line segment S , and we are done by Proposition 5.15. Otherwise, the polytope has (non-trivial) faces.

Recall that Π can be expressed as the intersection of t half-spaces (corresponding to linear conditions), such that the i -th half-space is given by

$$H_i = \{(q_1, \dots, q_r) \in \mathbb{R}^r : \sum_{j \in [r]} \alpha_j^{(i)} q_j \leq \beta_i\}.$$

Claim 5.16.1 *There exists $i_1, i_2 \in [t]$ and points $\bar{p}, \bar{q} \in \Pi$ that are internal to $\Delta^{(r)}$ such that \bar{p} satisfies $\sum_{j \in [r]} \alpha_j^{(i_k)} p_j = \beta_{i_k}$ for $k = 1, 2$, whereas \bar{q} satisfies $\sum_{j \in [r]} \alpha_j^{(i_1)} q_j = \beta_{i_1}$ and $\sum_{j \in [r]} \alpha_j^{(i_2)} q_j < \beta_{i_2}$.*

Indeed, we may use $\bar{p} = \bar{v}$ and let $\bar{q} \neq \bar{v}$ be an arbitrary point that is internal to a face that contains \bar{v} but is on the boundary of the polytope (w.r.t $\Delta^{(r)}$). The existence of such a face follows from the hypothesis that \bar{v} has no zero coordinates and is a vertex of a multi-dimensional polytope (which resides inside $\Delta^{(r)}$). Note that picking \bar{q} as internal to this face, guarantees that \bar{q} too has no zero coordinates. Without loss of generality, we may assume that $i_1 = 1$ and $i_2 = 2$, and proceed with any two points \bar{p} and \bar{q} that satisfy the above claim.

We first observe that the line segment, S , defined by \bar{p} and \bar{q} is on the boundary of Π , because for every $\bar{p}' \in S$ and $\epsilon > 0$ there exists a vector \bar{d} that is shorter than ϵ such that $\bar{p}' + \bar{d} \in \Delta^{(r)} \setminus \Pi$ (i.e., in particular, $\sum_{j \in [r]} \alpha_j^{(1)} (p'_j + d_j) > \beta_1$). Next, observe that the entire line that passes through S is not contained in Π , because for some $\epsilon > 0$ it holds that $\bar{p}' = (1 + \epsilon)\bar{p} + \epsilon\bar{q} \in \Delta^{(r)} \setminus \Pi$ (i.e., in particular, $\sum_{j \in [r]} \alpha_j^{(2)} p'_j > \beta_2$). Thus, the corollary follows by (the second item of) Proposition 5.15. ■

6 More on Graph Properties in the Adjacency Rep. Model

In this section we utilize the POTs for general distributions, presented in Section 5, in order to obtain additional results regarding POTs for graph properties (in the adjacency representation model, aka the dense graph model). Let us start with a useful claim regarding the distributions of random induced subgraphs of two close graphs.

Claim 6.1 *Let $H = ([n], F)$ be a fixed graph. For every two graphs $G = ([N], E)$ and $G' = ([N], E')$ such that $N \geq n$, if G and G' are ϵ -close, then*

$$|\rho_H(G) - \rho_H(G')| = \frac{|\text{ind}_H(G) - \text{ind}_H(G')|}{\binom{N}{n}} \leq \epsilon n^2,$$

where $\text{ind}_H(G) = \rho_H(G) \cdot \binom{N}{n}$ denotes the number of induced copies of H in G .

The straightforward proof of this claim appears in Appendix A.6. Note that the converse of Claim 6.1 does not hold in general (e.g., consider a regular N -vertex graph consisting of a single super-cycle of length $2n + 2$ versus one that consists of two disjoint super-cycles of length $n + 1$).³⁴ However, in certain special cases that we shall encounter, a (quantitatively weaker) converse does hold.

³⁴See definition preceding Proposition 3.4.

Testing collections of cliques. We consider subclasses of the property called “clique collection”, which was studied in [GR09a, GR09b]. A graph is a clique collection if it consists of isolated cliques, and we shall be interested in graphs that are further restricted.

Definition 6.2 (density-parameterized clique collections): *Denote by $\mathcal{CC}^{\leq t}$ the class of all graphs that consist of at most t isolated cliques. For $\bar{\rho} = (\rho_1, \dots, \rho_t) \in \Delta^{(t)}$, where $\Delta^{(t)}$ is as in Eq. (8), denote by $\mathcal{CC}(\bar{\rho})$ the class of graphs that consists of exactly t isolated cliques of densities ρ_1, \dots, ρ_t . For $D \subseteq \Delta^{(t)}$ define $\mathcal{CC}(D) = \cup_{\bar{\rho} \in D} \mathcal{CC}(\bar{\rho})$.*

More generally, let $D : \mathbb{N} \rightarrow \mathcal{P}(\Delta^{(t)})$ be a function that assigns to each $N \in \mathbb{N}$ a subset of $\Delta^{(t)}$. Denote by $\mathcal{CC}(D)$ the subclass of $\mathcal{CC}^{\leq t}$ such that an N -vertex graph G that consists of t isolated cliques of densities ρ_1, \dots, ρ_t belongs to $\mathcal{CC}(D)$ if and only if $(\rho_1, \dots, \rho_t) \in D(N)$.

We first show in Proposition 6.4 that for any finite set $D \subseteq \Delta^{(t)}$ the property $\mathcal{CC}(D)$ has a POT. This result generalizes Proposition 3.3, which deals with a special case where all cliques are required to have equal sizes (i.e., $\rho_i = \frac{1}{t}$ for all $i \in [t]$).³⁵ Next, we restrict ourselves to $t = 2$ and construct POTs for a larger family of subclasses of $\mathcal{CC}^{\leq 2}$. Specifically, we consider $D : \mathbb{N} \rightarrow \mathcal{P}(\Delta^{(2)})$ such that for every $N \in \mathbb{N}$ the set $D(N)$ consists of a constant number of intervals (or segments) of $\Delta^{(2)}$. We obtain POTs for $\mathcal{CC}(D)$ in two extreme cases: (1) when these intervals cover a relatively small portion of $\Delta^{(2)}$ (see Proposition 6.7), and (2) when these intervals cover almost all of $\Delta^{(2)}$ (see Proposition 6.9). In contrast, when these intervals cover a (non-trivial) constant fraction of $\Delta^{(2)}$, no POT exists for $\mathcal{CC}(D)$ (see Proposition 6.12).

6.1 Testing $\mathcal{CC}(D)$ for finite sets $D \subseteq \Delta^{(t)}$.

Let us start with by sketching the proposed POT for $\mathcal{CC}(D)$, where $D \subseteq \Delta^{(t)}$ is a finite set. On an input graph G , the tester checks that the distribution of the subgraphs of G induced by $t + 1$ vertices fits a distribution that would have been observed when considering some graph in $\mathcal{CC}(D)$. (This fit will be checked by a distribution tester as described in Section 5.) In particular, $G \in \mathcal{CC}^{\leq t}$ must be $\{P_3, I_{t+1}\}$ -free, where P_3 is a path with 3 vertices, and I_{t+1} is an independent set of $t + 1$ vertices. Now, consider two cases for G that is far from $\mathcal{CC}(D)$.

1. Suppose first that G is far from the $\mathcal{CC}^{\leq t}$. Then, by the result of [GR09b], the graph G contains many induced copies of either P_3 or I_{t+1} , which means that our test will observe a distribution different from any of the ones expected (i.e., corresponding to some graph in $\mathcal{CC}(D)$). Thus, the distribution tester will accept with probability that is noticeably smaller than its threshold probability.
2. Suppose now that G belongs to $\mathcal{CC}^{\leq t}$ (or is very close to it), and yet is far from $\mathcal{CC}(D)$. Intuitively, in such a case the distribution induced by $(t + 1)$ -vertex subgraphs of G is far from all distributions induced by $\mathcal{CC}(D)$. (This intuition, which reverses Claim 6.1 in the special case of graphs in $\mathcal{CC}^{\leq t}$, is proved in Lemma 6.3.2.) Thus, again, the distribution tester will also accept with probability that is noticeably smaller than the threshold probability.

Hence, for any $D \subseteq \Delta^{(t)}$ (not merely a finite one), constructing a POT for the graph property $\mathcal{CC}(D)$ reduces to constructing a POT for the class of distributions of random $(t + 1)$ -vertex induced subgraphs of graphs in $\mathcal{CC}(D)$, where the reduction amounts to sampling the distribution of

³⁵Actually, Proposition 3.3 refers to the complementary graphs (i.e., regular complete t -partite graphs).

subgraphs induced by a random set of $t + 1$ vertices in the graph.³⁶ Indeed, we shall rely on the above conclusion both in this subsection and in subsequent ones.

For sake of technical convenience (or simplicity), we shall consider a distribution that slightly differs from the one given by the induced subgraphs of G on $t + 1$ vertices. Specifically, we shall consider the distribution of subgraphs induced by selecting $t + 1$ random vertices in the graph with repetitions (rather than without repetitions), and treat the (rare) event in which two sampled vertices collide as if these vertices are different but connected by an edge.

Definition 6.3 (\mathcal{S}_G^k) *For a given graph $G = ([N], E)$, define distribution \mathcal{S}_G^k as following: The distribution is supported on unlabeled graphs with k vertices. The sampler picks k vertices $v_1, \dots, v_k \in [N]$ uniformly at random with repetitions, and outputs the graph $([k], E_S)$, where $(i, j) \in E_S$ if and only if either $(v_i, v_j) \in E$ or $v_i = v_j$.*

For $\bar{\rho} = (\rho_1, \dots, \rho_t) \in \Delta^{(t)}$ and a graph $G = ([N], E) \in \mathcal{CC}(\bar{\rho})$, define $\mathcal{S}_{\bar{\rho}}^{t+1}$ to be \mathcal{S}_G^{t+1} .

Note that the latter definition is independent of N , and depends only on $\bar{\rho}$.

For a graph $H = ([k], E_H)$, we shall use interchangeably the notations $\Pr[\mathcal{S}_G^k = H]$ and $\mathcal{S}_G^k(H)$ to denote the probability that \mathcal{S}_G^k is isomorphic to H . For example, if $G \in \mathcal{CC}(\bar{\rho})$, we may write $\mathcal{S}_G^{t+1}(I_{t+1}) = 0$, $\mathcal{S}_G^{t+1}(K_{t+1}) = \sum_{i \in [t]} \rho_i^{t+1}$, and $\Pr[\mathcal{S}_G^{t+1} \text{ contains induced } P_3] = 0$. One can easily adapt the proof of Claim 6.1 to show an analogous claim for \mathcal{S}_G^k (and all H 's):

Claim 6.3.1 *Fix $k \geq 2$. For every two graphs $G = ([N], E)$ and $G' = ([N], E')$, if G and G' are ϵ -close, then for every graph H on k vertices it holds that*

$$\left| \mathcal{S}_G^k(H) - \mathcal{S}_{G'}^k(H) \right| \leq \epsilon k^2.$$

Hence, the statistical distance between \mathcal{S}_G^k and $\mathcal{S}_{G'}^k$ is bounded by

$$\text{dist}(\mathcal{S}_G^k, \mathcal{S}_{G'}^k) \leq O(\epsilon),$$

where the constant in the $O()$ notation depends only on k .

A key step in the proof of Proposition 6.4 (and other results in this section), is that for graphs in $\mathcal{CC}^{\leq t}$ (a weak quantitative version of) the inverse claim also holds; that is, if $G, G' \in \mathcal{CC}^{\leq t}$ and $\text{dist}(\mathcal{S}_G^{t+1}, \mathcal{S}_{G'}^{t+1})$ is small, then the graphs G and G' are close. Quantitatively, we prove the following lemma.

Lemma 6.3.2 *Fix $t \geq 2$ and let $\bar{\rho} = (\rho_1, \dots, \rho_t)$. Let $G' = ([N], E') \in \mathcal{CC}^{\leq t}$, and assume that $\text{dist}(\mathcal{S}_{\bar{\rho}}^{t+1}, \mathcal{S}_{G'}^{t+1}) \leq \epsilon'$. Then, G' is $O(\sqrt[t]{\epsilon'})$ -close to $\mathcal{CC}(\bar{\rho})$, where the constant in the $O()$ notation depends only on t .*

The lemma generalizes the structural claims of Proposition 3.3, which deals with a special case in which all ρ_i 's are equal. We defer the proof of the lemma to Appendix A.7, and show how it implies existence of a POT for $\mathcal{CC}(D)$, where D is a finite subset of $\Delta^{(t)}$.

Proposition 6.4 (POT for $\mathcal{CC}(D)$ when $D \subseteq \Delta^{(t)}$ is finite): *Fix $t \geq 1$ and let $D \subseteq \Delta^{(t)}$ be a finite set of size k . Then, the property $\mathcal{CC}(D)$ has a POT that makes $k \cdot (t^2 + t)$ queries and has polynomial detection probability.*

³⁶Actually, it suffices to construct a POT for a class of distributions that contain all distributions induced by graphs in $\mathcal{CC}(D)$ and none of the distributions induced by graphs that are not in $\mathcal{CC}(D)$. This relaxation of the POT-construction task simplifies our exposition.

Proof: On input G the POT for $\mathcal{CC}(D)$ tests that the distribution \mathcal{S}_G^{t+1} belongs to the set $\mathcal{S}_D = \{\mathcal{S}_{\bar{p}}^{t+1} : \bar{p} \in D\}$. That is, by considering distributions whose domain is the set of unlabeled graphs with $t + 1$ vertices, our goal is to test that \mathcal{S}_G^{t+1} belongs to the finite set \mathcal{S}_D . Using Corollary 5.7, we obtain a test that takes $2k$ independent samples of \mathcal{S}_G^{t+1} , accepts all distributions in \mathcal{S}_D with probability at least c (for some $c > 0$), and has quadratic detection probability. Therefore, for $G \in \mathcal{CC}(D)$, the test accepts G with probability at least c . Since each sample from \mathcal{S}_G^{t+1} requires $\binom{t+1}{2}$ queries to G , the test makes total of $2k \binom{t+1}{2}$ queries to G .

Assume that on input $G = ([N], E)$ the test accepts with probability $c - \epsilon$. Our aim is to show that G is $O(\epsilon^{1/t^2})$ -close to $\mathcal{CC}(D)$. We first note that by Corollary 5.7, if the test accepts with probability $c - \epsilon$, then there is some $\bar{p} \in D$ such that

$$\text{dist}(\mathcal{S}_G^{t+1}, \mathcal{S}_{\bar{p}}^{t+1}) = O(\sqrt{\epsilon}). \quad (16)$$

In particular

$$\Pr[\mathcal{S}_G^{t+1} = I_{t+1}] = O(\sqrt{\epsilon}) \text{ and } \Pr[\mathcal{S}_G^k \text{ contains induced } P_3] = O(\sqrt{\epsilon}).$$

Thus, by [GR09b, Proposition 4.11], the graph G is $O(\epsilon^{1/t})$ -close to $\mathcal{CC}^{\leq t}$. Let us fix a graph $G' = ([N], E') \in \mathcal{CC}^{\leq t}$ such that G' is $O(\epsilon^{1/t})$ -close to G . Then, according to Claim 6.3.1, the statistical distance between the distributions \mathcal{S}_G^{t+1} and $\mathcal{S}_{G'}^{t+1}$ is

$$\text{dist}(\mathcal{S}_G^{t+1}, \mathcal{S}_{G'}^{t+1}) = O(\epsilon^{1/t}). \quad (17)$$

Hence, by the triangle inequality, the statistical distance between $\mathcal{S}_{G'}^{t+1}$ and $\mathcal{S}_{\bar{p}}^{t+1}$ is bounded by

$$\text{dist}(\mathcal{S}_{G'}^{t+1}, \mathcal{S}_{\bar{p}}^{t+1}) \leq \text{dist}(\mathcal{S}_{G'}^{t+1}, \mathcal{S}_G^{t+1}) + \text{dist}(\mathcal{S}_G^{t+1}, \mathcal{S}_{\bar{p}}^{t+1}) \leq O(\epsilon^{1/t}), \quad (18)$$

where in the second inequality we used Eq. (16) and Eq. (17).

So far we have shown that $G' = ([N], E') \in \mathcal{CC}^{\leq t}$ satisfies $\text{dist}(\mathcal{S}_{G'}^{t+1}, \mathcal{S}_{\bar{p}}^{t+1}) \leq O(\epsilon^{1/t})$ for some $\bar{p} \in D$. By applying Lemma 6.3.2 with G' and \bar{p} , we conclude that G' is $O(\epsilon^{1/t^2})$ -close to $\mathcal{CC}(\bar{p})$. Recalling that G' is $O(\epsilon^{1/t})$ -close to G , we infer, using the triangle inequality, that G is $O(\epsilon^{1/t^2})$ -close to $\mathcal{CC}(\bar{p})$. This completes the proof of the proposition. \blacksquare

Digest. For the sake of future use, it is good to distill the essence of the proof of Proposition 6.4. We start by establishing the following corollary to Lemma 6.3.2, and then conclude with a reduction of testing $\mathcal{CC}(D)$ to testing $\mathcal{S}_D^{t+1} \stackrel{\text{def}}{=} \{\mathcal{S}_G^{t+1} : G \in \mathcal{CC}(D)\}$. (Actually, we may also reduce the testing of $\mathcal{CC}(D)$ to testing any class of distributions that contains \mathcal{S}_D^{t+1} but does not contain any distribution \mathcal{S}_G^{t+1} such that $G \notin \mathcal{CC}(D)$.)

Corollary 6.5 *Fix $t \geq 2$ and $D \subseteq \Delta^{(t)}$. Suppose that $G = ([N], E)$ is ϵ' -close to $\mathcal{CC}^{\leq t}$ but ϵ -far from $\mathcal{CC}(D)$. Then, \mathcal{S}_G^{t+1} is $(\Omega(\epsilon^t) - O(\epsilon'))$ -far from \mathcal{S}_D^{t+1} , where the constants in the $\Omega()$ and $O()$ notation depend only on t .*

Proof: Let $G' \in \mathcal{CC}^{\leq t}$ be ϵ' -close to G , and suppose that $\mathcal{S}_{G'}^{t+1}$ is δ -close to \mathcal{S}_D^{t+1} . Then, by Lemma 6.3.2, the graph G' is $O(\delta^{1/t})$ -close to $\mathcal{CC}(D)$. It follows that G is $(\epsilon' + O(\delta^{1/t}))$ -close to $\mathcal{CC}(D)$, which implies $\delta = \Omega(\epsilon - \epsilon')^t$. On the other hand, by Claim 6.3.1, we have $\text{dist}(\mathcal{S}_G^{t+1}, \mathcal{S}_{G'}^{t+1}) = O(\epsilon')$, and hence \mathcal{S}_G^{t+1} is $(\Omega(\epsilon - \epsilon')^t - O(\epsilon'))$ -far from \mathcal{S}_D^{t+1} . The claim follows. \blacksquare

Corollary 6.6 *Fix $t \geq 2$ and $D \subseteq \Delta^{(t)}$. If $G = ([N], E)$ is ϵ -far from $\mathcal{CC}^{\leq t}$, then \mathcal{S}_G^{t+1} is $\Omega(\epsilon^{t^2})$ -far from \mathcal{S}_D^{t+1} .*

Proof: Let $\epsilon' = \epsilon^{2t}$. If G is ϵ' -far from $\mathcal{CC}^{\leq t}$, then \mathcal{S}_G^{t+1} is $\Omega((\epsilon')^{t/2})$ -far from \mathcal{S}_D^{t+1} (by [GR09b, Proposition 4.11]). On the other hand, if G is ϵ' -close to $\mathcal{CC}^{\leq t}$, then \mathcal{S}_G^{t+1} is $(\Omega(\epsilon^t) - O(\epsilon'))$ -far from \mathcal{S}_D^{t+1} , (by Corollary 6.5). In both cases the claim follows. ■

6.2 Testing $\mathcal{CC}(D)$ for unbounded sets $D(N) \subseteq \Delta^{(2)}$

In this section we restrict ourselves to subclasses of $\mathcal{CC}^{\leq 2}$, namely the graphs consisting of two disjoint cliques. Recall the previous section where we have shown a POT for $\mathcal{CC}(D)$ where $D \subseteq \Delta^{(t)}$ is a finite set. Here we extend Proposition 6.4 for the special case of $t = 2$ by considering subclasses of $\mathcal{CC}^{\leq 2}$ where we allow the number of N -vertex graphs in the subclass to grow (slowly) with N .

We introduce the following notation. Let $\alpha, \beta : \mathbb{N} \rightarrow [0, \frac{1}{2}]$ be two functions that satisfy $0 \leq \alpha(N) \leq \beta(N) \leq \frac{1}{2}$ for all $N \in \mathbb{N}$. For each $N \in \mathbb{N}$ define $D_{\alpha(N), \beta(N)} \subseteq \Delta^{(2)}$ to be the set of all pairs $(\rho, 1 - \rho) \in \Delta^{(2)}$ such that $\min(\rho, 1 - \rho) \in [\alpha(N), \beta(N)]$, and define $D_{\alpha, \beta} : \mathbb{N} \rightarrow \mathcal{P}(\Delta^{(2)})$ to be $D_{\alpha, \beta}(N) = D_{\alpha(N), \beta(N)}$. Then, the class $\mathcal{CC}(D_{\alpha, \beta})$ consists of all graphs in $\mathcal{CC}^{\leq 2}$ such that an N -vertex graph G belongs to $\mathcal{CC}(D_{\alpha, \beta})$ if and only if $G \in \mathcal{CC}(D_{\alpha(N), \beta(N)})$, i.e., the density of its smaller clique belongs to the interval $[\alpha(N), \beta(N)]$.

Similarly, let $\bar{\alpha}, \bar{\beta} : \mathbb{N} \rightarrow [0, \frac{1}{2}]^t$ such that $\alpha_1(N) \leq \beta_1(N) < \dots < \alpha_t(N) \leq \beta_t(N)$ for all $N \in \mathbb{N}$. Define $D_{\bar{\alpha}, \bar{\beta}} : \mathbb{N} \rightarrow \Delta^{(2)}$ to be $D_{\bar{\alpha}, \bar{\beta}}(N) = \cup_{i \in [t]} D_{\alpha_i, \beta_i}(N)$. That is, $(\rho, 1 - \rho) \in D_{\bar{\alpha}, \bar{\beta}}(N)$ if and only if $\min(\rho, 1 - \rho) \in [\alpha_i(N), \beta_i(N)]$ for some $i \in [t]$. Then, the class $\mathcal{CC}(D_{\bar{\alpha}, \bar{\beta}})$ consists of all graphs in $\mathcal{CC}^{\leq 2}$ such that an N -vertex graph belongs to $\mathcal{CC}(D_{\bar{\alpha}, \bar{\beta}})$ if and only if the density of the smaller clique belongs to $[\alpha_i(N), \beta_i(N)]$ for some $i \in [t]$, i.e., $\mathcal{CC}(D_{\bar{\alpha}, \bar{\beta}}) = \cup_{i \in [t]} \mathcal{CC}(D_{\alpha_i, \beta_i})$.

We show in Proposition 6.7 that for any $d \in (0, 1]$ and for any $\bar{\alpha}, \bar{\beta} : \mathbb{N} \rightarrow [0, \frac{1}{2}]$ as above that satisfy $\beta_i(N) - \alpha_i(N) \leq N^{-d}$ for all $i \in [t]$ the class $\mathcal{CC}(D_{\bar{\alpha}, \bar{\beta}})$ has a POT whose parameters depend on d and t . The special case of $d = 1$ corresponds to $D_{\bar{\alpha}, \bar{\beta}}$ being a finite set, which has already been covered in Proposition 6.4. Therefore, the result of this section generalizes Proposition 6.4 in the special case of $\mathcal{CC}^{\leq 2}$.

Proposition 6.7 (POT for $\mathcal{CC}(D_{\bar{\alpha}, \bar{\beta}})$, when $\sum_i \beta_i(N) - \alpha_i(N) < N^{-\Omega(1)}$): *Let $\bar{\alpha}, \bar{\beta} : \mathbb{N} \rightarrow [0, \frac{1}{2}]$ and $d \in (0, 1]$ be such that for every $N \in \mathbb{N}$ it holds that $0 \leq \beta_i(N) - \alpha_i(N) \leq N^{-d}$ for all $i \in [t]$. Then, the class $\mathcal{CC}(D_{\bar{\alpha}, \bar{\beta}})$ has a POT that makes $O(t/d)$ queries, and has detection probability $\varrho(\epsilon) = \Omega(\epsilon^{O(t/d)})$.*

For the sake of simplicity we shall first prove Proposition 6.7 in the special case of $t = 1$, and later discuss the generalization to larger values of t .

Proposition 6.8 (Proposition 6.7, the case of $t = 1$): *Let $\alpha, \beta : \mathbb{N} \rightarrow [0, \frac{1}{2}]$ and $d \in (0, 1]$ be such that for every $N \in \mathbb{N}$ it holds that $0 \leq \beta(N) - \alpha(N) \leq N^{-d}$. Then, the class $\mathcal{CC}(D_{\alpha, \beta})$ has a POT that makes $O(1/d)$ queries, and has detection probability $\varrho(\epsilon) = \Omega(\epsilon^{O(1/d)})$.*

Proof: In light of Corollary 6.6, it suffices to present a POT for the class of distributions $\mathcal{S}_{D_{\alpha, \beta}}^3$ (which equals $\{\mathcal{S}_G^3 : G \in D_{\alpha, \beta}\}$), since testing G (for the graph property $\mathcal{CC}(D_{\alpha, \beta})$) is reduced to testing the distribution \mathcal{S}_G^3 (induced by random 3-subgraphs of the input graph G). We identify such distribution with $\mathcal{S}_G^3 = (q_0, q_1, q_2, q_3) \in \Delta^{(4)}$, where q_i denotes the density of the subgraph with 3 vertices and i edges in G . The existence of a POT for $\mathcal{S}_{D_{\alpha, \beta}}^3 \subseteq \Delta^{(4)}$ is proved by presenting a polynomial $P : \Delta^{(4)} \rightarrow \mathbb{R}$ such that P is non-negative on distributions in $\mathcal{S}_{D_{\alpha, \beta}}^3$ and is negative for distributions that are not in $\mathcal{S}_{D_{\alpha, \beta}}^3$. The POT for $\mathcal{S}_{D_{\alpha, \beta}}^3$ is obtained by invoking Theorem 5.1.

(Actually, we shall present a POT for a superset of $\mathcal{S}_{D_{\alpha,\beta}}^3$ that contains no distribution \mathcal{S}_G^3 such that $G \notin \mathcal{CC}(D_{\alpha,\beta})$.)

Denote by G_ρ the N -vertex graph consisting of two cliques of densities ρ and $1 - \rho$, and let $\kappa_3(\rho) = \rho^3 + (1 - \rho)^3$ be the K_3 -density in G_ρ . Let $m(N) = \frac{\kappa_3(\alpha(N)) + \kappa_3(\beta(N))}{2}$ be the average between $\kappa_3(\alpha(N))$ and $\kappa_3(\beta(N))$, and let $f(N) = \frac{|\kappa_3(\alpha(N)) - \kappa_3(\beta(N))|}{2}$ be the distance from $m(N)$ to each of $\kappa_3(\alpha(N))$ and $\kappa_3(\beta(N))$. The hypothesis $0 \leq \beta(N) - \alpha(N) \leq N^{-d}$ implies that $f(N) < 3 \cdot N^{-d}$. Using the above notation, it suffices to construct a polynomial $\mathbf{P} : \Delta^{(4)} \rightarrow \mathbb{R}$ such that $\mathbf{P}(q_0, q_1, q_2, q_3)$ is non-negative if and only if (1) $q_0 + q_2 < N^{-3}$ and (2) $|q_3 - m(N)| \leq f(N)$. When applying \mathbf{P} to a distribution induced by 3-subgraphs of any graph G condition (1) implies that $G \in \mathcal{CC}^{\leq 2}$, and condition (2) implies that G has the right K_3 -density. Now, define $\mathbf{P} : \Delta^{(4)} \rightarrow \mathbb{R}$ to be

$$\mathbf{P}(q_0, q_1, q_2, q_3) = f(N)^k - (q_3 - m(N))^k - q_0 - q_2, \quad (19)$$

where $k \in 2\mathbb{N}$ is an even integer such that $f(N)^k < N^{-3}/3$. By our choice of $f(N)$, and using the fact that $f(N) < 3N^{-d}$, it is enough to take $k = O(1/d)$. Note that \mathbf{P} is a polynomial of total degree $k = O(1/d)$, and all its coefficients are upper bounded by a constant that does not depend on N , e.g. by $\binom{k}{k/2} < 2^{O(1/d)}$. Thus, although the polynomial \mathbf{P} depends on N , its degree and an upper bound on its coefficient sizes do not depend on N . This observation is important, because we want to use the POT derived from \mathbf{P} in order to derive a POT for $\mathcal{CC}(D_{\alpha,\beta})$, while the latter POT invokes the former POT with a varying value of N .

Still, by Theorem 5.1, there exists a distribution tester \mathcal{T} that makes k queries, whose acceptance probability when given a distribution $\bar{q} = (q_0, q_1, q_2, q_3) \in \Delta^{(4)}$ is

$$\Pr[\mathcal{T} \text{ accepts } \bar{q}] = 0.5 + \delta \cdot (f(N)^k - (q_3 - m(N))^k - q_0 - q_2) \quad (20)$$

for some absolute constant $\delta > 0$ that only depends on d (since all coefficients of the polynomial \mathbf{P} have absolute value at most $2^{O(1/d)}$).³⁷ We cannot rely on Theorem 5.1 in order to obtain a bound on the detection probability of \mathcal{T} , since the bound provided in the theorem may depend arbitrarily on the polynomial \mathbf{P} , which in turn depends on N . Instead, we shall lower bound the detection probability of \mathcal{T} by directly referring to Eq. (20).

But first, we verify that, for every graph G , the polynomial \mathbf{P} is non-negative on $(q_0, q_1, q_2, q_3) = \mathcal{S}_G^3$ if and only if $G \in \mathcal{CC}(D_{\alpha,\beta})$.

Indeed, suppose first that $G = ([N], E) \in \mathcal{CC}(D_{\alpha,\beta})$. Then, the distribution $\mathcal{S}_G^3 = (q_0, q_1, q_2, q_3)$ induced by 3-subgraph of G satisfies $q_0 = q_2 = 0$ (since $G \in \mathcal{CC}^{\leq 2}$) and $|q_3 - m(N)| \leq f(N)$ (since $G = G_\rho$ for some $\rho \in [\alpha(N), \beta(N)]$, and so $q_3 = \kappa_3(\rho) \in [\kappa_3(\alpha(N)), \kappa_3(\beta(N))] = [m(N) \pm f(N)]$). Therefore, by Eq. (19), we have $\mathbf{P}(q_0, q_1, q_2, q_3) \leq f(N)^k - f(N)^k - 0$.

On the other hand, suppose that for some N -vertex graph G the polynomial \mathbf{P} is non-negative on the distribution $\mathcal{S}_G^3 = (q_0, q_1, q_2, q_3)$. Since k is even, this implies that $q_0 + q_2 \leq f(N)^k < N^{-3}$ as well as $|q_3 - m(N)| \leq f(N)$. Noting that q_i comes in multiples of N^{-3} , the fact that $q_0 + q_2 \leq f(N)^k < N^{-3}$ implies that $q_0 = q_2 = 0$ and hence $G \in \mathcal{CC}^{\leq 2}$ (since $\mathcal{CC}^{\leq 2}$ coincides with the class of graphs that are both P_3 -free and I_3 -free). Hence, $G = G_\rho$ for some $\rho \in [0, 0.5]$. Using $|q_3 - m(N)| \leq f(N)$, it follows that $\kappa_3(\rho) = q_3 \in [m(N) \pm f(N)] = [\kappa_3(\alpha(N)), \kappa_3(\beta(N))]$, which implies that $G \in \mathcal{CC}(D_{\alpha,\beta})$.

It is left to provide an upper bound on the value that \mathbf{P} takes on an arbitrary $(q_0, q_1, q_2, q_3) \in \Delta^{(4)}$ that is ϵ -far from the set of probabilities on which \mathbf{P} is non-negative. Again, it actually suffices to provide such a bound for distributions (q_0, q_1, q_2, q_3) that are of the form $\mathcal{S}_G^3 = (q_0, q_1, q_2, q_3)$

³⁷The proof of Theorem 5.1 asserts that $\delta > 2^{-k}/3B$, where B is an upper bound on the absolute value of all coefficients of the polynomial \mathbf{P} (and k is its degree).

for some N -vertex graph $G \in \mathcal{CC}(D_{\alpha,\beta})$. Suppose that \mathcal{S}_G^3 is at distance $\epsilon > 0$ from the non-negative region (of \mathbb{P}), which implies that it is ϵ -far from $\mathcal{S}_{D_{\alpha(N),\beta(N)}}^3$. Since the coordinates of \mathcal{S}_G^3 (and of any distribution in $\mathcal{S}_{D_{\alpha(N),\beta(N)}}^3$) are multiples of N^{-3} , it follows that $\epsilon \geq N^{-3}$. Now, let $(p_0, p_1, p_2, p_3) \in \mathcal{S}_{D_{\alpha(N),\beta(N)}}^3$ be closest to (q_0, q_1, q_2, q_3) . Recall that $(p_0, p_1, p_2, p_3) \in \mathcal{S}_{D_{\alpha(N),\beta(N)}}^3$ implies $p_0 = p_2 = 0$ and $|p_3 - m(N)| \leq f(N)$. Then, either $q_0 + q_2 \geq \epsilon/2$ or $|q_3 - p_3| \geq \epsilon/2$. In the first case (i.e., $q_0 + q_2 \geq \epsilon/2$), it follows that $\mathbb{P}(\mathcal{S}_G^3) \leq f(N)^k - \epsilon/2 < -\epsilon/6$ (since $f(N)^k < N^{-3}/3$ and $\epsilon \geq N^{-3}$). In the second case (i.e., $|q_3 - p_3| \geq \epsilon/2$), it must be that $|q_3 - m(N)| > f(N)$ (since, otherwise, $p'_0 = p'_2 = 0$ and $p'_3 = q_3$ yields $(p'_0, p'_1, p'_2, p'_3) \in \mathcal{S}_{D_{\alpha(N),\beta(N)}}^3$ that is closer to (q_0, q_1, q_2, q_3)). Likewise, it must be that $p_3 = m(N) + \text{sgn}(q_3 - m(N)) \cdot f(N)$. It follows that $|q_3 - m(N)| \geq f(N) + \epsilon/2$ (since $|p_3 - m(N)| = f(N) + |q_3 - p_3|$), which implies that $\mathbb{P}(\mathcal{S}_G^3) \leq f(N)^k - (f(N) + \epsilon/2)^k < -(\epsilon/2)^k$.

Finally, combining the above analysis with Corollary 6.6, we conclude that each graph in $\mathcal{CC}(D_{\alpha,\beta})$ is accepted with probability at least $1/2$, whereas each graph that is ϵ -far from $\mathcal{CC}(D_{\alpha,\beta})$ is accepted with probability at most $0.5 - \Omega(\epsilon^{O(1)})^k$. The claim follows. \blacksquare

Generalization to $t \geq 2$: We now show how the proof of Proposition 6.8 can be generalized to larger values of t , and so establish Proposition 6.7. The generalization follows the proof idea of Corollary 5.4 (regarding disjoint union of testable classes of distributions). For each $i \in [t]$, define the parameters $m_i(N) = \frac{\kappa_3(\alpha_i(N)) + \kappa_3(\beta_i(N))}{2}$ and $f_i(N) = \frac{|\kappa_3(\alpha_i(N)) - \kappa_3(\beta_i(N))|}{2}$ (analogously to the proof Proposition 6.8). Then, analogously to Eq. (19) define the polynomial $\mathbb{P} : \Delta^{(4)} \rightarrow \mathbb{R}$ to be

$$\mathbb{P}(q_0, q_1, q_2, q_3) = (-1)^{t+1} \cdot \prod_{i \in [t]} \left(f_i(N)^k - (q_3 - m_i(N))^k - q_0 - q_2 \right)$$

for $k \in 2\mathbb{N}$ chosen as in the proof of Proposition 6.8. By following the proof of Proposition 6.8 (and the proof of Corollary 5.3), the reader can verify that this polynomial \mathbb{P} yields a POT for $\mathcal{CC}(D_{\bar{\alpha},\bar{\beta}})$, which makes $O(t/d)$ queries and has detection probability $\varrho(\epsilon) = \Omega(\epsilon^{O(t/d)})$. We omit the details.

6.3 Testing $\mathcal{CC}(D)$ when $D(N) \subseteq \Delta^{(2)}$ is almost everything

Recall that by the characterization result of [GR09b], the class $\mathcal{CC}^{\leq 2}$ has a one-sided error POT. However, if we remove from $\mathcal{CC}^{\leq 2}$ even just a single N -vertex graph for every $N \in \mathbb{N}$, we will obtain a class that does not have a one-sided error POT. Nevertheless, as we prove in this section, such a class has a two-sided error POT. This fact is as a special case of Proposition 6.9.

Recall, that for $\bar{\alpha}, \bar{\beta} : \mathbb{N} \rightarrow [0, \frac{1}{2}]^t$ such that $\alpha_1(N) \leq \beta_1(N) < \dots < \alpha_t(N) \leq \beta_t(N)$ for all $N \in \mathbb{N}$, the class $\mathcal{CC}(D_{\bar{\alpha},\bar{\beta}}) = \cup_{i \in [t]} \mathcal{CC}(D_{\alpha_i, \beta_i})$ consists of all graphs in $\mathcal{CC}^{\leq 2}$ such that an N -vertex graph belongs to $\mathcal{CC}(D_{\bar{\alpha},\bar{\beta}})$ if and only if the density of the smaller clique belongs to $[\alpha_i(N), \beta_i(N)]$ for some $i \in [t]$. For example, the classes considered in the first paragraph of this section can be described as $\mathcal{CC}(D_{\bar{\alpha},\bar{\beta}})$ for some $\bar{\alpha}, \bar{\beta} : \mathbb{N} \rightarrow [0, \frac{1}{2}]^2$.

We show in Proposition 6.9 that for $\bar{\alpha}(N), \bar{\beta}(N) \in [0, \frac{1}{2}]^t$ as above, if for each $N \in \mathbb{N}$ almost all graphs on N vertices in $\mathcal{CC}^{\leq 2}$ belong to $\mathcal{CC}(D_{\bar{\alpha},\bar{\beta}})$, then $\mathcal{CC}(D_{\bar{\alpha},\bar{\beta}})$ has a two-sided error POT. In particular, the class of graphs obtained from $\mathcal{CC}^{\leq 2}$ by removing any constant number of graphs on N vertices (for every $N \in \mathbb{N}$) has a two-sided error POT.

Proposition 6.9 (POT for $\mathcal{CC}(D_{\bar{\alpha},\bar{\beta}})$, when $\sum_i \beta_i(N) - \alpha_i(N) > \frac{1}{2} - N^{-\Omega(1)}$): *Let $\bar{\alpha}, \bar{\beta} : \mathbb{N} \rightarrow [0, \frac{1}{2}]^t$ be such that for every $N \in \mathbb{N}$ it holds that $\alpha_1(N) \leq \beta_1(N) < \dots < \alpha_t(N) \leq \beta_t(N)$. If for some $d \in (0, 1]$ and all $N \in \mathbb{N}$ it holds that $\sum_{i \in [t]} (\beta_i(N) - \alpha_i(N)) > \frac{1}{2} - N^{-d}$, then the class $\mathcal{CC}(D_{\bar{\alpha},\bar{\beta}})$ has a POT that makes $6t$ queries and has polynomial detection probability (where the exponent of the polynomial depends on t and d).*

Similarly to the previous section, we shall first prove Proposition 6.9 in the special case of $t = 1$, and later discuss the generalization to larger values of t .

Proposition 6.10 (Proposition 6.9, the case of $t = 1$): *Let $\alpha, \beta : \mathbb{N} \rightarrow [0, \frac{1}{2}]$ and $d \in (0, 1]$ be such that for every $N \in \mathbb{N}$ it holds that $\alpha(N) < N^{-d}$ and $\beta(N) > \frac{1}{2} - N^{-d}$. Then, the class $\mathcal{CC}(D_{\alpha, \beta})$ has a POT that makes 6 queries, and has detection probability $\varrho(\epsilon) = \Omega(\epsilon^{8/d})$.*

Proof: In order to develop some intuition, note that the assumption $\alpha(N) < N^{-d}$ and $\beta(N) > \frac{1}{2} - N^{-d}$ implies that no graph $G \in \mathcal{CC}^{\leq 2}$ is too far from $\mathcal{CC}(D_{\alpha, \beta})$. Indeed, suppose that G consists of two cliques of sizes K and $N - K$, with $K \leq N - K$, such that either $K/N < \alpha(N)$ or $K/N > \beta(N)$. Then, the graph G is N^{-d} -close to $\mathcal{CC}(D_{\alpha, \beta})$. Hence, graphs that are in $\mathcal{CC}^{\leq 2} \setminus \mathcal{CC}(D_{\alpha, \beta})$ are N^{-d} -close to $\mathcal{CC}(D_{\alpha, \beta})$, and it suffices to detect such graphs with probability $1/\text{poly}(N)$ (since this probability is polynomially related to its distance from the property being tested). On the other hand, any graph that is ϵ -far from $\mathcal{CC}(D_{\alpha, \beta})$ for $\epsilon > 2/N^d$ must be $\frac{\epsilon}{2}$ -far from $\mathcal{CC}^{\leq 2}$, and it suffices to detect such graphs with probability $\text{poly}(\epsilon)$. Thus, it is enough to design a tester that satisfies the following three conditions, with respect to some threshold c :

1. Graphs in $\mathcal{CC}(D_{\alpha, \beta})$ are accepted with probability at least c ;
2. Graphs that are ϵ -far from $\mathcal{CC}^{\leq 2}$ are accepted with probability lower than $c - \Omega(\epsilon^2)$;
3. Graphs that are not in $\mathcal{CC}(D_{\alpha, \beta})$ are accepted with probability lower than $c - \Omega(N^{-8})$.

For sake of simplicity, we consider first the simpler case in which either $\alpha = 0$ or $\beta = 1/2$. In this case, the test we propose and analyze examines the subgraph induced by three random vertices and accepts according to some (carefully chosen) predetermined probabilities. That is, we associate four probabilities with the four possible 3-vertex subgraphs that can be seen such that p_i denotes the probability that the test accepts when seeing a subgraph with i edges. Since we wish to accept only graphs in $\mathcal{CC}^{\leq 2}$, we may set $p_0 = p_2 = 0$ and $p_1, p_3 > 0$. Thus, in designing such a test, our only freedom is in the choice of $\min(p_1, p_3) > 0$, since without loss of generality we may have $\max(p_1, p_3) = 1$. Note that in order to satisfy the relation between Items 1 and 2, we must require $p_1 \approx p_3$ (e.g., $|p_1 - p_3| < 1/N^3$).³⁸ Picking p_1 and p_3 such that $|p_1 - p_3| < 1/N^3$, means that each N -vertex graph in $\mathcal{CC}^{\leq 2}$ is accepted with probability approximately p_1 , whereas N -vertex graphs that are ϵ -far from $\mathcal{CC}^{\leq 2}$ are accepted with probability at most $p_1 - \text{poly}(\epsilon)$, which follows from [GR09b, Proposition 4.11] using the bound $\epsilon > 1/N^2$ (i.e., each N -vertex graph not in a class is at least $1/N^2$ -far from it). Finally, setting $p_1 > p_3$ favors graphs in $\mathcal{CC}^{\leq 2}$ that consist of two cliques of (approximately) the same size, whereas $p_1 < p_3$ favors graphs in $\mathcal{CC}^{\leq 2}$ that consist of two cliques such that one is significantly larger than the other. Note, however, that this favoring amounts to at most $|p_1 - p_3|$, but nevertheless this will suffice for Item 3.

In the general case, where both $\alpha > 0$ and $\beta < 1/2$, we view the distribution induced by 3-subgraphs of G as an element of $\Delta^{(4)}$. Thus, the tester needs to check that this distribution lies in some (carefully chosen) predetermined subset of $\Delta^{(4)}$. Note that graphs in $\mathcal{CC}^{\leq 2}$ are associated with distributions $(q_0, q_1, q_2, q_3) \in \Delta^{(4)}$ such that $q_0 = q_2 = 0$, where q_i is the probability that a random induced 3-vertex graph has i edges. Furthermore, graphs in $\mathcal{CC}(D_{\alpha, \beta})$ are associated with

³⁸To see why $p_1 \approx p_3$ must hold, consider the graph G_ρ consisting of two cliques such that the smaller one has density $\rho \leq 1/2$. This graph is accepted with probability $\kappa_3(\rho) \cdot p_3 + (1 - \kappa_3(\rho)) \cdot p_1 = p_3 - (p_3 - p_1)(1 - \kappa_3(\rho))$, where $\kappa_3(\rho) = \rho^3 + (1 - \rho)^3 \in [0.25, 1]$ is the K_3 -density of G_ρ . On the other hand, a graph obtained from G_ρ by an omission of a single edge is accepted with probability that is $\Theta(N^{-2})$ smaller. Thus, $|p_3 - p_1| = O(N^{-2})$ is required for guaranteeing that both $G_{0.01}$ and $G_{0.49}$ are accepted with probability that is higher than either of the graphs obtained from them by omission of a single edge.

a segment $I_{\alpha,\beta}$ of the line $\{(0, q_1, 0, 1 - q_1) : q_1 \in [0, 1]\}$. Thus, the desired test may be thought of as a test of $I_{\alpha,\beta} \subseteq \Delta^{(4)}$ that is required to (1) accept points on $I_{\alpha,\beta}$ with probability at least c , (2) accept points that are ϵ -far from the entire line (which passes through $I_{\alpha,\beta}$) with probability at most $c - \text{poly}(\epsilon)$, and (3) accept points that are on the line but outside $I_{\alpha,\beta}$ with probability at most $c - \text{poly}(1/N)$. Corollary 5.8 suggests that such a test can be obtained by considering an ellipsoid that contains an $\Theta(1/N^3)$ -neighborhood of $I_{\alpha,\beta}$. (This ellipsoid has a very long axis in the direction of $I_{\alpha,\beta}$, and is very slim in all the directions orthogonal to $I_{\alpha,\beta}$.) Note, however, that in the analysis we cannot just invoke Corollary 5.8, because we need stronger bounds; i.e., for points that are far from the long axis of the ellipsoid we need bounds that do not depend on the volume of the ellipsoid. (Such bounds are readily obtained by using the ideas that underly the proof of Corollary 5.8.)

We now turn to the actual proof. Let us fix $N \in \mathbb{N}$, and denote for simplicity $\alpha = \alpha(N)$ and $\beta = \beta(N)$. We restrict ourselves to designing a POT for N -vertex graphs of $\mathcal{CC}(D_{\alpha,\beta})$. Similarly to the proof of Proposition 6.8 the POT for $\mathcal{CC}(D_{\alpha,\beta})$ relies on testing the distribution induced by 3-subgraphs of a given graph G . Specifically, we consider the distributions $\mathcal{S}_G^3 = (q_0, q_1, q_2, q_3) \in \Delta^{(4)}$ induced by 3-subgraphs of G as given in Definition 6.3, where q_i denotes the density of the subgraph with 3 vertices and i edges in G . As in the proof of Proposition 6.8 we show that there exists a polynomial $\mathbf{P} : \Delta^{(4)} \rightarrow \mathbb{R}$ with bounded coefficients such that \mathbf{P} is positive on distributions induced by graphs in $\mathcal{CC}(D_{\alpha,\beta})$, and is bounded below zero for distributions induced by graphs that are far from $\mathcal{CC}(D_{\alpha,\beta})$. The POT is derived from such polynomial by following the recipe given in Theorem 5.1.

For every $\rho \in [0, \frac{1}{2}]$ let G_ρ be the graph consisting of two cliques of densities ρ and $1 - \rho$. Then, the distribution induced by G_ρ is

$$\mathcal{S}_{G_\rho}^3 = \mathcal{S}_{(\rho, 1-\rho)}^3 = (0, 3\rho(1-\rho), 0, \rho^3 + (1-\rho)^3).$$

Thus, by considering the segment $I_{\alpha,\beta} \subseteq \Delta^{(4)}$ contained in the line $\{(0, q_1, 0, 1 - q_1) : q_1 \in [0, 1]\}$ between the points $\mathcal{S}_{(\alpha, 1-\alpha)}^3$ and $\mathcal{S}_{(\beta, 1-\beta)}^3$ we see that $G \in \mathcal{CC}(D_{\alpha,\beta})$ if and only if $\mathcal{S}_G^3 \in I_{\alpha,\beta}$.

Next, define an ellipsoid $\mathcal{E}_{\alpha,\beta}$ such that for every N -vertex graph G it holds that $\mathcal{S}_G^3 \in I_{\alpha,\beta}$ if and only if $\mathcal{S}_G^3 \in \mathcal{E}_{\alpha,\beta}$. This is done by taking an ellipsoid that contains only distributions that are $N^{-3}/2$ -close to $I_{\alpha,\beta}$ (since for every two distinct graphs induce distributions that are $\Omega(N^{-2})$ from each other). Specifically, let $\overline{m} = (m_0, m_1, m_2, m_3) = \frac{1}{2} \left(\mathcal{S}_{(\alpha, 1-\alpha)}^3 + \mathcal{S}_{(\beta, 1-\beta)}^3 \right)$ be the midpoint of the interval $I_{\alpha,\beta}$, and let r denote the L2-distance between this midpoint and either endpoints of the segment $I_{\alpha,\beta}$. (Note that for $\alpha \approx 0$ and $\beta \approx \frac{1}{2}$ we have $I_{\alpha,\beta} \approx [(0, 0, 0, 1), (0, 0.75, 0, 0.25)]$ and so $r \approx \sqrt{2 \cdot (0.75/2)^2} \approx 0.53$.) Since we wish the ellipsoid to contain $I_{\alpha,\beta}$ and no point (q_0, q_1, q_2, q_3) that is $(1/2N^3)$ -far from this segment, we relax the conditions $q_0 = q_2 = 0$ and $(q_1 - m_1)^2 + (q_3 - m_3)^2 \leq r^2$ into $(q_0^2 + q_2^2) \cdot (2N^3)^2 + (q_1 - m_1)^2/r^2 + (q_3 - m_3)^2/r^2 \leq 1$. This yields the following ellipsoid $\mathcal{E}_{\alpha,\beta}$ (which contains $I_{\alpha,\beta}$, has radius r in the direction of $I_{\alpha,\beta}$ and radius $1/2N^3$ in the orthogonal directions):

$$\mathcal{E}_{\alpha,\beta} = \left\{ (q_0, q_1, q_2, q_3) \in \Delta^{(4)} : 4N^6 \cdot q_0^2 + \frac{1}{r^2} \cdot (q_1 - m_1)^2 + 4N^6 \cdot q_2^2 + \frac{1}{r^2} \cdot (q_3 - m_3)^2 \leq 1 \right\}.$$

Note that by applying Corollary 5.8 we obtain a POT for $\mathcal{E}_{\alpha,\beta}$ whose detection probability depends on the parameters of the ellipsoid, and in particular depends on N . Since we are interested in detection probability that is independent of N , we use a more careful analysis of the tester for $\mathcal{E}_{\alpha,\beta}$. Specifically we define a polynomial \mathbf{P} that is non-negative on $\mathcal{E}_{\alpha,\beta}$ and is bounded below zero for

distributions that is far from $\mathcal{E}_{\alpha,\beta}$. The ellipsoid $\mathcal{E}_{\alpha,\beta}$ defines naturally a polynomial $\mathbf{P}' : \Delta^{(4)} \rightarrow \mathbb{R}$

$$\mathbf{P}'(q_0, q_1, q_2, q_3) = 1 - \left(4N^6 \cdot q_0^2 + \frac{1}{r^2} \cdot (q_1 - m_1)^2 + 4N^6 \cdot q_2^2 + \frac{1}{r^2} \cdot (q_3 - m_3)^2 \right)$$

such that $\mathbf{P}'(q_0, q_1, q_2, q_3) \geq 0$ if and only $\bar{q} \in \mathcal{E}_{\alpha,\beta}$. However, in order to apply Theorem 5.1 it is required that the coefficients of \mathbf{P}' be upper bounded independently of N . Such polynomial $\mathbf{P} : \Delta^{(4)} \rightarrow \mathbb{R}$ is obtained from \mathbf{P}' by normalizing the coefficients so that all coefficients of \mathbf{P} are bounded by 1 in absolute value.

$$\mathbf{P}(q_0, q_1, q_2, q_3) = \frac{1}{4N^6} - \left(q_0^2 + \frac{1}{4r^2N^6} \cdot (q_1 - m_1)^2 + q_2^2 + \frac{1}{4r^2N^6} \cdot (q_3 - m_3)^2 \right). \quad (21)$$

Clearly for every $\bar{q} = (q_0, q_1, q_2, q_3) \in \Delta^{(4)}$ it holds that $\mathbf{P}(q_0, q_1, q_2, q_3) \geq 0$ if and only $\bar{q} \in \mathcal{E}_{\alpha,\beta}$. By Theorem 5.1 since all coefficients of the polynomial \mathbf{P} of degree $\deg(\mathbf{P}) = 2$ are bounded in absolute value, there exists a POT \mathcal{T} that accepts all distributions $\bar{q} = (q_0, q_1, q_2, q_3) \in \Delta^{(4)}$ with probability

$$\Pr[\mathcal{T} \text{ accepts } \bar{q}] = 0.5 + \delta \cdot \mathbf{P}(q_0, q_1, q_2, q_3).$$

for some absolute constant $\delta > 0$ (that depends only on the degree of \mathbf{P} and the maximal coefficient of \mathbf{P}). We claim that such tester gives us a POT for $\mathcal{CC}(D_{\alpha,\beta})$ with threshold probability 0.5. Indeed, if $G \in \mathcal{CC}_N(D_{\alpha,\beta})$, then the distribution $\mathcal{S}_G^3 = (q_0, q_1, q_2, q_3)$ belongs to $\mathcal{E}_{\alpha,\beta}$, and hence $\mathbf{P}(q_0, q_1, q_2, q_3) \geq 0$, thus implying that the tester accepts G with probability at least $\frac{1}{2}$.

Assume now that an N -vertex graph G is ϵ -far from $\mathcal{CC}(D_{\alpha,\beta})$ for some $\epsilon > 0$, and let $\mathcal{S}_G^3 = (q_0, q_1, q_2, q_3)$. We shall prove that $\mathbf{P}(q_0, q_1, q_2, q_3) \leq -\Omega(\epsilon^{8/d})$, thus implying that the tester accepts G with probability that is at most $0.5 - \Omega(\epsilon^{8/d})$. The proof is partitioned into two cases, depending on ϵ .

Case 1: $\epsilon > 2N^{-d}$. Then, by the observation in the beginning of the proof the graph G is $\frac{\epsilon}{2}$ -far from $\mathcal{CC}^{\leq 2}$. Thus, by [GR09b, Proposition 4.11] it holds that $q_0 + q_2 = \Omega(\epsilon)$, and hence, using $\frac{\epsilon}{2} > N^{-d} \geq N^{-1}$, we get

$$\mathbf{P}(q_0, q_1, q_2, q_3) \leq \frac{1}{4N^6} - (q_0^2 + q_2^2) \leq \frac{1}{4N^6} - \Omega(\epsilon^2) = -\Omega(\epsilon^2),$$

where the constants in different $\Omega()$ notations might be different.

Case 2: $\epsilon \leq 2N^{-d}$. We shall show next that $G \notin \mathcal{CC}(D_{\alpha,\beta})$ implies $\mathbf{P}(q_0, q_1, q_2, q_3) = -\Omega(N^{-8})$, which is $-\Omega(\epsilon^{8/d})$ by the case hypothesis.

We consider two subcases.

1. If $G \notin \mathcal{CC}^{\leq 2}$, then either $q_0 \geq N^{-3}$ or $q_2 \geq N^{-3}$ (because $q_0 + q_2 > 0$ whereas each comes at multiples of N^{-3}). Therefore, $\mathbf{P}(q_0, q_1, q_2, q_3) \leq \frac{1}{4N^6} - \frac{1}{N^6}$, which is definitely smaller than $-\Omega(N^{-8})$.
2. If $G \in \mathcal{CC}^{\leq 2}$, then $q_0 = q_2 = 0$ and $q_1 = 1 - q_3$. In this case, G consists of two cliques of sizes K and $N - K$ such that either $K \leq \alpha N - 1$ or $\beta N + 1 < K \leq \lfloor \frac{N}{2} \rfloor$. Assume that $K \geq \beta N + 1$ (the case $K \leq \alpha N - 1$ is handled analogously). Then, the K_3 -density of G is

$$q_3 = (K/N)^3 + (1 - K/N)^3 \geq (\beta + 1/N)^3 + (1 - (\beta + 1/N))^3 \geq \beta^3 + (1 - \beta)^3 + \Omega(N^{-2}),$$

where the first inequality is due to monotonicity of the function $x^3 + (1-x)^3$ for $x \in [0, 0.5]$, and the second inequality follows from $\beta \leq \frac{K-1}{N} \leq 0.5 - N^{-1}$. Hence, the distance of $\bar{q} = (0, 1 - q_3, 0, q_3)$ from \bar{m} is $r + \Omega(N^{-2})$, where $\Omega(N^{-2})$ is the distance of \bar{q} to $\mathcal{S}_{(\beta, 1-\beta)}^3$ to \bar{m} and r is the distance of $\mathcal{S}_{(\beta, 1-\beta)}^3$ to \bar{m} (since the three points lie on the line $(0, 1-x, 0, x)$). Hence, $(q_1 - m_1)^2 + (q_3 - m_3)^2 = (r + \Omega(N^{-2}))^2 > r^2 + \Omega(r^2 N^{-2})$, which implies the required bound on \mathbb{P} because

$$\mathbb{P}(q_0, q_1, q_2, q_3) \leq \frac{1}{4N^6} - \frac{(q_1 - m_1)^2}{4r^2 N^6} - \frac{(q_3 - m_3)^2}{4r^2 N^6} < \frac{1}{4N^6} - \frac{r^2 + \Omega(r^2 N^{-2})}{4r^2 N^6} < -\Omega(N^{-8}).$$

We have shown that if G is ϵ -far from $\mathcal{CC}(D_{\alpha, \beta})$, then the tester accepts G with probability that is at most $0.5 - \Omega(\epsilon^{8/d})$, and thus the tester has detection probability $\varrho(\epsilon) = \Omega(\epsilon^{8/d})$, as required. Since $\deg(\mathbb{P}) = 2$, it follows that given a graph G the distribution tester needs 2 samples from \mathcal{S}_G^3 , thus making 6 queries to the graph G itself. This completes the proof of the proposition. \blacksquare

Remark regarding the bounds on $\alpha(N)$ and $\beta(N)$: Note that we can relax the restrictions on $\alpha(N)$ and $\beta(N)$ at the cost of detection probability of the tester. Specifically, let $\mu : \mathbb{R} \rightarrow [0, 1]$ be a monotone function such that $\mu(x) \rightarrow 0$ as x grows to infinity. If we assume in Proposition 6.10 that $\alpha(N) < \mu(N)$ and $\beta(N) > \frac{1}{2} - \mu(N)$ for all $N \in \mathbb{N}$, then the POT described in the proof has detection probability $\varrho(\epsilon) = \text{poly}\left(\frac{1}{\mu^{-1}(\epsilon)}\right)$, where μ^{-1} denotes the inverse function of μ .

The following corollary follows from the proof of Proposition 6.10. Note that this corollary uses only trivial conditions regarding α and β (i.e., $0 \leq \alpha(N) \leq \beta(N) \leq 0.5$), but yields no POT.

Corollary 6.11 (For the discussion in Section 3.5): *Let $\bar{\alpha}, \bar{\beta} : \mathbb{N} \rightarrow [0, \frac{1}{2}]^t$ such that for every $N \in \mathbb{N}$ it holds that $\alpha_1(N) \leq \beta_1(N) < \dots < \alpha_t(N) \leq \beta_t(N)$ for all $N \in \mathbb{N}$. Then, there is a constant $b \in (0, 1)$ and for every $N \in \mathbb{N}$ there are $\bar{w} = (w_H)_{H:|V(H)|=O(1)}$ such that $\mathcal{CC}_N(D_{\bar{\alpha}, \bar{\beta}})$ is exactly the set of all N -vertex graphs in $\Pi_{\bar{w}, b}$.*

Proof: By the proof above there is a universal constant c and a (weak) tester for $\mathcal{CC}(D_{\bar{\alpha}, \bar{\beta}})$ that accepts graph $G \in \mathcal{CC}_N(D_{\bar{\alpha}, \bar{\beta}})$ with probability at least c , and accepts graph $G \notin \mathcal{CC}_N(D_{\bar{\alpha}, \bar{\beta}})$ with probability smaller than c . The corollary follows by letting $b = 1 - c$ and using the proof of Theorem 3.10 to obtain a characterization of $\mathcal{CC}_N(D_{\bar{\alpha}, \bar{\beta}})$ in terms of $\Pi_{\bar{w}, b}$. \blacksquare

Generalization to $t \geq 2$: In order to generalize the above proof to larger values of t , we will need to define a collection of t segments $\{I_{\alpha_i, \beta_i} : i \in [t]\}$, and t ellipsoids $\{\mathcal{E}_{\alpha_i, \beta_i} : i \in [t]\}$, instead of a single segment (and ellipsoid) as done in the proof of Proposition 6.10. The desired POT is obtained by constructing a tester analogous to the one in Proposition 5.5 for disjoint union of ellipsoids.

6.4 Impossibility results for subclasses of $\mathcal{CC}^{\leq 2}$

In this section we prove that for any constants $0 \leq \alpha < \beta \leq \frac{1}{2}$ such that $\beta - \alpha < 0.5$, the class $\mathcal{CC}(D_{\alpha, \beta})$ does not have a two-sided error POT. This impossibility result complements Propositions 6.7 and 6.9 by explaining why the POTs provided by the latter results apply to $\alpha, \beta : \mathbb{N} \rightarrow [0, 0.5]$ such that $\lim_{N \rightarrow \infty} (\beta(N) - \alpha(N)) \in \{0, 0.5\}$.

The argument uses Theorem 3.10, which allows us to consider only potential testers that on input a graph G rule based on the distribution induced by $O(1)$ -vertex subgraphs of G . We show

that if such a potential tester \mathcal{T} provides a characterization of $\mathcal{CC}(D_{\alpha,\beta})$ with respect to some threshold c (i.e., \mathcal{T} accepts G with probability at least c if and only if $G \in \mathcal{CC}(D_{\alpha,\beta})$), then \mathcal{T} there exist infinitely many graphs G that are $\Omega(1)$ -far from $\mathcal{CC}(D_{\alpha,\beta})$ such that \mathcal{T} accepts G with probability $c - O(1/|V(G)|^2)$. It follows that \mathcal{T} cannot be a POT.

Proposition 6.12 (classes $\mathcal{CC}(D_{\alpha,\beta})$ that have no POT): *Let $0 \leq \alpha < \beta \leq \frac{1}{2}$ be constants such that either $\alpha > 0$ or $\beta < \frac{1}{2}$. Then, the class $\mathcal{CC}(D_{\alpha,\beta})$ does not have a two-sided error POT.*

Proof: We start with an overview of the proof, where we assume towards the contradiction that there is a constant query tester \mathcal{T} for $\mathcal{CC}(D_{\alpha,\beta})$ with threshold probability c . Then (similarly to the proof of Theorem 3.10), we may assume that, for some constant t , the tester \mathcal{T} reads a subgraph of G induced by random uniformly chosen t vertices, and accepts a view $H = ([h], E_H)$ with probability $w_H \in [0, 1]$. Hence the probability that \mathcal{T} accepts G can be written as $\sum_H w_H \cdot \rho_H(G)$, where the sum is over all t -vertex graphs, $\rho_H(G)$ denotes the density of H as an induced subgraph of G , and $w_H \in [0, 1]$ for each t -vertex graph H .

We first claim that all N -vertex graphs $G \in \mathcal{CC}(D_{\alpha,\beta})$ are accepted with probability that is at most $c + O(N^{-2})$. This follows from the fact that every graph $G \in \mathcal{CC}(D_{\alpha,\beta})$ is (N^{-2}) -close to a graph G' that is not in $\mathcal{CC}(D_{\alpha,\beta})$, since we can remove an edge from the larger clique of G to obtain such G' . Therefore, the distribution $(\rho_H(G'))_{H:|V(H)|=t}$ is $O(N^{-2})$ -close to the distribution $(\rho_H(G))_{H:|V(H)|=t}$, and so $\sum_H w_H \cdot \rho_H(G') < c$ (per $G' \notin \mathcal{CC}(D_{\alpha,\beta})$) implies $\sum_H w_H \cdot \rho_H(G) = c + O(N^{-2})$.

In the second step we shall claim (see proof outline below) that since all graphs $G \in \mathcal{CC}(D_{\alpha,\beta})$ are accepted with probability that deviates from c by at most $O(N^{-2})$, it must be the case that all graphs in $\mathcal{CC}^{\leq 2}$ are accepted by \mathcal{T} with probability that is $O(N^{-2})$ -close to c , and hence with probability at least $c - O(N^{-2})$. In particular, if $\beta < \frac{1}{2}$, then the graph consisting of two cliques each of density $\frac{1}{2}$ is $\Omega(1)$ -far from $\mathcal{CC}(D_{\alpha,\beta})$, yet it is accepted with probability at least $c - O(N^{-2})$. Similarly, if $\alpha > 0$, then the graph $G = K_n$ is $\Omega(1)$ -far from $\mathcal{CC}(D_{\alpha,\beta})$, yet it is accepted with probability at least $c - O(N^{-2})$. Therefore, the detection probability of \mathcal{T} , on some N -vertex graphs that are $\Omega(1)$ -far from the property, is at most $O(N^{-2})$, thus implying that \mathcal{T} is not a POT for $\mathcal{CC}(D_{\alpha,\beta})$.

The second step is proven by focusing on the behavior of \mathcal{T} on the various graphs in $\mathcal{CC}^{\leq 2}$, while noting that this behavior (or rather \mathcal{T} 's acceptance probability) depends only on the density of the smallest clique, denoted ρ , which in turn determines a unique N -vertex graph in $\mathcal{CC}^{\leq 2}$, denoted G_ρ . Recall that the probability that \mathcal{T} accepts the graph G_ρ is a linear combination (with coefficients in $[0, 1]$) of the corresponding densities $(\rho_H(G_\rho))_{H:|V(H)|=t}$. Moreover, for every t -vertex graph H , the density $\rho_H(G_\rho)$ is a polynomial (in ρ) of degree at most t .³⁹ Therefore, the probability that \mathcal{T} accepts G_ρ can be written as a polynomial (in ρ) of degree at most t . Let us denote this polynomial by $P : [0, \frac{1}{2}] \rightarrow \mathbb{R}$. Recall that (by the first step) P is almost constant on the interval $[\alpha, \beta]$. We claim that this implies that P is almost constant also on the entire interval $[0, \frac{1}{2}]$, where the ‘‘almost’’ in the conclusion depends on $\deg(P)$, on the ratio between the length of the interval $\beta - \alpha$, and on the ‘‘almost’’ parameter in the hypothesis. (This claim follows almost immediately from Claim 6.12.1, which in turn follows by polynomial interpolation.) Now, using the first step by which $P(\rho)$ is $O(1/N^2)$ -close to c for every $\rho \in [\alpha, \beta]$, we infer that $P(\rho)$ is $O(1/N^2)$ -close to c for every $\rho \in [0, \frac{1}{2}]$, where the constants in the $O()$ notations might differ. Since, for every $\rho \in [0, \frac{1}{2}]$,

³⁹For example, for $t \geq 3$ the density of K_t in G_ρ is $\rho_{K_t}(G_\rho) = \frac{\binom{\rho N}{t}}{\binom{N}{t}} + \frac{\binom{(1-\rho)N}{t}}{\binom{N}{t}}$, and the density of H that consists of K_{t-1} with additional isolated vertex is $\rho_H(G_\rho) = \frac{\binom{\rho N}{t-1} \cdot (1-\rho)^N}{\binom{N}{t}} + \frac{\binom{(1-\rho)N}{t-1} \cdot \rho^N}{\binom{N}{t}}$.

we have $\Pr[\mathcal{T} \text{ accepts } G_\rho] = \mathbb{P}(\rho)$ it follows that all graphs in $\mathcal{CC}^{\leq 2}$ are accepted with probability that is $O(1/N^2)$ -close to the threshold. As explained above, this implies that \mathcal{T} is not a POT for $\mathcal{CC}(D_{\alpha,\beta})$. The detailed argument is given next.

Assume towards contradiction that $\mathcal{CC}(D_{\alpha,\beta})$ has a POT \mathcal{T} with threshold probability c . Then, as explained in the proof overview, we may assume that for some constant t and for every $N \in \mathbb{N}$ there exists a sequence $\bar{w} = (w_H)_{H:|V(H)|=t}$ taking values in $[0, 1]$, such that the acceptance probability of \mathcal{T} when given a graph G on N vertices can be written as

$$\Pr[\mathcal{T} \text{ accepts } G] = \sum_{H:|V(H)|=t} w_H \cdot \rho_H(G), \quad (22)$$

where the sum is over all unlabeled t -vertex graphs H , and $\rho_H(G)$ denotes the density of H as a subgraph in G .

Note that for every graph $G \in \mathcal{CC}_N(D_{\alpha,\beta})$ we can drop an internal edge of the larger clique to obtain a graph G' that does not belong to $\mathcal{CC}_N(D_{\alpha,\beta})$. Hence, the graphs G and G' are (N^{-2}) -close. Therefore, by Claim 6.1, the densities $\rho_H(G)$ and $\rho_H(G')$ differ by at most $t^2 \cdot N^{-2}$ for every H . We conclude that \mathcal{T} accepts the graph G and G' with almost the same probability. That is:

$$\begin{aligned} |\Pr[\mathcal{T} \text{ accepts } G] - \Pr[\mathcal{T} \text{ accepts } G']| &= \left| \sum_H w_H \cdot \rho_H(G) - \sum_H w_H \cdot \rho_H(G') \right| \\ &\leq \sum_H w_H \cdot |\rho_H(G) - \rho_H(G')| \\ &\leq r t^2 \cdot N^{-2}, \end{aligned}$$

where $r < 2^{t^2}$ denotes the number of unlabeled t -vertex graphs.

Since any graph $G' \notin \mathcal{CC}_N(D_{\alpha,\beta})$ must be accepted with probability smaller than c , we conclude that any graph $G \in \mathcal{CC}_N(D_{\alpha,\beta})$ is accepted by \mathcal{T} with probability at most $c + O(N^{-2})$, where the constant in the $O()$ notation depends only on t . This implies the following inequality:

$$c \leq \Pr[\mathcal{T} \text{ accepts } G] \leq c + r t^2 \cdot N^{-2} \quad \text{for every } G \in \mathcal{CC}_N(D_{\alpha,\beta}). \quad (23)$$

In order to complete the proof we shall prove that for all N -vertex graphs $G \in \mathcal{CC}^{\leq 2}$ the acceptance probability of \mathcal{T} is $O(N^{-2})$ -close to the threshold c , where the constant in the $O()$ notation depends only on t and $\beta - \alpha$. As explained in the proof overview, since α, β are constants, there is a graph in $\mathcal{CC}^{\leq 2}$ that is $\Omega(1)$ -far from $\mathcal{CC}(D_{\alpha,\beta})$. Yet, since this graph is in $\mathcal{CC}^{\leq 2}$, it is accepted with probability at least $c - O(N^{-2})$, thus implying that \mathcal{T} is not a POT for $\mathcal{CC}(D_{\alpha,\beta})$.

In light of the above, we now focus on the behavior of \mathcal{T} only N -vertex input graphs that are in $\mathcal{CC}^{\leq 2}$. For every $\rho \in [0, \frac{1}{2}]$ such that $\rho N \in \mathbb{N}$, let G_ρ be the N -vertex graph in $\mathcal{CC}^{\leq 2}$ with cliques of densities ρ and $1 - \rho$. Then, as noted in the proof overview, for every t -vertex graph H the density of H in G_ρ is a polynomial (in ρ) of degree t , and thus, by Eq. (22), the probability that \mathcal{T} accepts the input graph G_ρ is also a polynomial of degree t . Consider a polynomial $\mathbb{P} : [0, \frac{1}{2}] \rightarrow \mathbb{R}$ defined as

$$\mathbb{P}(\rho) = \Pr[\mathcal{T} \text{ accepts } G_\rho] - c.$$

Recall that we have shown, in the first step of the proof, that $\Pr[\mathcal{T} \text{ accepts } G_\rho]$ is $(1/N^2)$ -close to c for every $\rho \in [\alpha, \beta]$ that satisfies $\rho N \in \mathbb{N}$. Specifically, by Eq. (23), the polynomial \mathbb{P} satisfies the following condition:

$$\mathbb{P}(\rho) \in [0, \frac{r t^2}{N^2}] \quad \text{for all } \rho \in [\alpha, \beta] \text{ that satisfy } \rho N \in \mathbb{N}. \quad (24)$$

It follows from the next claim that a polynomial $P(\rho)$ that satisfies Eq. (24) cannot deviate from zero by more than $O(N^{-2})$ also on a larger interval $[0, \frac{1}{2}]$.

Claim 6.12.1 *Let $P : [0, \frac{1}{2}] \rightarrow \mathbb{R}$ be a polynomial of degree at most t . Assume that for some $\epsilon, \delta > 0$ there are $t+1$ points $\rho_1, \dots, \rho_{t+1} \in [0, \frac{1}{2}]$ such that $|\rho_i - \rho_j| \geq \delta$ for all $i \neq j \in [t+1]$, and $|P(\rho_i)| \leq \epsilon$ for all $i \in [t+1]$. Then, for every $x \in [0, \frac{1}{2}]$, it holds that $|P(x)| < (t+1) \frac{1}{2^t \delta^t} \epsilon$.*

Before proving the claim, let us see how it allows us to complete the proof of Proposition 6.12. Let $\delta = \frac{\beta - \alpha}{t}$, and assume that N is large enough (e.g., $N > \frac{4}{\delta}$). Let us choose $t+1$ values $\alpha \leq \rho_1 < \rho_2 \dots < \rho_{t+1} \leq \beta$ that satisfy $\rho_i N \in \mathbb{N}$ for all $i \in [t+1]$ and $|\rho_{i+1} - \rho_i| \geq \delta - \frac{2}{N} > \frac{1}{2}\delta$ for all $i \in [d]$.⁴⁰ Then, by applying Claim 6.12.1 (and recalling that $P(\rho) \in [0, \frac{rt^2}{N^2}]$ for all $\rho \in [\alpha, \beta]$ that satisfy $\rho N \in \mathbb{N}$), we conclude that

$$|P(\rho)| \leq (t+1) \frac{1}{\delta^t} \cdot \frac{rt^2}{N^2} = (t+1) \left(\frac{t}{\beta - \alpha} \right)^t \cdot rt^2 \cdot N^{-2} \quad \text{for all } \rho \in [0, \frac{1}{2}]. \quad (25)$$

Therefore, the tester \mathcal{T} accepts all N -vertex graphs $G_\rho \in \mathcal{CC}^{\leq 2}$ with probability

$$\Pr[\mathcal{T} \text{ accepts } G_\rho] = c + P(\rho) > c - O(N^{-2}),$$

where the constant in the $O()$ notation depends only on t and $\beta - \alpha$. Since there are graphs in $\mathcal{CC}^{\leq 2}$ that are $\Omega(1)$ -far from $\mathcal{CC}(D_{\alpha, \beta})$, we conclude that \mathcal{T} is not a POT for $\mathcal{CC}(D_{\alpha, \beta})$. \blacksquare

Extension to smaller intervals (i.e., $\beta(N) = \alpha(N) + N^{-o(1)}$). The proof of Proposition 6.12 extends also to the case that α and β are functions that are relatively close. The point is that the only dependence on $\beta - \alpha$ occurs when we use the hypothesis that $(t+1) \left(\frac{t}{\beta - \alpha} \right)^t \cdot rt^2 = o(N^2)$, which implied that $\Omega(1)$ -far graphs are accepted with probability $c - o(1)$. Recalling that t and r are constants (which are determined by the query complexity of the potential tester), we infer that the argument holds as long as $\beta(N) = \alpha(N) + o(N^{2/t})$. Since we should fail each potential POT (i.e., each constant t), we can support any $\beta(N) = \alpha(N) + N^{-o(1)}$, which perfectly complements Proposition 6.8.

We return to the proof of Claim 6.12.1.

Proof of Claim 6.12.1 The proof uses interpolation of polynomials. Specifically, if we are given the values of P in $t+1$ points $\rho_1, \dots, \rho_{t+1} \in [0, \frac{1}{2}]$, then the polynomial P can be written as

$$P(x) = \sum_{i \in [t+1]} \left(\prod_{j \neq i} \frac{x - \rho_j}{\rho_i - \rho_j} \right) P(\rho_i).$$

Therefore, for every $x \in [0, \frac{1}{2}]$ we can upper bound $|P(x)|$ as follows:

$$|P(x)| \leq \sum_{i \in [t+1]} \left| \left(\prod_{j \neq i} \frac{x - \rho_j}{\rho_i - \rho_j} \right) P(\rho_i) \right|$$

⁴⁰This can be done by letting $\rho_i = \alpha + (i-1)\delta \in [\alpha, \beta]$ for all $i \in [t+1]$ (recall $\delta = \frac{\beta - \alpha}{t}$). Then $|\rho_i - \rho_j| \geq \delta$ for all $i \neq j \in [t+1]$. Note that ρ_i 's might not satisfy the condition $\rho_i N \in \mathbb{N}$. However, by modifying ρ_i by at most $\frac{1}{N}$ we can obtain $\rho_i \in [\alpha, \beta]$ for which $\rho_i N \in \mathbb{N}$ holds. Such modification changes the distance between ρ_i and ρ_j by at most $\frac{2}{N}$.

$$\begin{aligned}
[\text{using } |\rho_i - \rho_j| \geq \delta \text{ and } |\mathbb{P}(\rho_i)| \leq \epsilon] &\leq \sum_{i \in [t+1]} \left| \prod_{j \neq i} \frac{x - \rho_j}{\delta} \right| \cdot \epsilon \\
[\text{using } x, \rho_j \in [0, \frac{1}{2}]] &\leq \sum_{i \in [t+1]} \left(\prod_{j \neq i} \frac{1}{2\delta} \right) \cdot \epsilon \\
&= (t+1) \left(\frac{1}{2\delta} \right)^t \cdot \epsilon.
\end{aligned}$$

This completes the proof of the claim. ■

Acknowledgments

We are grateful to Dana Ron for collaboration in early stages of this research.

References

- [AFKS] N. Alon, E. Fischer, M. Krivelevich and M. Szegedy. Efficient Testing of Large Graphs. *Combinatorica*, Vol. 20, pages 451–476, 2000.
- [BCR] J. Bochnak, M. Coste, and M. Roy. *Real Algebraic Geometry*. Springer, 1998.
- [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. Extended abstract in *37th FOCS*, 1996.
- [GR97] O. Goldreich and D. Ron. Property Testing in Bounded Degree Graphs. *Algorithmica*, Vol. 32 (2), pages 302–343, 2002. Extended abstract in *29th STOC*, 1997.
- [GR09a] O. Goldreich and D. Ron. Algorithmic Aspects of Property Testing in the Dense Graphs Model. *SIAM Journal on Computing*, Vol. 40, No. 2, pages 376–445, 2011. Extended abstract in *13th RANDOM*, LNCS 5687, 2009.
- [GR09b] O. Goldreich and D. Ron. On Proximity Oblivious Testing. *SIAM Journal on Computing*, Vol. 40, No. 2, pages 534–566, 2011. Extended abstract in *41st STOC*, 2009.
- [GT03] O. Goldreich and L. Trevisan. Three theorems regarding testing graph properties. *Random Structures and Algorithms*, Vol. 23 (1), pages 23–57, August 2003.
- [GT05] O. Goldreich and L. Trevisan. Errata to [GT03]. Manuscript, August 2005. Available from http://www.wisdom.weizmann.ac.il/~oded/p_ttt.html
- [Ost] A. M. Ostrowski. *Solution of Equations and Systems of Equations*. Academic Press, 1960.
- [R1] D. Ron. Property testing: A learning theory perspective. *Foundations and Trends in Machine Learning*, 1(3), pages 307–402, 2008.
- [R2] D. Ron. Algorithmic and analysis techniques in property testing. *Foundations and Trends in TCS*, 5(2), pages 73–205, 2009.
- [RS] R. Rubinfeld and M. Sudan. Robust characterization of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2), pages 252–271, 1996.
- [Sol] P. Solernó. Effective Łojasiewicz Inequalities in Semialgebraic Geometry. *Applicable Algebra in Engineering, Communication and Computing*, Vol. 2 (1), pages 1–14, 1990.
- [WH] Y. Wu and C. Hadjicostis. On Solving Composite Power Polynomial Equations. *Math. Comput.*, 74(250):853–868, 2005.

Appendices: Proofs of Technical Claims

A.1 Proof of Claim 3.1.1

The following claim improves over a similar claim that appeared in [GGR, Apdx D].

Claim 3.1.1, restated: *If NK is even and $\sum_{v \in [N]} |d_G(v) - K| \leq \epsilon' \cdot N^2$, then G is $6\epsilon'$ -close to the set of K -regular N -vertex graphs.*

Proof: We modify G in three stages, while keeping track of the number of edge modifications. In the first stage we reduce all vertex degrees to at most K , by scanning all vertices and omitting $d_G(v) - K$ edges incident at each vertex $v \in H \stackrel{\text{def}}{=} \{u : d_G(u) > K\}$. Since $\sum_{v \in H} (d_G(v) - K) \leq \epsilon' N^2$, we obtain a graph G' that is $2\epsilon'$ -close to G such that $d_{G'}(v) \leq K$ holds for each vertex v , because every omitted edge reduces $\sum_{v \in H} (d_G(v) - K)$ by at least one unit. Furthermore, $\sum_{v \in [N]} |d_{G'}(v) - K| \leq \epsilon' \cdot N^2$, because each omitted edge $\{u, v\}$ reduces either $|d(u) - K|$ or $|d(v) - K|$ (while possibly increasing the other by one unit).

In the second stage, we insert an edge between every pair of vertices that are currently non-adjacent and have both degree smaller than K . Thus, we obtain a graph G'' that is ϵ' -close to G' such that $\{v : d_{G''}(v) < K\}$ is a clique (and $d_{G''}(v) \leq K$ for all v).

In the third stage, we iteratively increase the degrees of vertices that have degree less than K while preserving the degrees of all other vertices. Denoting by $\Gamma(v)$ the current set of neighbours of vertex v , we distinguish two cases.

Case 1: There exists a single vertex of degree less than K . Denoting this vertex by v , we note that $|\Gamma(v)| \leq K - 2$ must hold. We shall show that there exists two vertices u, w such that $\{u, w\}$ is an edge in the current graph but $u, w \notin \Gamma(v) \cup \{v\}$. Adding the edges $\{u, v\}$ and $\{w, v\}$ to the graph, while omitting the edge $\{u, w\}$, we increase $|\Gamma(v)|$ by two, while preserving the degrees of all other vertices.

We show the existence of two such vertices by starting with an arbitrary vertex $u \notin \Gamma(v) \cup \{v\}$. Vertex u has K neighbors (since $u \neq v$)⁴¹, and these neighbors cannot all be in $\Gamma(v) \cup \{v\}$ (which has size at most $K - 1$). Thus, there exists $w \in \Gamma(u) \setminus (\Gamma(v) \cup \{v\})$, and we are done.

Case 2: There exist at least two vertices of degree less than K . Let v_1 and v_2 be two vertices such that $|\Gamma(v_i)| \leq K - 1$ holds for both $i = 1, 2$. Note that $\{v_1, v_2\}$ is an edge in the current graph, since the set of vertices of degree less than K constitute a clique. We shall show that there exists two vertices u_1, u_2 such that $\{u_1, u_2\}$ is an edge in the current graph but neither $\{v_1, u_1\}$ nor $\{v_2, u_2\}$ are edges (and so $|\Gamma(u_1)| = |\Gamma(u_2)| = K$). Adding the edges $\{u_1, v_1\}$ and $\{u_2, v_2\}$ to the graph, while omitting the edge $\{u_1, u_2\}$, we increase $|\Gamma(v_i)|$ by one (for each $i = 1, 2$), while preserving the degrees of all other vertices.

We show the existence of two such vertices by starting with an arbitrary vertex $u_1 \notin \Gamma(v_1) \cup \{v_1, v_2\}$. Such a vertex exists since $v_2 \in \Gamma(v_1)$ and so $|\Gamma(v_1) \cup \{v_1, v_2\}| \leq K < N$. Vertex u_1 has K neighbors (since $u_1 \notin \Gamma(v_1)$, whereas all vertices of lower degree are neighbors of v_1). Note that $\Gamma(u_1)$ cannot be contained in $\Gamma(v_2) \cup \{v_2\}$, because $v_1 \notin \Gamma(u_1)$ whereas $v_1 \in \Gamma(v_2)$ (and $\Gamma(u_1) \subseteq \Gamma(v_2) \cup \{v_2\}$ would have implied $\Gamma(u_1) \subseteq \Gamma(v_2) \cup \{v_2\} \setminus \{v_1\}$, which is impossible since $|\Gamma(u_1)| = K$ whereas $|\Gamma(v_2) \cup \{v_2\} \setminus \{v_1\}| \leq K - 1$). Thus, there exists $u_2 \in \Gamma(u_1) \setminus (\Gamma(v_2) \cup \{v_2\})$.

⁴¹Recall that, by the case hypothesis, all vertices other than v have degree K .

Thus, in each step of the third stage, we decrease $\sum_{v \in [N]} |d_{G''}(v) - K|$ by two units, while preserving both the invariances established in the second stage (i.e., $\{v : d_{G''}(v) < K\}$ is a clique and $d_{G''}(v) \leq K$ for all v). Since in each step we modified three edges, we conclude that G'' is $3\epsilon'$ -close to $\mathcal{R}_N^{(k)}$, and the claim follows (by recalling that G is $3\epsilon'$ -close to G''). \blacksquare

A.2 Proofs of Propositions 3.4 and 3.5

Proposition 3.4, restated: *Let $\ell \geq 3$ be an integer and $t \in \mathbb{T}_\ell$, where*

$$\mathbb{T}_\ell = \begin{cases} \{1, 2, 3\} & \text{if } \ell \equiv 1 \pmod{2} \\ \{1\} & \text{if } \ell \equiv 2 \pmod{4} \\ \mathbb{N} & \text{if } \ell \equiv 0 \pmod{4} \end{cases} \quad (26)$$

and $k(N) = 2N/t\ell$. Then:

1. *The set $\mathcal{SC}_\ell \mathcal{C} \cap \mathcal{R}^{(k)}$ equals the set of graphs that consists of t super-cycle of length ℓ , each containing N/t vertices, such that clouds that are at distance four apart have equal size. Furthermore, if $\ell \not\equiv 0 \pmod{4}$, then each cloud has size $N/t\ell$.*
2. *If a graph $G = ([N], E)$ is δ -close to both $\mathcal{SC}_\ell \mathcal{C}$ and $\mathcal{R}^{(k)}$, then G is $O(\sqrt{\delta})$ -close to $\mathcal{SC}_\ell \mathcal{C} \cap \mathcal{R}^{(k)}$, where the hidden constant depends polynomially on $t\ell$.*

Thus, $\mathcal{SC}_\ell \mathcal{C} \cap \mathcal{R}^{(k)}$ has a two sided error POT.

Proof: Once Item 2 is proved, we use Theorem 3.2 to conclude that $\mathcal{SC}_\ell \mathcal{C} \cap \mathcal{R}^{(k)}$ has a two sided error POT. The proof of Item 2 is facilitated by Item 1, which anyhow serves as a good warm-up towards Item 2. Our exposition breaks down to several cases (i.e., $\ell = 4$, other $\ell \equiv 0 \pmod{4}$, and $\ell \not\equiv 0 \pmod{4}$), where in each case we first prove Item 1 and next prove Item 2.

We start with the case that $\ell = 4$. Recalling that a super-cycle of length four can be viewed as a bi-clique, we consider the class $\mathcal{BCC} \cap \mathcal{R}^{(k)}$, where \mathcal{BCC} denote the class of graphs that are each a collection of isolated bi-cliques. Suppose that $G \in \mathcal{BCC} \cap \mathcal{R}^{(k)}$, and consider the sequences of pairs $(S_0^{(1)}, S_1^{(1)})$ through $(S_0^{(m)}, S_1^{(m)})$ that are guaranteed by $G \in \mathcal{BCC}$. That is, vertices u and v are connected in G if and only if there exist $j \in [m]$ and $i \in \{0, 1\}$ such that $u \in S_i^{(j)}$ and $v \in S_{1-i}^{(j)}$. Using $G \in \mathcal{R}^{(k)}$, we infer that for every $j \in [m]$ and every $i \in \{0, 1\}$ it holds that $|S_i^{(j)}| = k(N) = N/2t$. (Indeed, if $N/2k(N)$ is not an integer, then $\mathcal{BCC} \cap \mathcal{R}^{(k)} = \emptyset$.) Thus, G may be viewed as a collection of t super-cycles of length four in which each cloud has size $N/4t$, and Item 1 follows.

Towards Item 2, suppose that G is δ -close to both \mathcal{BCC} and $\mathcal{R}^{(k)}$, and let $G' \in \mathcal{BCC}$ be δ -close to G . Then, G' is 2δ -close to $\mathcal{R}^{(k)}$, and thus $\sum_{v \in [N]} |d_{G'}(v) - (N/2t)| \leq 2\delta \cdot N^2$, where $d_{G'}(v)$ denotes the degree of v in G' . Fixing any integer q (e.g., $q = \Theta(1/\sqrt{\delta})$), we call a vertex v **good** if $|d_{G'}(v) - (N/2t)| \leq N/q$, and note that all but at most $2q\delta \cdot N$ vertices are good. Consider the sequences of pairs $(S_0^{(1)}, S_1^{(1)})$ through $(S_0^{(m)}, S_1^{(m)})$ that are guaranteed by $G' \in \mathcal{BCC}$. If vertex $v \in S_i^{(j)}$ is good, then $S_{1-i}^{(j)}$ must have size at least $(N/2t) - (N/q) > 2\delta N$, and so must also contain a good vertex. Hence, any biclique containing a good vertex has $(N/2t) \pm (N/q)$ vertices on each side (i.e., each cloud). Letting m' denote the number of bicliques that contain good vertices and assuming $q > 2(t+1)^2$ and $q < \sqrt{1/\delta}$, it follows that $m' = t$ (because $m' > t$ is ruled out by $2(t+1)((N/2t) - (N/q)) > N$, whereas $m' < t$ is ruled out by $2(t-1)((N/2t) + (N/q)) < N - 2q\delta N$). Moving all vertices to these $m' = t$ bicliques and modifying the edges accordingly, we obtain a graph G'' that is $2q\delta$ -close to G' . Furthermore, each biclique of G'' has $(N/2t) \pm (N/q)$ good vertices on

each side. Thus, by moving at most $2t \cdot (N/q) + 2q\delta N = 2(tN/q + q\delta)N$ vertices and modifying the edges accordingly, we obtain a graph in $\mathcal{BCC} \cap \mathcal{R}^{(k)}$. It follows that G is $(\delta + 4q\delta + 2t/q)$ -close to $\mathcal{BCC} \cap \mathcal{R}^{(k)}$. Using $q = 1/\sqrt{4\delta}$, the claim follows in the current case (i.e., for $\ell = 4$).

We now turn to the general case that $\ell \equiv 0 \pmod{4}$. Starting with Item 1, suppose that $G \in \mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$ and consider the sequences of clouds $(S_0^{(1)}, \dots, S_{\ell-1}^{(1)})$ through $(S_0^{(m)}, \dots, S_{\ell-1}^{(m)})$ that are guaranteed by $G \in \mathcal{SC}_\ell\mathcal{C}$. Using $G \in \mathcal{R}^{(k)}$, we infer that for every $j \in [m]$ and every $i \in \{0, 1, \dots, \ell - 1\}$ it holds that

$$|S_{i-1 \bmod \ell}^{(j)}| + |S_{i+1 \bmod \ell}^{(j)}| = k(N) \quad (27)$$

and so $|S_i^{(j)}| = |S_{i+4 \bmod \ell}^{(j)}|$ holds for every i, j . Combining the latter with Eq. (27), it follows that $\sum_i |S_i^{(j)}| = (\ell/2) \cdot k(N) = N/t$ for every j , which establishes Item 1. (Indeed, if $2N/t\ell$ is not an integer, then $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)} = \emptyset$.)

Turning to Item 2, suppose that G is δ -close to both $\mathcal{SC}_\ell\mathcal{C}$ and $\mathcal{R}^{(k)}$, and let $G' \in \mathcal{SC}_\ell\mathcal{C}$ be δ -close to G . Then, G' is 2δ -close to $\mathcal{R}^{(k)}$, and thus all but at most $2q\delta \cdot N$ vertices have degree $k(N) \pm (N/q)$. (We shall again use $q = \Theta(1/\sqrt{\delta})$.) We call these non-exceptional vertices *good*. Consider the sequences of clouds $(S_0^{(1)}, \dots, S_{\ell-1}^{(1)})$ through $(S_0^{(m)}, \dots, S_{\ell-1}^{(m)})$ that are guaranteed by $G' \in \mathcal{SC}_\ell\mathcal{C}$. We say that a cloud $S_i^{(j)}$ is *small* (resp., *big*) if $|S_i^{(j)}| < 3q\delta N$ (resp., if $|S_i^{(j)}| \geq 3q\delta N$), and note that big clouds contain good vertices. For each $j \in [m]$, we consider the following three cases.

Case 1: all clouds in the j^{th} super-cycle are small. Assuming that $k(N) - N/q > 6q\delta \cdot N$, we conclude that vertices on such super-cycles are not good, and thus their total number is at most $2q\delta N$.

Case 2: the j^{th} super-cycle contains four consecutive clouds that are big; that is, there exists an i_0 such that the clouds $S_{i_0}^{(j)}, S_{i_0+1 \bmod \ell}^{(j)}, S_{i_0+2 \bmod \ell}^{(j)}, S_{i_0+3 \bmod \ell}^{(j)}$ are all big. In this case, we can proceed analogously to the perfect case (where all degrees equal $k(N)$), and infer that $|S_i^{(j)}| + |S_{i+2 \bmod \ell}^{(j)}| = k(N) \pm N/q$ holds for every i , and the number of vertices residing on this super-cycle is $(\ell/2) \cdot (k(N) \pm (N/q)) = (N/t) \pm (\ell N/2q)$. Details follow.

The forgoing claim is established in $\ell - 3$ iterations, where in the $(i + 1)^{\text{st}}$ iteration we use hypotheses regarding the sets $S_{i_0+i \bmod \ell}^{(j)}, S_{i_0+i+1 \bmod \ell}^{(j)}, S_{i_0+i+2 \bmod \ell}^{(j)}, S_{i_0+i+3 \bmod \ell}^{(j)}$ and make an inference regarding the set $S_{i_0+i+4 \bmod \ell}^{(j)}$. Specifically, we assume that sets $S_{i_0+i+1 \bmod \ell}^{(j)}$ and $S_{i_0+i+3 \bmod \ell}^{(j)}$ contain good vertices and that $|S_{i_0+i}^{(j)}| \geq 3q\delta N - (i - 1)N/2q$, and infer that $|S_{i_0+i+4 \bmod \ell}^{(j)}| = |S_{i_0+i}^{(j)}| \pm 2N/q$ and thus $|S_{i_0+i+4 \bmod \ell}^{(j)}| \geq 3q\delta N - 2iN/q$. Using $i < \ell$ and $2\ell N/q < q\delta N$, it follows that $S_{i_0+i+4 \bmod \ell}^{(j)}$ contain good vertices. Having inferred that all clouds in the j^{th} super-cycle contain good vertices, we infer that $|S_i^{(j)}| + |S_{i+2 \bmod \ell}^{(j)}| = k(N) \pm N/q$ holds for every i .

Case 3: the j^{th} super-cycle has big clouds but no four consecutive clouds are big. We shall show that in this case the set of these super-cycles is close to one that satisfies the conclusion of Case 2.

Focusing on one such super-cycle (i.e., the j^{th} super-cycle), suppose that the i^{th} cloud is big and the $i - 1^{\text{st}}$ cloud is small. Then, the $i + 1^{\text{st}}$ cloud must be big (since $|S_{i-1 \bmod \ell}^{(j)}| + |S_{i+1 \bmod \ell}^{(j)}| \geq k(N) - N/q > 6q\delta N$), and so either the $i + 3^{\text{rd}}$ or the $i + 2^{\text{nd}}$ cloud must be small (because otherwise Case 2 holds). It follows that $|S_{i+1 \bmod \ell}^{(j)}| = k(N) \pm N/q \pm 3q\delta N$ and

$|S_i^{(j)}| + |S_{i+2 \bmod \ell}^{(j)}| = k(N) \pm N/q$, which means that these three sets form a bi-clique with approximately $k(N)$ vertices on each side and approximately $2k(n) = 4N/t\ell$ vertices in total.

Considering the approximate number of vertices counted in each of the three cases, we may ignore Case 1 and conclude that each super-cycle of Case 2 contributes (approximately) N/t vertices whereas the contribution of Case 3 comes in multiples of $4 \cdot (N/t\ell)$. Thus, if we have $t' \leq t$ super-cycles in Case 2, then we must have $(t - t') \cdot \ell/4$ bicliques in Case 3. But in such a case, we can rearrange these bicliques in $t - t'$ super-cycles of length ℓ such that each super-cycle contains $\ell/4$ bicliques and $\ell/2$ small clouds such that each two consecutive bicliques are connected via a super-path of two small clouds (i.e., the resulting super-cycle will consist of ℓ clouds such that the i^{th} cloud has size $k(N) \pm (N/q) \pm 3q\delta N$ if $\lfloor i/4 \rfloor \in \{0, 1\}$, and is small otherwise). Indeed, this may require changing the edges of all vertices that reside in small clouds, but the number of such vertices is less than $t\ell \cdot 3q\delta N$.

Thus, G' is $3t\ell q\delta$ -close to a graph that consists of t super-cycles in which each vertex has degree $k(N) \pm (N/q) \pm 3q\delta N$. It follows that G' is $t\ell \cdot (6q\delta + 1/q)$ -close to $\mathcal{R}^{(k)}$, and the claim follows (for any $\ell \equiv 0 \pmod{4}$).

Finally, we turn to the case that $\ell \not\equiv 0 \pmod{4}$. Using notations as in the case of $\ell \equiv 0 \pmod{4}$, we again have $|S_i^{(j)}| = |S_{i+4 \bmod \ell}^{(j)}|$ for every i, j . However, here (using $\ell \not\equiv 0 \pmod{4}$), we can infer that $|S_i^{(j)}| = |S_{i+2 \bmod \ell}^{(j)}|$, and (combining this with Eq. (27)) it follows that $|S_i^{(j)}| = k(N)/2 = N/t\ell$ holds for each i, j . This establishes Item 1. (Indeed, if $N/t\ell$ is not an integer, then $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)} = \emptyset$ follows.)

When proving Item 2, we again let G be δ -close to both $\mathcal{SC}_\ell\mathcal{C}$ and $\mathcal{R}^{(k)}$, and let $G' \in \mathcal{SC}_\ell\mathcal{C}$ be δ -close to G . Using the same notions of ‘good’ and ‘small’ as before, we again consider the same three cases regarding each super-cycle $(S_0^{(j)}, \dots, S_{\ell-1}^{(j)})$ of G' , where again $q = \Theta(1/\sqrt{\delta})$.

Case 1: all clouds in the j^{th} super-cycle are small. Again, we conclude that the total number of vertices on such super-cycles is at most $2q\delta N$.

Case 2: the j^{th} super-cycle contains four consecutive clouds that are big. Again, it follows that all clouds on this super-cycle contain good vertices, and $|S_i^{(j)}| + |S_{i+2 \bmod \ell}^{(j)}| = k(N) \pm N/q$ for every i . However, here we can also infer that each cloud has size $(N/t\ell) \pm (\ell N/q)$. This holds because, using $\ell \not\equiv 0 \pmod{4}$, it holds that $|S_i^{(j)}| = |S_{i+2 \bmod \ell}^{(j)}| \pm 2(\ell - 1)N/q$ for every i , and $|S_i^{(j)}| = (k(N) \pm (2\ell N/q))/2 = (N/t\ell) \pm (\ell N/q)$ follows. This means that, in this case, all clouds are big.

Case 3: the j^{th} super-cycle has both small and big clouds. Here we shall show that this case is actually impossible.

As in the case of $\ell \equiv 0 \pmod{4}$, we first observe that the approximate number of vertices that resides on super-cycles that satisfy Case 3 is a multiple of $4N/t\ell$. Thus, if we have $t' \geq 0$ super-cycles in Case 2, then $t' \cdot \ell \equiv t\ell \pmod{4}$, whereas $t' \leq t \in \mathbb{T}_\ell$. By considering both cases for $\ell \not\equiv 0 \pmod{4}$, we infer that $t' = t$ must hold⁴², which leaves no room for Case 3.

We conclude that almost all vertices reside in one of t super-cycles, which in turn contain clouds that are each of size $(N/t\ell) \pm (\ell N/q)$. Thus, G' is $t\ell^2 N/q$ -close to $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$, and the claim follows (when using $q = \Theta(1/\sqrt{\delta})$ again). ■

⁴²If $\ell \equiv 1 \pmod{2}$, then $t'\ell \equiv t\ell \pmod{4}$ implies $t' \equiv t \pmod{4}$, which implies $t' = t$ (since $t < 4$). If $\ell \equiv 2 \pmod{4}$, then $t'\ell \equiv t\ell \pmod{4}$ implies $t' \equiv t \pmod{2}$, which implies $t' = t$ (since $t = 1$).

Proposition 3.5, restated: *Let $\ell \geq 3$ and $t \in \mathbb{N} \setminus \mathbb{T}_\ell$, where \mathbb{T}_ℓ is as in Eq. (26). Then, for any integer $k(N) = 2N/t\ell$, there exists an N -vertex graph in $\mathcal{R}^{(k)}$ that is $O(1/N)$ -close to $\mathcal{SC}_\ell\mathcal{C}$ but $\Omega(1)$ -far from $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$.*

Proof: Fix any $\ell \not\equiv 0 \pmod{4}$ and $t \in \mathbb{N} \setminus \mathbb{T}_\ell$. Then, $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$ consists of super-cycles of length ℓ such that each cloud has size $N/t\ell$ (see the proof of Item 1 of Proposition 3.4, and note that it does not use the hypothesis $t \in \mathbb{T}_\ell$).⁴³ Indeed, if $N/t\ell$ is not an integer, then $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)} = \emptyset$ and we are done. Otherwise, let $i = 4$ if ℓ is odd and $i = 2$ otherwise, and note that $i \leq t$ (since $t \geq 4$ if ℓ is odd and $t \geq 2$ if $\ell \equiv 2 \pmod{4}$). Note that $i\ell \equiv 0 \pmod{4}$, and consider the N -vertex graph $G \in \mathcal{R}^{(k)}$ that consists of $t - i$ super-cycles of length ℓ with clouds of size $N/t\ell$ and $i\ell/4$ bicliques with $2N/t\ell$ vertices on each side. Observing that we can transform each such biclique into a super-cycle of length ℓ by moving $\ell - 2$ vertices into singleton clouds, it follows that G is $O(1/N)$ -close to $\mathcal{SC}_\ell\mathcal{C}$ (i.e., the distance of G to $\mathcal{SC}_\ell\mathcal{C}$ is smaller than $2i\ell^2 \cdot (2k(N)/N^2) \leq 32\ell/tN$). The claim follows by observing that G is $\Omega(1)$ -far from $\mathcal{SC}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$ (because the collection of $i\ell/4$ bi-cliques in G is far from a collection of i super-cycles of length ℓ with clouds of size $N/t\ell$). ■

A.3 Proof of Proposition 5.5

Proposition 5.5, restated: *Let Π_1, \dots, Π_k be disjoint classes of distributions with domain $[r]$. Assume that for each $i \in [k]$ the class Π_i has a two-sided error POT that makes t_i queries and has detection probability ϱ_i . Then, their union $\Pi = \cup_{i \in [k]} \Pi_i$ has a two-sided error POT that makes $\sum_{i \in [k]} t_i$ queries and has detection probability $\Omega(\min\{\varrho_i : i \in [k]\})$.*

Before proving the proposition, let us make a comment on the proof of Theorem 5.1. The proof gives us more information than what is stated in the theorem. Specifically, the proof gives us some relations between the class of distributions Π and the polynomial \mathbb{P} in the statement of the theorem. This extra information is summarized in Proposition A.14 below. The following definition will be convenient in the statement of the proposition.

Definition A.13 *Let $\mathbb{P}(q_1, \dots, q_r) : \Delta^{(r)} \rightarrow \mathbb{R}$ be a polynomial of total degree at most t . Write*

$$\mathbb{P}(q_1, \dots, q_r) = \sum_{m: \deg(m) \leq t} \alpha_m \cdot m(q_1, \dots, q_r),$$

where the sum is over monomials m of total degree at most t , and $\alpha_m \in \mathbb{R}$ for each monomial m . The polynomial \mathbb{P} is called γ -normalized if $|\alpha_m| \leq \gamma$ for each monomial m .

Note that, for every γ -normalized polynomial $\mathbb{P} : \Delta^{(r)} \rightarrow \mathbb{R}$ of degree t , there is a homogeneous $2^t\gamma$ -normalized polynomial \mathbb{P}^* of the same degree such that $\mathbb{P}(\bar{q}) \geq 0$ if and only if $\mathbb{P}^*(\bar{q}) \geq 0$ (for every \bar{q} in the domain). Specifically, suppose that $\mathbb{P}(q_1, \dots, q_r) = \sum_m \alpha_m \cdot m(q_1, \dots, q_r)$ is a non-zero polynomial of degree t . By multiplying each monomial m of degree $d < t$ by $(\sum_{i \in [r]} q_i)^{t-d}$, we obtain a homogeneous polynomial $\mathbb{P}^*(q_1, \dots, q_r) = \sum_{\bar{v}=(v_1, \dots, v_t) \in [r]^t} (\prod_{i=1}^t q_{v_i}) \cdot \beta_{\bar{v}}$ of degree t that agrees with \mathbb{P} on every point in $\Delta^{(r)}$. Note that for each $v \in [r]^t$ we have $\beta_{\bar{v}} = \sum_m \alpha_m$, where the sum is over all monomials m that divide $\prod_{i=1}^t q_{v_i}$. (Indeed, the above reproduces an argument that was outlined in the proof of Theorem 5.1.) We note that the proof of Theorem 5.1 actually establishes the following.

⁴³The hypothesis $t \in \mathbb{T}_\ell$ is only used when establishing Item 2 of Proposition 3.4 (for the case of $\ell \not\equiv 0 \pmod{4}$).

Proposition A.14 (Theorem 5.1, revised): *A property Π of distributions with domain $[r]$ has a two-sided error POT that makes t queries and has detection probability $\varrho : (0, 1] \rightarrow (0, 1]$ if and only if there exists a 1-normalized polynomial $P : \Delta^{(r)} \rightarrow \mathbb{R}$ of degree t satisfying the following conditions:*

1. $P(\bar{q}) \geq 0$ for every $\bar{q} \in \Pi$.
2. If $\bar{q} \in \Delta^{(r)}$ is ϵ -far from Π , then $P(\bar{q}) < -\varrho(\epsilon)$.

Moreover, if P is a γ -normalized polynomial of degree t , then $\Pi = \{\bar{q} \in \Delta^{(r)} : P(\bar{q}) \geq 0\}$ has a two-sided error POT \mathcal{T}_Π that makes t queries, whose acceptance probability when testing $\bar{q} \in \Delta^{(r)}$ can be written as

$$\Pr[\mathcal{T}_\Pi \text{ accepts } \bar{q}] = \frac{1}{2} + \frac{1}{2^{t+1}\gamma} P(\bar{q}).$$

We now turn to the proof of Proposition 5.5.

Proof: We give a proof for the special case of $k = 2$. The proof for larger values of k follows by induction on k .

Since each of the classes Π_i ($i = 1, 2$) has a two-sided error POT, by Proposition A.14 there are 1-normalized polynomials $P_i : \Delta^{(r)} \rightarrow \mathbb{R}$ of degree t_i , such that $P_i(\bar{q}) \geq 0$ for all $\bar{q} \in \Pi_i$, and $P_i(\bar{q}) < -\varrho_i(\epsilon)$ for any $\bar{q} \in \Delta^{(r)}$ that is ϵ -far from Π_i . Define a polynomial $P : \Delta^{(r)} \rightarrow \mathbb{R}$ of degree $t_1 + t_2$ to be

$$P(\bar{q}) = -\delta \cdot (P_1(\bar{q}) \cdot P_2(\bar{q})),$$

where $\delta > 0$ is some constant (e.g., $\delta = 2^{-2(t_1+t_2+1)}$), assuring that P is $2^{-t_1-t_2-1}$ -normalized. By applying Proposition A.14 with the polynomial P , we obtain a POT \mathcal{T}_Π for $\Pi = \Pi_1 \cup \Pi_2$ that makes $t_1 + t_2$ queries, as the total degree of P is at most $t_1 + t_2$.

We show below that the detection probability of \mathcal{T}_Π is $\Omega(\min(\varrho_1, \varrho_2))$. By Proposition A.14 it is enough to show that any $\bar{q} \in \Delta^{(r)}$ that is ϵ -far from Π satisfies the inequality $P(\bar{q}) < -\Omega(\min(\varrho_1(\epsilon), \varrho_2(\epsilon)))$.

By continuity of P_1 , since $\Pi_1 = \{\bar{q} : P_1(\bar{q}) \geq 0\}$ is the preimage of closed set $[0, \infty) \subseteq \mathbb{R}$ under P_1 , it follows that Π_1 is a closed subset of $\Delta^{(r)}$. Similarly Π_2 is also a closed subsets of $\Delta^{(r)}$. Therefore, since Π_1 and Π_2 are disjoint closed sets, there exists some $\gamma > 0$ such that $\text{dist}(q^{(1)}, q^{(2)}) > \gamma$ for all $q^{(1)} \in \Pi_1$ and $q^{(2)} \in \Pi_2$.

Fix $\epsilon < \gamma/2$ and let $\bar{q} = (q_1, \dots, q_r)$ be a distribution satisfying $\text{dist}(\bar{q}, \Pi) = \epsilon$. If $\text{dist}(\bar{q}, \Pi_1) = \epsilon$, then, by letting $\eta = \text{dist}(\bar{q}, \Pi_2)$, the triangle inequality gives us $\eta > \gamma/2$. Therefore, by using the bounds $P_1(\bar{q}) < -\varrho_1(\epsilon)$ and $P_2(\bar{q}) < -\varrho_2(\eta)$ we have

$$\begin{aligned} P(\bar{q}) &= -\delta \cdot (P_1(\bar{q}) \cdot P_2(\bar{q})) \\ &< -\delta \cdot \varrho_1(\epsilon) \cdot \varrho_2(\eta) \\ &\leq -\delta \cdot \varrho_1(\epsilon) \cdot \varrho_2(\gamma/2), \end{aligned}$$

where the second inequality follows from monotonicity of ϱ_2 . Similarly, if $\text{dist}(\bar{q}, \Pi_2) = \epsilon$, then $P(\bar{q}) < -\delta \cdot \varrho_2(\epsilon) \cdot \varrho_1(\gamma/2)$. In both cases we have $P(\bar{q}) < -\Omega(\min(\varrho_1(\epsilon), \varrho_2(\epsilon)))$, where the constant in the $\Omega()$ notation depends only on Π_1 and Π_2 (i.e., it is $\delta \cdot \min(\varrho_1(\gamma/2), \varrho_2(\gamma/2))$). This completes the proof of Proposition 5.5. ■

A.4 Strengthening Corollary 5.8

In this section we strengthen Corollary 5.8 on testing classes of distributions defined by an ellipsoid. Recall the definition of $\Pi(\bar{p}, \bar{B})$ in Eq. (14). For $\bar{p} = (p_1, \dots, p_r)$ and $\bar{B} = (B_0; B_1 \dots B_r) \in \mathbb{R}^{r+1}$, such that $B_0 \geq 0$ and $B_i > 0$ for all $i \in [r]$, let $\Pi(\bar{p}, \bar{B})$ be the ellipsoid in $\Delta^{(r)}$ defined as

$$\Pi(\bar{p}, \bar{B}) = \{\bar{q} = (q_1, \dots, q_r) : \sum_{i \in [r]} B_i (q_i - p_i)^2 \leq B_0\}.$$

Proposition A.15 *Fix $r \geq 2$, and let $\bar{p} = (p_1, \dots, p_r)$ and $\bar{B} = (B_0; B_1 \dots B_r) \in \mathbb{R}^{r+1}$, such that $B_i > 0$ for all $i = 0, 1, \dots, r$. Then, the property $\Pi(\bar{p}, \bar{B})$ has a two-sided error POT, that makes two queries and has linear detection probability.*

Proof: As in the proof of Corollary 5.8, define a polynomial P in r variables, that is non-negative for all points (q_1, \dots, q_r) in the ellipsoid, and negative outside the ellipsoid. Namely

$$P(q_1, \dots, q_r) = B_0 - \sum_{i \in [r]} B_i (q_i - p_i)^2.$$

Clearly $\Pi(\bar{p}, \bar{B}) = \{\bar{q} : P(\bar{q}) \geq 0\}$. The following claim completes the proof of the proposition.

Claim A.15.1 *For any $\bar{q} = (q_1, \dots, q_r)$ that is ϵ -far from $\Pi(\bar{p}, \bar{B})$, it holds that*

$$\sum_{i \in [r]} B_i (q_i - p_i)^2 > B_0 + \Omega(\epsilon), \quad (28)$$

where the constant in the $\Omega(\cdot)$ notation depends only on \bar{B} .

According to Claim A.15.1 for any $\bar{q} = (q_1, \dots, q_r) \in \Delta^{(r)}$ that is ϵ -far from $\Pi(\bar{p}, \bar{B})$ it holds that $P(q_1, \dots, q_r) < -\Omega(\epsilon)$. The proposition follows by normalizing P and applying the characterization given in Proposition A.14. ■

We return to the proof of Claim A.15.1.

Proof: Consider the distribution $\bar{q}' = (q'_1, \dots, q'_r)$, which is a convex combination of \bar{p} and \bar{q} , such that $\sum_{i \in [r]} B_i (q'_i - p_i)^2 = B_0$. The expression $\sum_{i \in [r]} B_i (q_i - p_i)^2$ can be bounded from below as follows:

$$\begin{aligned} \sum_{i \in [r]} B_i (q_i - p_i)^2 &= \sum_{i \in [r]} B_i ((q'_i - p_i) + (q_i - q'_i))^2 \\ &= \sum_{i \in [r]} B_i (q'_i - p_i)^2 + \sum_{i \in [r]} B_i (q_i - q'_i)^2 + 2 \sum_{i \in [r]} B_i (q'_i - p_i)(q_i - q'_i) \\ &\geq \sum_{i \in [r]} B_i (q'_i - p_i)^2 + 2 \sum_{i \in [r]} B_i (q'_i - p_i)(q_i - q'_i). \end{aligned}$$

Since \bar{q}' is a convex combination of \bar{p} and \bar{q} , for each $i \in [r]$ it holds that $q'_i \geq p_i$ if and only if $q'_i \leq q_i$, and in particular $(q'_i - p_i)(q_i - q'_i) \geq 0$. Therefore $\sum_{i \in [r]} B_i (q_i - p_i)^2$ can be bounded from below by

$$\sum_{i \in [r]} B_i (q_i - p_i)^2 \geq \sum_{i \in [r]} B_i (q'_i - p_i)^2 + 2 \min_{i \in [r]} \{B_i\} \cdot \left(\sum_{i \in [r]} (q'_i - p_i)(q_i - q'_i) \right). \quad (29)$$

By the choice of $\bar{q}' \in \Delta^{(r)}$ the first sum equals to B_0 . Therefore, in order to prove the claim it is enough to show that $\sum_{i \in [r]} (q'_i - p_i)(q_i - q'_i) > \Omega(\epsilon)$, where the constant in the $\Omega()$ notation depends only on \bar{B} .

Using again the fact that \bar{q}' is a convex combination of \bar{p} and \bar{q} , we conclude that the vectors $(\bar{q}' - \bar{p})$ and $(\bar{q} - \bar{q}')$ are co-linear, and thus

$$\sum_{i \in [r]} (q'_i - p_i)(q_i - q'_i) = \sqrt{\sum_{i \in [r]} (q'_i - p_i)^2} \cdot \sqrt{\sum_{i \in [r]} (q_i - q'_i)^2}. \quad (30)$$

We bound from below each term of the product separately. The first term is upper bounded as following.

$$\sqrt{\sum_{i \in [r]} (q'_i - p_i)^2} \geq \sqrt{\sum_{i \in [r]} \frac{B_i}{\max_{j \in [r]} \{B_j\}} \cdot (q'_i - p_i)^2} = \sqrt{\frac{B_0}{\max_{j \in [r]} \{B_j\}}}.$$

where the equality is by the choice of \bar{q}' that satisfies $\sum_{i \in [r]} B_i (q'_i - p_i)^2 = B_0$. The second term can be upper bounded by applying Cauchy-Schwarz inequality.

$$\sqrt{r} \sqrt{\sum_{i \in [r]} (q_i - q'_i)^2} \geq \sum_{i \in [r]} |q_i - q'_i| > 2\epsilon,$$

where the second inequality uses the assumption that $\bar{q}' \in \Pi^{(p,B)}$, and \bar{q} is ϵ -far from $\Pi^{(p,B)}$. This implies a lower bound on Eq. (30). Specifically, we have

$$\sum_{i \in [r]} (q'_i - p_i)(q_i - q'_i) \geq \frac{2}{\sqrt{r}} \cdot \sqrt{\frac{B_0}{\max_{j \in [r]} \{B_j\}}} \cdot \epsilon.$$

Then, by plugging it in Eq. (29) we complete the proof of the claim. \blacksquare

A.5 Testable classes of distributions are not closed under taking complements

Following the remark in the end of Section 5.2, we describe a class of ternary distributions $\Pi \subseteq \Delta^{(3)}$ that has a POT, while $\text{cl}(\Delta^{(3)} \setminus \Pi)$ does not have one⁴⁴. We start with the following claim.

Claim A.16 *Let $\mathcal{D} = \{(x, y) \in [0, 1]^2 : x + y \leq 1\}$ be a subset of \mathbb{R}^2 . For $\alpha \in (0, 1)$ let $A = \{(x, y) \in \mathcal{D} : \mathbb{P}(x, y) \geq 0\}$, where $\mathbb{P} : \mathcal{D} \rightarrow \mathbb{R}$ is the polynomial $\mathbb{P}(x, y) = y^2 - (x - \alpha) \cdot x^2$. Then, there is no real polynomial \mathbb{Q} such that $\text{cl}(\mathcal{D} \setminus A) = \{(x, y) \in \mathcal{D} : \mathbb{Q}(x, y) \geq 0\}$, where $\text{cl}(\mathcal{D} \setminus A)$ is the closure of the complement⁴⁵ of A in \mathcal{D} .*

Proof: Note first that A can be written as

$$A = \{(x, y) \in \mathcal{D} | x \leq \alpha\} \cup B,$$

where

$$B = \left\{ (x, y) \in \mathcal{D} \mid x \geq \alpha, y \geq x\sqrt{x - \alpha} \right\}.$$

⁴⁴By $\text{cl}(\Delta^{(3)} \setminus \Pi)$ we refer to the set of all $(q_1, q_2, q_3) \in \Delta^{(3)}$, such that for all $\epsilon > 0$ there is $(q'_1, q'_2, q'_3) \in \Delta^{(3)} \setminus \Pi$ that satisfies $\frac{1}{2}(|q_1 - q'_1| + |q_2 - q'_2| + |q_3 - q'_3|) < \epsilon$.

⁴⁵By $\text{cl}(\mathcal{D} \setminus A)$ we refer to the set of all $(x, y) \in \mathcal{D}$, such that for all $\epsilon > 0$ there is $(x', y') \in \mathcal{D} \setminus A$ that satisfies $\frac{1}{2}(|x - x'| + |y - y'|) < \epsilon$.

In particular the boundary of A is $\partial A = \left\{ \left(x, x\sqrt{(x-\alpha)} \right) \mid x \in [\alpha, 1] \right\}$, which can also be written as

$$\partial A = \{ (x^2, x^2 \cdot (x - \alpha)) \mid x \in [\sqrt{\alpha}, 1] \}.$$

Assume towards contradiction that there is a polynomial Q that satisfies the condition stated in the claim, namely $\text{cl}(\mathcal{D} \setminus A) = \{(x, y) \in \mathcal{D} : Q(x, y) \geq 0\}$. Then, in particular (1) Q must be zero on the boundary of A , and (2) for any $(x, y) \in A \setminus \partial A$, it must be the case that $Q(x, y) < 0$. We prove below that no polynomial satisfies these two conditions simultaneously. Specifically we show that any polynomial satisfying (1), must vanish at the point $(0, 0) \in A \setminus \partial A$, thus contradicting condition (2).

Let Q be a polynomial that vanishes on ∂A . Note that the polynomial P is irreducible⁴⁶, and the two polynomials P and Q agree on the curve $\partial A = \{(x^2, x^2 \cdot (x - \alpha)) \mid x \in [\sqrt{\alpha}, 1]\}$. Therefore, since the two polynomials have infinitely many common zeros, by Bezout's theorem, they have a common non-trivial factor, i.e. there is a non-constant polynomial R , such that $P = R \cdot P'$ and $Q = R \cdot Q'$, for some polynomials P' and Q' . However, since P is irreducible, we conclude that P' is some constant polynomial and $R = cP$ for some non-zero constant $c \in \mathbb{R}$, and thus Q can be written as $Q = c \cdot P \cdot Q'$. Therefore, since P vanishes at $(0, 0)$ it follows that Q also vanishes at $(0, 0)$. The claim follows. ■

Using Claim A.16 we exhibit a property of ternary distributions Π that has a POT, while $\text{cl}(\Delta^{(3)} \setminus \Pi)$ does not have one.

Proposition A.17 *Let $\alpha \in (0, 1)$ and let $P(x, y) = y^2 - (x - \alpha) \cdot x^2$ be as in Claim A.16. Define $\Pi \subseteq \Delta^{(3)}$ to be*

$$\Pi = \{(q_1, q_2, q_3) \in \Delta^{(3)} : P(q_1, q_2) \geq 0\}.$$

Then, Π has a two-sided error POT, while the property $\text{cl}(\Delta^{(3)} \setminus \Pi)$ does not have one.

Proof: Clearly, by Theorem 5.1, Π has a two-sided error POT. In order to prove that $\Pi' := \text{cl}(\Delta^{(3)} \setminus \Pi)$ does not have a two-sided error POT, it is enough to show that there is no polynomial $P' : \Delta^{(3)} \rightarrow \mathbb{R}$, that satisfies $\Pi' = \{(q_1, q_2, q_3) \in \Delta^{(3)} : P'(q_1, q_2, q_3) \geq 0\}$, which follows easily from Claim A.16. Details follow.

Assume towards contradiction that such polynomial P' exists. Define a real polynomial $Q : \mathcal{D} \rightarrow \mathbb{R}$ to be

$$Q(x, y) = P'(x, y, 1 - x - y),$$

where $\mathcal{D} = \{(x, y) \in [0, 1]^2 : x + y \leq 1\}$, as in Claim A.16. Note that $(x, y, 1 - x - y) \in \Delta^{(3)}$ for all $(x, y) \in \mathcal{D}$, and thus Q is well defined.

Let $A = \{(x, y) \in \mathcal{D} : P(x, y) \geq 0\}$. It is a routine exercise to show that $Q(x, y) \geq 0$ if and only if $(x, y) \in \text{cl}(\mathcal{D} \setminus A)$, by using the definition of Q , as well as the definitions of $\text{cl}(\Delta^{(3)} \setminus \Pi)$ and $\text{cl}(\mathcal{D} \setminus A)$. By Claim A.16, such polynomial does not exist, thus contradicting the assumption that Π' has a POT. ■

⁴⁶Namely, P cannot be written as a product of two polynomials of smaller degree. This can be checked by writing P either as $P(x, y) = (y^2 + ay + \sum_{i=0}^3 b_i x^i)(\sum_{i=0}^3 c_i x^i)$ or $P(x, y) = (y + \sum_{i=0}^3 d_i x^i)(y + \sum_{i=0}^3 e_i x^i)$, and verifying that P has no non-trivial factorizations.

A.6 Proof of Claim 6.1

Claim 6.1, restated: Let $H = ([n], F)$ be a fixed graph. For every two graphs $G = ([N], E)$ and $G' = ([N], E')$ such that $N \geq qn$, if G and G' are ϵ -close, then

$$|\rho_H(G) - \rho_H(G')| = \frac{|\text{ind}_H(G) - \text{ind}_H(G')|}{N \text{ chosen}} \leq \epsilon n^2,$$

where $\text{ind}_H(G) = \rho_H(G) \cdot \binom{N}{n}$ denotes the number of induced copies of H in G .

Proof: Assume that G and G' are ϵ -close. Then, there is a sequence of graphs on N vertices $(G_0 = G, G_1, \dots, G_t = G')$, where $t \leq \epsilon N^2$, such that G_i and G_{i+1} differ by exactly one edge, for every $i \in \{0, \dots, t-1\}$. Note that for every such i we have

$$|\text{ind}_H(G_i) - \text{ind}_H(G_{i+1})| \leq \binom{N-2}{n-2}$$

as one pair of vertices is contained in at most $\binom{N-2}{n-2}$ subgraphs on n vertices. Therefore, by the triangle inequality, we have

$$\frac{|\text{ind}_H(G) - \text{ind}_H(G')|}{\binom{N}{n}} \leq \sum_{i=1}^t \frac{|\text{ind}_H(G_i) - \text{ind}_H(G_{i+1})|}{\binom{N}{n}} \leq \epsilon N^2 \frac{\binom{N-2}{n-2}}{\binom{N}{n}} \leq \epsilon n^2.$$

The claim follows. \blacksquare

A.7 Proof of Lemma 6.3.2

Lemma 6.3.2, restated: Fix $t \geq 2$ and let $\bar{\rho} = (\rho_1, \dots, \rho_t)$. Let $G' = ([N], E') \in \mathcal{CC}^{\leq t}$, and assume that $\text{dist}(\mathcal{S}_{\bar{\rho}}^{t+1}, \mathcal{S}_{G'}^{t+1}) \leq \epsilon'$. Then, G' is $O(\sqrt[t]{\epsilon'})$ -close to $\mathcal{CC}(\bar{\rho})$, where the constant in the $O()$ notation depends only on t .

We first observe that for any $\bar{\rho} = (\rho_1, \dots, \rho_t)$ and $\bar{\rho}' = (\rho'_1, \dots, \rho'_t)$, the distance between (N -vertex) graphs in $\mathcal{CC}^{\leq t}$ that have the corresponding clique-densities is upper bounded by $\delta \stackrel{\text{def}}{=} \sum_{i \in [t]} |\rho_i - \rho'_i|$. This holds because it suffices to move δN vertices among the cliques of one graph in order to result in the other. This observation, which will be used several times in the section, is summarized in the following claim.

Claim A.18 Fix $t \geq 2$. For given $\bar{\rho} = (\rho_1, \dots, \rho_t)$ and $\bar{\rho}' = (\rho'_1, \dots, \rho'_t)$, let $G = ([N], E) \in \mathcal{CC}(\bar{\rho})$ and $G' = ([N], E') \in \mathcal{CC}(\bar{\rho}')$. If $\sum_{i \in [t]} |\rho_i - \rho'_i| < \epsilon$, then G is ϵ -close to G' .

A warm-up (the case $t = 2$). Before proving Lemma 6.3.2 for all $t \geq 2$, let us consider a special case of $t = 2$, i.e., the graph G' consists of two cliques of densities ρ' and $1 - \rho'$. Then

$$\begin{aligned} \alpha'_0 &:= \Pr[\mathcal{S}_G^3 = I_3] = 0 \\ \alpha'_1 &:= \Pr[\mathcal{S}_G^3 \text{ is a graph with exactly one edge}] = \rho'(1 - \rho') \\ \alpha'_2 &:= \Pr[\mathcal{S}_G^3 = P_3] = 0 \\ \alpha'_3 &:= \Pr[\mathcal{S}_G^3 = K_3] = (\rho')^3 + (1 - \rho')^3. \end{aligned}$$

By the hypothesis of Lemma 6.3.2 we have $|\rho'(1 - \rho') - \rho(1 - \rho)| < \epsilon'$, which implies either $|\rho - \rho'| < O(\sqrt{\epsilon'})$ or $|\rho - (1 - \rho')| < O(\sqrt{\epsilon'})$. By applying Claim A.18 we conclude the lemma for the special case of $t = 2$.

The general case (i.e., $t \geq 2$). We shall proceed in two steps corresponding to the two parts of the following lemma, while noting that the second part coincides with Lemma 6.3.2.

Lemma A.19 (Lemma 6.3.2, rephrased) *Fix $t \geq 2$ and let $G = ([N], E) \in \mathcal{CC}^{\leq t}$. Then*

1. *The distribution \mathcal{S}_G^{t+1} uniquely defines the graph G . That is, the distribution \mathcal{S}_G^{t+1} uniquely determines $\bar{\rho}$ such that $G \in \mathcal{CC}(\bar{\rho})$.*
2. *If $G' = ([N], E') \in \mathcal{CC}^{\leq t}$ and the statistical distance between \mathcal{S}_G^{t+1} and $\mathcal{S}_{G'}^{t+1}$ is at most ϵ' , then G' is $O(\sqrt[t]{\epsilon'})$ -close to G , where the constant in the $O()$ notation depends only on t .*

Proof: Since $G \in \mathcal{CC}^{\leq t}$, it holds that $G \in \mathcal{CC}(\bar{\rho})$ for some $\bar{\rho} = (\rho_1, \dots, \rho_t)$. In the first item we need to show that \mathcal{S}_G^{t+1} uniquely defines $\bar{\rho}$. In order to achieve this goal, we use the following two claims. (which will be proved later).

Claim A.19.1 *There are $s_1, s_2, \dots, s_t \in [0, 1]$ that depend only on \mathcal{S}_G^{t+1} such that*

$$\sum_{i \in [t]} (\rho_i)^k = s_k \text{ for } k = 1, \dots, t. \quad (31)$$

More specifically, each s_k can be expressed as

$$s_k = \sum_H \alpha_k(H) \mathcal{S}_G^{t+1}(H), \quad (32)$$

where the sum is over graphs H with $t + 1$ vertices, and $\alpha_k(H)$ depends only on k and H .

Claim A.19.2 *For any given $s_1, \dots, s_t \in \mathbb{R}$ the system of equation in Eq. (31) (in variables ρ_1, \dots, ρ_t) has a unique solution over the complex numbers, up to a permutation of the variables.*

Given the two claims, consider the unique solution of the system Eq. (31). Since the clique densities of G constitutes a solution, it follows that it is the unique solution, which proves the first item of the lemma.

For the second item, let $G' \in \mathcal{CC}^{\leq t}$. Then we can write $G' \in \mathcal{CC}(\bar{\rho}')$ for some $\bar{\rho}' = (\rho'_1, \dots, \rho'_t)$. Define s'_k analogously to Claim A.19.1, such that $\sum_{i \in [t]} (\rho'_i)^k = s'_k$ for $k = 1, \dots, t$. Applying expression Eq. (32) for every $k = 1, \dots, t$ we have

$$\begin{aligned} |s_k - s'_k| &= \left| \sum_H \alpha_k(H) \mathcal{S}_G^{t+1}(H) - \sum_H \alpha_k(H) \mathcal{S}_{G'}^{t+1}(H) \right| \\ &\leq \sum_H \alpha_k(H) \cdot |\mathcal{S}_G^{t+1}(H) - \mathcal{S}_{G'}^{t+1}(H)| \\ &= 2 \max_H \{\alpha_k(H)\} \cdot \text{dist}(\mathcal{S}_G^{t+1}(H), \mathcal{S}_{G'}^{t+1}(H)) \\ &= O(\epsilon'), \end{aligned}$$

where the constant in the $O()$ notation depends only on t . The following claim (to be proved later) allows us to complete the proof of the lemma. (Actually, given Claim A.19.1, the following claim is the actual core of the lemma.)

Claim A.19.3 Let $s_1, \dots, s_t, s'_1, \dots, s'_t \in [0, 1]$ such that, for every $k \in [t]$, we have $|s_k - s'_k| \leq O(\epsilon')$. Assume that $(\rho_1, \dots, \rho_t) \in [0, 1]^t$ satisfy

$$\sum_{i \in [t]} (\rho_i)^k = s_k \text{ for } k = 1, \dots, t, \quad (33)$$

and $(\rho'_1, \dots, \rho'_t) \in [0, 1]^t$ satisfy

$$\sum_{i \in [t]} (\rho'_i)^k = s'_k \text{ for } k = 1, \dots, t. \quad (34)$$

Then, there is a permutation π of the index set $\{1, \dots, t\}$, that satisfies $|\rho_i - \rho'_{\pi(i)}| = O(\sqrt[t]{\epsilon'})$ for every $i \in [t]$, where the constant in the $O()$ notation depends only on t .

By Claim A.19.3 above, we have $G \in \mathcal{CC}(\bar{\rho})$ and $G' \in \mathcal{CC}(\bar{\rho}')$ such that

$$\sum_{i \in [t]} |\rho_i - \rho'_i| = O(\sqrt[t]{\epsilon'}).$$

By applying Claim A.18, we conclude that G and G' are $O(\sqrt[t]{\epsilon'})$ -close, where the constant in the $O()$ notation depends only on t . This completes the proof of the lemma. \blacksquare

We now turn to the proofs of the subclaims stated during the proof of Lemma A.19. The proofs rely on basic results regarding symmetric polynomials, power sums, and continuity of the roots in algebraic equation.

Proof of Claim A.19.1 For the special case of $k = 1$ we have $s_1 = 1$, as $\sum_{i \in [t]} \rho_i = 1$. Therefore we let $k \in \{2, \dots, t\}$ and consider the following polynomial.

$$\left(\sum_{i \in [t]} (\rho_i)^k \right) \cdot \left(\sum_{j \in [t]} \rho_j \right)^{t+1-k}. \quad (35)$$

Since $\sum_{i \in [t]} \rho_i = 1$, the expression in Eq. (35) equals to $\sum_{i \in [t]} (\rho_i)^k$. Thus it is enough to show that Eq. (35) can be written as

$$\sum_H \alpha_k(H) \mathcal{S}_G^{t+1}(H)$$

for some $\alpha_k(H)$. Let $H = ([t+1], E_H) \in \mathcal{CC}^{\leq t}$ be a graph with at most t cliques of sizes (c_1, \dots, c_t) , where some of c_i 's might be zeros. Then

$$\mathcal{S}_G^{t+1}(H) = \frac{1}{K_H} \sum_{\sigma \in S_t} \prod_{i \in [t]} (\rho_{\sigma(i)})^{c_{\sigma(i)}}, \quad (36)$$

where S_t denotes the set of permutations on $[t]$, and $K_H \geq 1$ is some constant⁴⁷ that depends only on (c_1, \dots, c_t) .

It is a standard fact that the set $\{\mathcal{S}_G^{t+1}(H) : H = ([t], E)\}$, considered as polynomials in the variables ρ_1, \dots, ρ_t , forms a basis for the space of symmetric polynomial of degree $t+1$. Thus, since the expression Eq. (35) is a symmetric homogeneous polynomial of degree $t+1$, it can be

⁴⁷In fact, $K_H = k_1! \cdot k_2! \cdots k_t!$, where $k_i = |\{j : c_j = i\}|$ is the number of i -cliques in H .

written as a linear combination of polynomials in Eq. (36). Therefore for every k the power sum $\sum_{i \in [t]} (\rho_i)^k$ can be written as $\sum_H \alpha_k(H) \mathcal{S}_G^{t+1}(H)$ for some coefficients $\alpha_k(H)$. ■

Proof of Claim A.19.2 The system of equations Eq. (31), can be solved using Newton's identities, by reducing the problem to a univariate monic polynomial of degree t . More explicitly, define the matrix $M_k(\bar{s}) \in \mathbb{R}^{k \times k}$ as

$$M_k(\bar{s}) = \begin{bmatrix} s_1 & 1 & 0 & \dots & 0 & 0 \\ s_2 & s_1 & 2 & \dots & 0 & 0 \\ s_3 & s_2 & s_1 & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ s_{k-1} & s_{k-2} & s_{k-3} & \dots & s_1 & k-1 \\ s_k & s_{k-1} & s_{k-2} & \dots & s_2 & s_1 \end{bmatrix} \quad \text{for } k = 1 \dots t. \quad (37)$$

Define $\Lambda_k(\bar{s})$ to be

$$\Lambda_k = \Lambda_k(\bar{s}) = \frac{(-1)^k}{k!} \det[M_k(\bar{s})], \quad (38)$$

and consider the polynomial

$$P_{\bar{s}}(x) = x^t + \Lambda_1 x^{t-1} + \Lambda_2 x^{t-2} + \dots + \Lambda_t.$$

Then, the t roots of $P_{\bar{s}}$ are the unique, up to a permutation, solution of the system Eq. (31) (for details see e.g. [WH]). ■

Proof of Claim A.19.3 As explained in the proof of Claim A.19.2, the vector (ρ_1, \dots, ρ_t) consist exactly of all the roots of the monic polynomial

$$P_{\bar{s}}(x) = x^t + \Lambda_1 x^{t-1} + \Lambda_2 x^{t-2} + \dots + \Lambda_t, \quad (39)$$

where Λ_k 's are given in Eq. (38). That is $P_{\bar{s}}(x) = \prod_{i=1}^t (x - \rho_i)$. Similarly, the vector $(\rho'_1, \dots, \rho'_t)$ contains exactly all the roots of the polynomial

$$P_{\bar{s}'}(x) = x^t + \Lambda'_1 x^{t-1} + \Lambda'_2 x^{t-2} + \dots + \Lambda'_t \quad (40)$$

with $\Lambda'_k = \Lambda'_k(\bar{s}')$ defined analogously, i.e., $P_{\bar{s}'}(x) = \prod_{i=1}^t (x - \rho'_i)$.

Next, we claim that

$$|\Lambda_k - \Lambda'_k| = O(\epsilon') \quad \text{for } k = 1, \dots, t. \quad (41)$$

Indeed, by the assumption of the claim we have $|s_i - s'_i| < O(\epsilon')$ for every $i = 1, \dots, t$. Thus, considering Eq. (37) and Eq. (38), we have

$$\begin{aligned} |\Lambda_k - \Lambda'_k| &= \left| \frac{(-1)^k}{k!} \det[M_k(\bar{s})] - \frac{(-1)^k}{k!} \det[M_k(\bar{s}')] \right| \\ [\text{triangle inequality}] &\leq \frac{1}{k!} \sum_{\sigma \in S_k} \left| \prod_{i=1}^t (M(\bar{s}))_{i, \sigma(i)} - \prod_{j=1}^t (M(\bar{s}'))_{j, \sigma(j)} \right| \\ &\leq \max_{\sigma \in S_k} \left[\prod_{i=1}^t (M(\bar{s}))_{i, \sigma(i)} - \prod_{j=1}^t (M(\bar{s}'))_{j, \sigma(j)} \right] \\ &\leq O(\epsilon'), \end{aligned}$$

where the last inequality relies on the fact that all entries of the matrices $M(\bar{s})$ and $M(\bar{s}')$ are between 0 and $k-1$, together with the observation that for all $m_1, \dots, m_k \in [0, k-1]$ and $\epsilon_1, \dots, \epsilon_k$, such that $|\epsilon_i| < O(\epsilon')$ for all i , we have $|\prod_{i \in [k]} (m_i + \epsilon_i) - \prod_{i \in [k]} m_i| = O(\epsilon')$.

Having established Eq. (41), we have two monic polynomials $P_{\bar{s}}$ and $P_{\bar{s}'}$ in Eq. (39) and Eq. (40), whose coefficients differ by at most $O(\epsilon')$. Recall, the roots of $P_{\bar{s}}$ are (ρ_1, \dots, ρ_t) , and the roots of $P_{\bar{s}'}$ are $(\rho'_1, \dots, \rho'_t)$. The claim follows from continuity of the roots of monic polynomials. Specifically, we quote a theorem due to Ostrowski [Ost, Appendix A].

Theorem A.20 *Let f and g be two monic polynomials $f(z) = \sum_{i=1}^t a_i z^i$ and $g(z) = \sum_{i=1}^t b_i z^i$ such that $a_t = b_t = 1$. Let x_1, \dots, x_t be the t roots of $f(z)$ (with multiplicities), and let y_1, \dots, y_t be the roots of $g(z)$. For*

$$\gamma = \max_{i \in [t]} \{|x_i|, |y_i|\} \quad (42)$$

introduce the expression

$$\epsilon = \sqrt[t]{\sum_{i=1}^t |b_i - a_i| \gamma^i}. \quad (43)$$

Then, the roots x_i and y_i can be ordered in such a way that

$$|x_i - y_i| < 2t\epsilon \quad \text{for } i = 1, \dots, t. \quad (44)$$

We apply Theorem A.20 to the polynomials $P_{\bar{s}}$ and $P_{\bar{s}'}$. By Eq. (41), the coefficients of the polynomials differ by at most $O(\epsilon')$. Recall that the roots of the polynomials are the densities of the cliques in G and G' , and hence they all lie in the interval $[0, 1]$. Therefore, γ in Eq. (42) is bounded by 1, and hence ϵ in Eq. (43) is bounded by $O(\sqrt[t]{\epsilon'})$. Hence, by Theorem A.20 there is an ordering of the roots such that $|\rho_i - \rho'_i| < O(\sqrt[t]{\epsilon'})$. This completes the proof of Claim A.19.3. ■