Two-Sided Error Proximity Oblivious Testing*

Oded Goldreich† Igor Shinkar‡

June 9, 2014

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 $F : (0,1] \to (0,1]$ 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.

*A preliminary version of this work has appeared as TR12-021 of *ECCC*; an extended abstract has appeared in the proceedings of *16th RANDOM*, Springer LNCS, Vol. 7408, pages 565–578, 2012.

†Department of Computer Science, Weizmann Institute of Science, Rehovot, Israel. oded.goldreich@weizmann.ac.il Partially supported by the Israel Science Foundation (grants No. 1041/08 and 671/13).

‡Department of Computer Science, Weizmann Institute of Science, Rehovot, Israel. igor.shinkar@weizmann.ac.il Research supported by ERC grant number 239985.

Contents

1 Introduction

In the last two decades, the area of property testing has attracted much attention (see, e.g., a couple of recent surveys [R1, R2]). Loosely speaking, property testing typically refers to sub-linear time probabilistic algorithms for deciding whether a given object has a predetermined property or is far from any object having this property. Such algorithms, called testers, obtain 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].

1.1 The notion of a Proximity Oblivious Tester (POT)

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 done below.

In the following definition, POTs are defined 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 the length of the input. The tester, denoted T, is a probabilistic strategy that gets an explicit input N that specifies the domain of the tested object (viewed as a function) as well as oracle access to the object itself; by $T^f(N)$ we denote a random variable that represents the output of T on explicit input N and oracle access to $f : \{1, ..., N\} \to \{0, 1\}^*$. We do n ot assume that this strategy is computable, let alone efficiently computable.[2] In the second item (of Definition 1.1), $\delta(f, g)$ denotes the relative distance between the functions f and g, the constant $c \in (0, 1]$ is a lower bound on the probability that objects

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

[2] Indeed, an equivalent formulation may introduce a different tester T_N for each value of $N \in \mathbb{N}$. Also note that, although the tester's query complexity is independent of N (and of ϵ), the tester must "know" N in order to make adequate queries to its oracle.

having the property are accepted, and the function $\varrho : (0,1] \to (0,1]$ represents the gap between the acceptance probability of objects having the property and objects far from it. Specifically, while objects having the property are accepted with probability at least c, objects that are ϵ-far from the property are accepted with probability at most $c - \varrho(\epsilon)$.

Definition 1.1 (POT, generalized): *Let $\Pi = \bigcup_{N\in\mathbb{N}} \Pi_N$, where Π_N contains functions defined over the domain $[N] \stackrel{\text{def}}{=} \{1, ..., N\}$, and let $\varrho : (0,1] \to (0,1]$ be monotone. A* two-sided error POT with detection probability ϱ for Π *is a probabilistic strategy 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 $f : [N] \to \{0,1\}^*$ 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.

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 $O(1/\varrho(\epsilon))$ 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.

1.2 On the power of two-sided error POTs

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.

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.

2. Regarding graph properties in the adjacency matrix model, we consider regular graphs, regular graphs of a prescribed degree, and subsets of such regular graphs (e.g., regular graphs that consists of a collection of bicliques). Another class of properties 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).

3. Regarding 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).

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. A more detailed account of these and other results is provided next.

1.3 An overview of our results

In this section and throughout the rest of this paper, unless stated differently, a POT means one with two-sided error probability.

We first consider POTs for symmetric properties of Boolean functions, where a property $\Pi = \cup_{N\in\mathbb{N}}\Pi_N$ is symmetric if for every $f \in \Pi_N$ and every permutation $\pi : [N] \to [N]$ it holds that $f \circ \pi \in \Pi_N$ (where $(f \circ \pi)(x) \stackrel{\text{def}}{=} f(\pi(x))$). Each symmetric property of Boolean functions, $\Pi = \cup_{N\in\mathbb{N}}\Pi_N$, is characterized by a sequence of sets $(S_N)_{N\in\mathbb{N}}$ such that for every $f : [N] \to \{0, 1\}$ it holds that $f \in \Pi_N$ if and only if $|\{x \in [N] : f(x) = 1\}| \in S_N$. We say that a set of natural numbers is t-consecutive if it can be partitioned into at most t sequences of consecutive numbers (e.g., $\{1, 2, 3, 5, 7, 8\}$ is 3-consecutive but not 2-consecutive).

Theorem 1.2 (POTs for symmetric properties of Boolean functions): *Let $\Pi = \cup_{N\in\mathbb{N}}\Pi_N$ be a symmetric property that is chracterized by the sequence of sets $(S_N)_{N\in\mathbb{N}}$. Then, Π has a POT if and only if there exists a constant t such that each S_N is t-consecutive.*

Theorem 1.2 is proved by relating uniform symmetric properties of Boolean functions to properties of distributions that assume values in $\{0, 1\}$, whereas chracterization of the binary distributions that have a POT is provided in Theorem 2.5. Jumping ahead, we mention that this relation generalized to the relation between functions with range Σ and distributions that assume values in Σ.

We next turn to testing graph properties in the adjacency matrix model (as defined in [GGR]). Here we present POTs for several properties that refer to regular graphs including all regular graphs, regular graphs of a prescribed degree, and some subsets of the latter.

[3]Consider, 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).

Theorem 1.3 (POTs for certain sets of regular graphs, in the adjacency matrix model): *The following graph properties have a POT.*

1. *The set of all regular graphs.*
2. *The set of all $\kappa \cdot N$-regular N-vertex graphs, for any constant κ.*
3. *The set of all regular complete t-partite graphs, for any constant $t \geq 2$.*

Item 1 of Theorem 1.3 appears as Theorem 3.5, Item 2 appears as Theorem 3.1, and Item 3 is derived by combining Theorem 3.2 (which states a general condition) with Proposition 3.3 (which shows that the conditions of Theorem 3.2 hold for complete t-partite graphs).

An altogether different class of properties that have POTs is the class of properties that upper bounded the density of the occurrences of some fixed graph as an induced subgraph. Specifically, for any fixed graph H and a generic graph G, let $\rho_H(G)$ denote the density of H as an induced subgraph of G. Let $\Pi_{H,\tau}$ denote the set of graphs G that satisfy $\rho_H(G) \leq \tau$. Recall that Alon *et al.* [AFKS] showed that, for every fixed H, the set $\Pi_{H,0}$ has a one-sided error POT, albeit their lower bound on the detection probability of this POT is very weak (i.e., a graph that is δ-far from $\Pi_{H,\tau}$ is rejected with probability $1/\mathtt{T}(\text{poly}(1/\delta))$, where $\mathtt{T}(m)$ is a tower of m exponents). Here we provide a much sharper bound for the case of $\tau > 0$ (while using an elementary proof and a two-sided error POT, which is necessary in this case).

Theorem 1.4 (a POT for $\Pi_{H,\tau}$, still in the adjacency matrix model): *For every fixed graph H with n vertices and for any constant $\tau > 0$, the property $\Pi_{H,\tau}$ has a POT. Furthermore, this POT accepts each graph in $\Pi_{H,\tau}$ with probability at least $1 - \tau$ and accepts graphs that are δ-far from $\Pi_{H,\tau}$ with probability at most $1 - \tau - (\tau n/3) \cdot \delta$.*[4]

Theorem 1.4 follows from Theorem 3.9, which relates the distance of a graph G from $\Pi_{H,\tau}$ to the density of H as an induced subgraph in G.

We also consider testing graph properties in the bounded-degree graph model (as defined in [GR97]). In this case, our results are obtained by simple reductions to the problem of testing binary distributions. Loosely speaking, the main result in this model is a POT for properties that refer to the number of isolated subgraphs that equal one of the graphs in some fixed family of graphs. For details, see Section 4 (and Theorem 4.3).

Theorems 1.3, 1.4 and 4.3 refer to the density of the occurrence of some specific patterns in the tested graph (e.g., Theorem 1.3 refers to the density of edges incident at various vertices, and Theorem 1.4 refers to the density of occurrence of a fixed graph as an induced subgraph). In each case, we consider the density of a single pattern, and so the tested density corresponds to a binary distribution. But when one wishes to refer to a number of densities that correspond to the occurrences of different patterns, multi-valued distributions arise. Indeed, a property may be defined by an arbitrary condition of the form "pattern A occurs between 10%-20% of the time, whereas pattern B occurs at least twice as often as pattern C." This motivates the study of POTs for properties of distributions over an arbitrary fixed-size domain (rather than over a binary domain), which is initiated in Section 5.

It turns out that POTs for properties of multi-valued distributions are more exceptional than their binary-valued analogues. As hinted above, Theorem 2.5 asserts that properties (of binary distributions) that correspond to intervals (representing the probability that the outcome is 1) have POTs. It is tempting to hope that properties of ternary distributions that correspond to

[4] This bound assumes that the tested graph has more than $6/\delta$ vertices.

rectangles (representing 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. In general, the question of whether or not a property of r-valued distributions has a POT is closely related to the question of whether there exists a polynomial that is non-negative on the distributions having the property and negative otherwise (where r-valued distributions are viewed as sets in $\mathbb{R}^r$).

Theorem 1.5 (POTs for testing multi-valued distributions, a coarse version of Theorem 5.1): *Let* Π *be an arbitrary set of distributions over* $[r]$*, viewed as the set of all non-negative* r*-sequences that sum up to 1. Then,* Π *has a POT if and only if there is a polynomial* $\mathtt{P} : \mathbb{R}^r \to \mathbb{R}$ *such that for every distribution* $\overline{q} = (q_1, \ldots, q_r)$ *it holds that* $\mathtt{P}(q_1, \ldots, q_r) \geq 0$ *if and only if* $\overline{q} \in \Pi$*. Furthermore, if the total degree of* $\mathtt{P}$ *is* t*, then* Π *has a two-sided error POT that makes* t *queries and has polynomial detection probability, where the power of the polynomial depends on* $\mathtt{P}$*.*

In particular, properties that correspond to sets of sequences that are within a given (positive) distance from a given sequence have a POT if distance is measured according to an L_p-norm for any integer $p \geq 2$, but have no POT if distance is measured according to the L_1-norm or the max-norm (see Corollary 5.3).

Conclusion. 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. In particular, Theorem 2.5 provides a characterization of binary distributions having a POT, and Theorem 5.1 provides a less effective characterization w.r.t multi-valued distributions, whereas Theorem 3.12 may be viewed as a programmatic step in the context of graph properties. Indeed, the current work is merely a first exploration of the notion of two-sided error POTs.

1.4 Organization

The rest of this paper is organized in four sections: Section 2 deals with testing properties of Boolean functions, while focusing on the study of POTs for symmetric properties, which is closely related to the study of POTs for properties of binary distributions. Section 3 deals with testing graph properties in the adjacency matrix model (of [GGR]). Section 4 deals with testing graph properties in the bounded-degree model (of [GR97]). Section 5 revisits the study of POTs for distributions, extending it to multi-valued distributions.

We note that Section 3 contains explicit statements of results and open problems that were only implicit in our technical report [GS12]. On the other hand, Section 5 contains only some of our results regarding the study of POTs for multi-valued distributions and its applications; more can be found in our technical report [GS12].

2 Classes of Boolean Functions

As mentioned in the Introduction, 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 $\tau - \epsilon$. (This is the case since a function that has a q fraction of

1-values is accepted with probability q, whereas a function that is ϵ-far from the aforementioned set must have $q < \tau - \epsilon$.)

A more telling example refers to the set of all Boolean functions 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 accepts with probability α_i if the sum of the answers equals i, where α_0, α_1 and α_2 are selected in a suitable manner.[5] To analyze this tester, consider a function with a q fraction of 1-values, and note that the probability that this function is accepted equals

$$P(q) \stackrel{\text{def}}{=} \alpha_2 \cdot q^2 + \alpha_1 \cdot 2q(1-q) + \alpha_0 \cdot (1-q)^2 .$$

Pick $\alpha_0, \alpha_1, \alpha_2 \in [0,1]$ such that $P(q)$ is quadratic form maximized at $(\tau_1 + \tau_2)/2$. This implies that $P(\tau_1) = P(\tau_2)$ and $P(q) < P(\tau_1)$ if and only if $q \notin [\tau_1, \tau_2]$. Furthermore, if q is ϵ-far from being in the interval (i.e., either $q < \tau_1 - \epsilon$ or $q > \tau_2 + \epsilon$), then $P(q) = P(\tau_1) - \Omega(\epsilon)$.

In general, we consider properties that are each specified by a sequence of t density thresholds, denoted $\overline{\tau} = (\tau_1, ..., \tau_t)$, such that $0 < \tau_1 < \tau_2 < \cdots < \tau_t < 1$. The corresponding property, denoted $\mathcal{B}_{\overline{\tau}}$, consists of all Boolean functions $f : [N] \to \{0,1\}$ 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).[6] Specifically, the corresponding set of distributions, denoted $\mathcal{D}_{\overline{\tau}}$, 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 $\mathcal{B}_{\overline{\tau}}$) correspond to samples obtained from the tested distribution. (Here, and throughout the paper we identify random variables with the distributions of their values.)

2.1 A generic tester and its analysis

Given a sequence of t density thresholds $\overline{\tau} = (\tau_1, ..., \tau_t)$, a generic tester for $\mathcal{D}_{\overline{\tau}}$ 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 $\overline{\alpha} = (\alpha_0, \alpha_1, ..., \alpha_k)$. 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 $\overline{\alpha}$ such that the resulting tester, denoted $T_{\overline{\alpha}}$, is a POT for $\mathcal{D}_{\overline{\tau}}$. (We shall show that $k = t$ is sufficient and necessary.) The key quantity to analyze is the probability that this tester (i.e., $T_{\overline{\alpha}}$) accepts a distribution that is 1 with probability q. This accepting probability, denoted $\mathsf{P}_{\overline{\alpha}}(q)$, satisfies

$$\mathsf{P}_{\overline{\alpha}}(q) = \sum_{i=0}^{k} \binom{k}{i} \cdot q^i (1-q)^{k-i} \cdot \alpha_i. \tag{1}$$

Indeed, the function $\mathsf{P}_{\overline{\alpha}}$ is a degree k polynomial. Since 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 $\mathcal{D}_{\overline{\tau}}$ and mean that the distribution that is 1 with

[5] For example, one may select $(\alpha_0, \alpha_1, \alpha_2) = (0, 1, \frac{2(\tau_1+\tau_2-1)}{\tau_1+\tau_2})$ 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.

[6] In this case, the distance between distributions is merely the standard notion of statistical distance.

probability q is in $\mathcal{D}_{\overline{\tau}}$). Thus, $T_{\overline{\alpha}}$ is a POT for $\mathcal{D}_{\overline{\tau}}$ if every distribution that is ϵ-far from $\mathcal{D}_{\overline{\tau}}$ is accepted with probability at most $c - \varrho(\epsilon)$, where $c \stackrel{\rm def}{=} \min_{q \in \mathcal{D}_{\overline{\tau}}}\{\mathtt{P}_{\overline{\alpha}}(q)\}$ and $\varrho : (0,1] \to (0,1]$ is some monotone function.

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

Proposition 2.1 (on the existence of $\overline{\alpha}$ such that $\mathtt{P}_{\overline{\alpha}}$ is "good"): *For every sequence $\overline{\tau} = (\tau_1, ..., \tau_t)$ such that $0 < \tau_1 < \tau_2 < \cdots < \tau_t < 1$, there exists a sequence $\overline{\alpha} = (\alpha_0, \alpha_1, ..., \alpha_t) \in [0,1]^{t+1}$ such that the following two conditions hold*

1. *For every $i \in [t]$, it holds that $\mathtt{P}_{\overline{\alpha}}(\tau_i) = \mathtt{P}_{\overline{\alpha}}(\tau_1)$.*
2. *The function $\mathtt{P}_{\overline{\alpha}}$ 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 $\alpha_0, \alpha_1, ..., \alpha_t$. This system is not contradictory, since the uniform vector, denoted $\overline{u}$, is a solution (i.e., $\overline{\alpha} = ((t+1)^{-1}, ..., (t+1)^{-1})$ satisfies $\mathtt{P}_{\overline{\alpha}}(\tau_i) = (t+1)^{-1}$). Thus, this ($t-1$ dimensional) system has also a solution that is linearly independent of $\overline{u}$. Denoting such a solution by $\overline{s}$, consider arbitrary $\beta \neq 0$ and γ such that $\beta\overline{s} + \gamma\overline{u} \in [0,1]^{t+1} \setminus \{0^{t+1}\}$. Note that $\overline{\alpha} \stackrel{\rm def}{=} \beta\overline{s} + \gamma\overline{u}$ satisfies the linear system and is not spanned by $\overline{u}$. To establish Condition 2, we show that only vectors $\overline{\alpha}$ that are spanned by $\overline{u}$ yield a constant function $\mathtt{P}_{\overline{\alpha}}$. To see this fact, write $\mathtt{P}_{\overline{\alpha}}(q)$ as a polynomial in q, obtaining:

$$\mathtt{P}_{\overline{\alpha}}(q) = \sum_{d=0}^{t} (-1)^d \binom{t}{d} \cdot \left(\sum_{i=0}^{d} (-1)^i \binom{d}{i} \cdot \alpha_i \right) \cdot q^d. \tag{2}$$

Hence, if $\mathtt{P}_{\overline{\alpha}}$ 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 equation in $t+1$ unknowns. Thus, the only solutions to this system are vectors that are spanned by $\overline{u}$, and the claim follows. ■

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

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

Proof: Let $\overline{\alpha} = (\alpha_0, \alpha_1, ..., \alpha_t) \in [0,1]^{t+1}$ be as guaranteed by Proposition 2.1. Then, the (degree t) polynomial $\mathtt{P}_{\overline{\alpha}}$ "oscillates" in $[0,1]$, while obtaining the value $\mathtt{P}_{\overline{\alpha}}(\tau_1)$ on the t points $\tau_1, \tau_2, ..., \tau_t$ (and only on these points). Thus, for every $i \in [t]$ and all sufficiently small $\epsilon > 0$, exactly one of the values $\mathtt{P}_{\overline{\alpha}}(\tau_i - \epsilon)$ and $\mathtt{P}_{\overline{\alpha}}(\tau_i + \epsilon)$ is larger than $\mathtt{P}_{\overline{\alpha}}(\tau_1)$ (and the other is smaller than it). Without loss of generality, it holds that $\mathtt{P}_{\overline{\alpha}}(q) \geq \mathtt{P}_{\overline{\alpha}}(\tau_1)$ for every q in $\mathcal{D}_{\overline{\tau}}$ and $\mathtt{P}_{\overline{\alpha}}(q) < \mathtt{P}_{\overline{\alpha}}(\tau_1)$ otherwise.[7] Furthermore, we claim that there exists a constant γ such that, for any probability q that is ϵ-far from $\mathcal{D}_{\overline{\tau}}$, it holds that $\mathtt{P}_{\overline{\alpha}}(q) \leq \mathtt{P}_{\overline{\alpha}}(\tau_1) - \gamma \cdot \epsilon$. This claim can be proved by considering the Taylor expansion of $\mathtt{P}_{\overline{\alpha}}$; specifically, expanding $\mathtt{P}_{\overline{\alpha}}(q)$ based on the value at τ_i yields

$$\mathtt{P}_{\overline{\alpha}}(q) = \mathtt{P}_{\overline{\alpha}}(\tau_i) + \mathtt{P}'_{\overline{\alpha}}(\tau_i) \cdot (q - \tau_i) + \sum_{j=2}^{t} \frac{\mathtt{P}^{(j)}_{\overline{\alpha}}(\tau_i)}{j!} \cdot (q - \tau_i)^j, \tag{3}$$

[7]Otherwise, use $\mathtt{P}_{(1-\alpha_0,...,1-\alpha_t)}$, which equals $1 - \mathtt{P}_{\overline{\alpha}}$.

where $\mathtt{P}'_{\overline{\alpha}}$ is the derivative of $\mathtt{P}_{\overline{\alpha}}$ and $\mathtt{P}^{(j)}_{\overline{\alpha}}$ is the $j^{\rm th}$ derivative of $\mathtt{P}_{\overline{\alpha}}$. By the above, $\mathtt{P}'_{\overline{\alpha}}(\tau_i) \neq 0$ (for all $i \in [t]$); furthermore, $\mathtt{P}'_{\overline{\alpha}}(\tau_i)$ is positive if and only if i is odd. Let $v \stackrel{\rm def}{=} \min_{i\in[t]}\{|\mathtt{P}'_{\overline{\alpha}}(\tau_i)|\} > 0$ and $w \stackrel{\rm def}{=} \max_{i\in[t],j\geq 2}\{|\mathtt{P}^{(j)}_{\overline{\alpha}}(\tau_i)|/j!\}$. (We may assume that $w > 0$ since otherwise $\mathtt{P}_{\overline{\alpha}}(q) = \mathtt{P}_{\overline{\alpha}}(\tau_i)+\mathtt{P}'_{\overline{\alpha}}(\tau_i)\cdot(q-\tau_i)$, and the claim follows.) Hence, for all sufficiently small $\epsilon > 0$ (say for $\epsilon \leq \min(1, v)/3w$), if $q \in [\tau_i \pm \epsilon]$, then $\sum_{j=2}^{t} \frac{|\mathtt{P}^{(j)}_{\overline{\alpha}}(\tau_i)|}{j!}\cdot(q-\tau_i)^j$ is upper bounded by $\sum_{j\geq 2} w\cdot\epsilon\cdot(v/3w)\cdot(1/3)^{j-2} = v\cdot\epsilon/2$; and so, $\mathtt{P}_{\overline{\alpha}}(q) = \mathtt{P}_{\overline{\alpha}}(\tau_i) + \mathtt{P}'_{\overline{\alpha}}(\tau_i)\cdot(q-\tau_i) \pm v\cdot\epsilon/2$. It follows that, for every $j \leq \lceil t/2\rceil$, it holds that $\mathtt{P}_{\overline{\alpha}}(\tau_{2j-1}-\epsilon) \leq \mathtt{P}_{\overline{\alpha}}(\tau_{2j-1}) - v\cdot\epsilon/2$ and $\mathtt{P}_{\overline{\alpha}}(\tau_{2j}+\epsilon) \leq \mathtt{P}_{\overline{\alpha}}(\tau_{2j}) - v\cdot\epsilon/2$ (since $\mathtt{P}'_{\overline{\alpha}}(\tau_{2j-1}-\epsilon) \geq v$ and $\mathtt{P}'_{\overline{\alpha}}(\tau_{2j}+\epsilon) \leq -v$). Using $\gamma = \min(1, v)/3tw$, the claim holds for all $\epsilon \leq 1$. ■

Sample optimality: We first note that no generality is lost by focusing on generic testers (for $\mathcal{D}_{\overline{\tau}}$), because a k-sample POT can be made generic by letting α_i be the average probability that the original POT accepts when seeing exactly i samples that equal 1.[8] Recall that we analyzed a generic tester that uses $k = t$ samples for testing a property parameterized by t thresholds (i.e., $\overline{\tau} = (\tau_1, ..., \tau_t)$). The foregoing considerations can be employed to show that using t samples (i.e., $k \geq t$) is necessary. Specifically, recall that the t conditions listed in Proposition 2.1 represent necessary conditions regarding the behavior of a generic tester $T_{\overline{\alpha}}$, which means that for $\overline{\alpha} = (\alpha_0, \alpha_1, ..., \alpha_k)$ the degree k polynomial $\mathtt{P}_{\overline{\alpha}}$ must be non-constant and attain the same value on t different points (i.e., the τ_i's). These conditions imply that the solution space of a specific linear system of $t-1$ equations (i.e., $\mathtt{P}_{\overline{\alpha}}(\tau_i) = \mathtt{P}_{\overline{\alpha}}(\tau_1)$ for $i = 2, ..., t$) in $k+1$ variables (i.e., $\alpha_0, ..., \alpha_k$) must have dimension at least two. Rewriting $\mathtt{P}_{\overline{\alpha}}(q)$ as $(1-q)^k\cdot\sum_{i=0}^{k}\binom{k}{i}\cdot(q/(1-q))^i\cdot\alpha_i$, we observe that the $\mathtt{P}_{\overline{\alpha}}(\tau_i)$'s are linearly independent combinations of $\alpha'_0, ..., \alpha'_k$, where $\alpha'_i = \binom{k}{i}\alpha_i$. It follows that the dimension of the solution space is at most $k+1-(t-1)$, which implies $k-t+2 \geq 2$.

The case of $t = 2$: The considerations underlying the proof of Theorem 2.2 imply that in the case of $t = 2$ the polynomial $\mathtt{P}_{\overline{\alpha}}$ is quadratic and equals $\mathtt{P}_{\overline{\alpha}}(q) = \alpha_0 - 2(\alpha_0-\alpha_1)\cdot q + (\alpha_0 - 2\alpha_1 + \alpha_2)\cdot q^2$ (cf. Eq. (2)). Thus, $\mathtt{P}_{\overline{\alpha}}$ obtains its maximum at the point $\tau \stackrel{\rm def}{=} (\tau_1+\tau_2)/2$, which in turn equals $\frac{2(\alpha_0-\alpha_1)}{2\cdot(\alpha_0-2\alpha_1+\alpha_2)}$. The derivative of $\mathtt{P}_{\overline{\alpha}}$ at τ_2 (and likewise $-\mathtt{P}'_{\overline{\alpha}}(\tau_1)$) equals

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

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 $\alpha_0 = \frac{1-2\tau}{1-\tau} \in [0, 1]$).[9] In both cases, letting $\gamma = \max(\tau, 1-\tau) \in [0, 5, 1)$, we obtain

$$
-\mathtt{P}'_{\overline{\alpha}}(\tau_2) = \frac{\tau_2-\tau_1}{\tau}\cdot(\alpha_1-\alpha_0)
$$

[8]Indeed, the original tester may accepts with probability β_σ when the sequence of samples equals $\sigma \in \{0, 1\}^k$. We derive a generic tester by letting α_i be the average of the β_σ's taken over all $\sigma \in \{0, 1\}^k$ of Hamming weight i.

[9]Note that this is the setting used in Footnote 5.

$$= \frac{\tau_2 - \tau_1}{\gamma} \geq 2(\tau_2 - \tau_1),$$

which means that distributions that are ϵ-far from $\mathcal{D}_{\overline{\tau}}$ 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 sets $D_{\overline{\tau}}$ such that $\overline{\tau} = (\tau_1, ..., \tau_t)$ and $0 < \tau_1 < \tau_2 < \cdots < \tau_t < 1$. Recall that this set 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 $\mathcal{D}_{\overline{\tau}}$ all distributions X such that $\Pr[X = 1] \leq \tau_2$. In such case we define the sequence $\overline{\alpha} = (\alpha_0, \alpha_1, ..., \alpha_{t-1})$ as guaranteed by Proposition 2.1 such that the polynomial $\mathtt{P}_{\overline{\alpha}}$ of degree $t-1$ is non-constant and satisfies $\mathtt{P}_{\overline{\alpha}}(\tau_i) = \mathtt{P}_{\overline{\alpha}}(\tau_2)$ for all $i \geq 2$. Analogously we treat the case of t being even and $\tau_t = 1$, which corresponds to including in $\mathcal{D}_{\overline{\tau}}$ all distributions X such that $\Pr[X = 1] \geq \tau_{t-1}$. In both cases the induced tester $T_{\overline{\alpha}}$ is a POT for $\mathcal{D}_{\overline{\tau}}$.

We consider also the case where $\tau_{2i-1} = \tau_{2i}$ for some i's; that is, some of the allowed intervals can be collapsed to single points. Consider, for example, the class of distributions $D_{\tau,\tau}$, for some $\tau \in (0, 1)$. The foregoing design of a POT for D_{τ_1,τ_2} can be easily adapted for the case of $D_{\tau,\tau}$. Specifically, rather than ensuring that $\mathtt{P}_{\overline{\alpha}}(\tau_1) = \mathtt{P}_{\overline{\alpha}}(\tau_2)$, we ensure that $\mathtt{P}_{\overline{\alpha}}$ obtains a maximum at τ (equiv., $\mathtt{P}'_{\overline{\alpha}}(\tau) = 0$), which is actually what we did in the case of $t = 2$ in Section 2.1 for $\tau \stackrel{\rm def}{=} (\tau_1 + \tau_2)/2$. 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 note that $\mathtt{P}_{\overline{\alpha}}(\tau \pm \epsilon)$ equals $\mathtt{P}_{\overline{\alpha}}(\tau) + \mathtt{P}^{(2)}_{\overline{\alpha}}(\tau) \cdot \epsilon^2/2$, where $\mathtt{P}^{(2)}_{\overline{\alpha}}(\tau) = 2(\alpha_0 - 2\alpha_1 + \alpha_2) = 2(\alpha_0 - \alpha_1)/\tau$, which equals $-2/\tau$ if $\tau \geq 1/2$ and $-2/(1-\tau)$ otherwise. Hence, $\mathtt{P}_{\overline{\alpha}}(q) \leq \mathtt{P}_{\overline{\alpha}}(\tau) - (q - \tau)^2$, and we get

Proposition 2.3 (POT for $D_{\tau,\tau}$): *For every $\tau \in (0, 1)$, there exists a two-sample POT with quadratic detection probability* (actually, $\varrho(\delta) = \delta^2$ will do).

(A much simpler presentation is begging in the case that $\tau = 1/2$.)[10] More generally, we get

Theorem 2.4 (POT for $D_{\overline{\tau}}$, general case): *For every sequence $\overline{\tau} = (\tau_1, ..., \tau_t) \in [0, 1]^t$ such that $\tau_1 \leq \tau_2 < \tau_3 \leq \tau_4 < \cdots < \tau_{t-1} \leq \tau_t$, there exists a sequence $\overline{\alpha} = (\alpha_0, \alpha_1, ..., \alpha_t) \in [0, 1]^{t+1}$ such that $T_{\overline{\alpha}}$ is a POT with quadratic detection probability for $D_{\overline{\tau}}$. Furthermore, if $\tau_{2i-1} = \tau_{2i}$ for every $i \in [\lceil t/2 \rceil]$, then $\mathtt{P}_{\overline{\alpha}}(q) = \mathtt{P}_{\overline{\alpha}}(\tau_1)$ for every q in $\mathcal{D}_{\overline{\tau}}$.*

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 $\mathtt{P}_{\overline{\alpha}}$ at $t - |J|$ different points and $|J|$ additional equalities that arise from equalities imposed on the values of $\mathtt{P}'_{\overline{\alpha}}$ at $|J|$ different points. The same considerations (as in the proof of Theorem 2.2) imply the existence of a solution $\overline{\tau}$ such that $\mathtt{P}_{\overline{\alpha}}$ is not a constant function, but here the analysis of $\mathtt{P}_{\overline{\alpha}}(\tau_j \pm \epsilon)$ 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 $\mathtt{P}^{(2)}_{\overline{\alpha}}(\tau_{2j}) < 0$ (which holds since $\mathtt{P}'_{\overline{\alpha}}(\tau_{2j-1}) = \mathtt{P}'_{\overline{\alpha}}(\tau_{2j}) = 0$ but $\mathtt{P}'_{\overline{\alpha}}(\tau_{2j-1} - e) > 0$ and $\mathtt{P}'_{\overline{\alpha}}(\tau_{2j} + \epsilon) < 0$ for all sufficiently small $\epsilon > 0$) In the latter case, for sufficiently small $\epsilon > 0$ and all $q \in [\tau_{2j} \pm \epsilon]$, we get $\mathtt{P}_{\overline{\alpha}}(q) < \mathtt{P}_{\overline{\alpha}}(\tau_{2j}) - \mathtt{P}^{(2)}_{\overline{\alpha}}(\tau_{2j}) \cdot (q - \tau_{2j})^2/4$, since $\sum_{i=3}^{t} \frac{|\mathtt{P}^{(i)}_{\overline{\alpha}}(\tau_{2j})|}{i!} \cdot (q - \tau_{2j})^i$ is upper bounded by $\mathtt{P}^{(2)}_{\overline{\alpha}}(\tau_{2j}) \cdot (q - \tau_{2j})^2/2$. ■

[10] In 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 $2q(1-q) = \frac{1}{2} - 2(q - 0.5)^2$.

2.3 POTs can test only intervals

In this section we show that the only testable sets 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 2.4.

Theorem 2.5 (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 2.4. For the other direction, assume that $\mathcal{T}$ is POT for $\mathcal{D}_S$ that makes k queries. Then, for a view $\overline{b} = (b_1,\ldots,b_k) \in \{0,1\}^k$, the tester $\mathcal{T}$ accepts this view with some probability, denoted $\alpha_{\overline{b}} \in [0,1]$. Note that when testing a distribution X such that $\Pr[X=1]=p$, the probability of seing this view is $p^{w(\overline{b})}(1-p)^{k-w(\overline{b})}$, where $w(\overline{b}) = \sum_j b_j$ denotes the number of 1's in $\overline{b}$. Hence, when given a distribution X such that $\Pr[X=1]=p$, the acceptance probability of $\mathcal{T}$ on X is

$$\Pr[\mathcal{T} \text{ accepts } X] = \sum_{i=0}^{k} \left(\sum_{\overline{b}\in\{0,1\}^k : w(\overline{b})=i} \alpha_{\overline{b}} \right) \cdot p^i (1-p)^{k-i}, \tag{4}$$

which is a polynomial of degree k (in p). Thus, for every $c \in \mathbb{R}$, the set of points $p \in [0,1]$ on which the value of this polynomial is at least c equals a union of up to $\lceil (k+1)/2 \rceil$ intervals. In particular, this holds when 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. ■

2.4 Proof of Theorem 1.2

As should be clear by now, the positive part of Theorem 2.5 implies the positive part of Theorem 1.2: That is, a POT for the symmetric property Π is obtained by sampling elements in $[n]$, querying the Boolean function for their value, and invoking the corresponding distribution-POT.

The opposite direction require a little more care, since a tester for the (function) property Π may avoid repeated samples, while a distribution-tester can not do so. Furthermore, the behavior of the former may depend on n, whereas the behavior of the latter may not depend on the unknown sample space. Still, by considering a sufficiently large n, these effects become negligible. Hence, we may just mimic the argument used in the proof of the corresponding part of Theorem 2.5. The key observation is that the probability that a k-query tester accepts a random function $f:[n]\to\{0,1\}$ that evaluates to 1 on exactly m inputs equals

$$\sum_{i=0}^{k} \left(\sum_{\overline{b}\in\{0,1\}^k : w(\overline{b})=i} \alpha_{\overline{b}} \right) \cdot \left(\prod_{j=0}^{i-1} \frac{m-j}{n-j} \right) \cdot \left(\prod_{j=0}^{k-i-1} \frac{n-m-j}{n-j} \right) \tag{5}$$

where the $\alpha_{\overline{b}}$ are as in Eq. (4) (except that they may depend on n), and we assume (w.l.o.g.) that the tester always makes k queries and never makes the same query twice. Letting $\rho = m/n$ denote the density of 1-values in f, observe that Eq. (5) is a polynomial of degree k in ρ. The theorem follows (exactly as in the case of Theorem 2.5).

3 Graph Properties (in the Adjacency Matrix Model)

Symmetric properties of Boolean functions induce graph properties (in the adjacency matrix 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 $g : [N] \times [N] \to \{0, 1\}$ 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 set of k-regular graphs

For every function $k : \mathbb{N} \to \mathbb{N}$, 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]).[11] In contrast, we show that it has a two-sided error POT.

Theorem 3.1 (a POT for $\mathcal{R}^{(k)}$): *For every function $k : \mathbb{N} \to \mathbb{R}$, 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 $k(N)$ and $N \cdot k(N)/2$ are both integers (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 $\tau = k(N)/N$ and proceeds as follows.

1. Selects uniformly a vertex $s \in [N]$ and consider the Boolean function $f_s : [N] \to \{0, 1\}$ such that $f_s(v) = 1$ if and only if $\{s, v\} \in E$.

2. Invokes the POT of Proposition 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 set $\mathcal{D}_{\tau,\tau}$.

   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.)[12]

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 $\mathtt{P}(q)$ denote the probability that the POT invoked in Step 2 accepts a random variable X such that $\Pr[X = x] = q$, and recall that $\mathtt{P}(q) \leq \mathtt{P}(\tau) - (q - \tau)^2$ (cf. Proposition 2.3). Then, the probability that our graph tester accepts the graph G equals

[11] Specifically, 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 $k(N) < N - 1$.

[12] Recall that we may use the setting outlined at the beginning of Section 2.2 preceding Proposition 2.3. Specifically we set $(\alpha_0, \alpha_1, \alpha_2) = (0, 1, (2\tau - 1)/\tau)$ if $\tau \geq 1/2$ and $(\alpha_0, \alpha_1, \alpha_2) = (\frac{1-2\tau}{1-\tau}, 1, 0)$ otherwise.

$$\frac{1}{N} \cdot \sum_{s \in [N]} \mathtt{P}(d_G(s)/N), \tag{6}$$

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 $\mathtt{P}(\tau)$. 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 $\mathtt{P}(\tau) - \Omega(\epsilon^2)$.

Claim 3.1.1 ("local" versus "global" distance from $\mathcal{R}^{(k)}$): *If $k(N)$ and $k(N)N/2$ are natural numbers and $\sum_{v \in [N]} |d_G(v) - k(N)| \le \epsilon' \cdot N^2$, 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, using $\mathtt{P}(q) \le \mathtt{P}(\tau) - (q-\tau)^2$, we have:

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

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 $\mathtt{P}(\tau) - (\epsilon/6)^2$. The theorem follows. ■

3.2 Other regular graph properties

The two-sided error POT guaranteed by Theorem 3.1 can be combined with one-sided error 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 generic POT for $\mathcal{R}^{(k)} \cap \Pi$): *Let Π be a graph property that has a one-sided error POT and $k : \mathbb{N} \to \mathbb{R}$. Suppose that there exists a monotone function $F : (0,1] \to (0,1]$ 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.*

We note that the "intersection" condition made in Theorem 3.2 does not hold in general. For example, consider $\Pi = \mathcal{BCC} \cup \mathcal{CC}$, where $\mathcal{CC}$ is the set of all graphs that consist of a collection of isolated cliques and $\mathcal{BCC}$ is the set of all graphs that that consist of a collection of isolated bicliques.

Both $\mathcal{CC}$ and $\mathcal{BCC}$ were studied in [GR09a], and since each is a subgraph-freeness property it follows that each has one-sided error POTs (see [GR09b, Thm. 4.7]), and so does their union (i.e., Π).[13] 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 (for even N) the graph G that consists of two $N/2$-vertex cliques is 0.49-far from $\mathcal{R}^{(k)} \cap \Pi$. On the other hand, G is in $\mathcal{CC} \subset \Pi$ and is $1/N$-close to $\mathcal{R}^{(k)}$ (by virtue of adding a perfect matching between the two cliques). Hence, Π does not satisfy the "intersection" condition (w.r.t $\mathcal{R}^{(k)}$ for $k(N) = N/2$). We stress that this fact does not mean that $\Pi \cap \mathcal{R}^{(k)}$ has no POT; actually, $\Pi \cap \mathcal{R}^{(k)} = \mathcal{BCC} \cap \mathcal{R}^{(k)}$ does have a POT, since $\mathcal{BCC}$ satisfies the "intersection" condition (w.r.t this $\mathcal{R}^{(k)}$).[14] Other properties that satisfy the condition appear in Proposition 3.3; one such property is the set Π consisting of all complete tripartite graphs where the condition holds with respect to $k(N) = 2N/3$.

Proof: On input N and oracle access to an N-vertex graph $G = ([N], E)$, the tester proceeds as follows (while assuming that $k(N)$ is an integer and $k(N)N$ is even).[15]

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

2. Inkoves 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 ("inteersection") 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 $k(N) = (t-1)N/t$.*

[13] Indeed, the class having one-sided error POTs is closed under union (e.g., invoke both POTs and accept iff at least one of them accepts). This is not necessarily the case for two-sided error POTs; see Section 5.2.

[14] If G is δ-close to both $\mathcal{BCC}$ and $\mathcal{R}^{(k)}$, then there exists a graph $G' \in \mathcal{BCC}$ that is δ-close to G. It follows that G' contains a biclique with $(0.5 \pm O(\delta)) \cdot N$ vertices on each side, which implies that G' is $O(\delta)$-close to $\mathcal{BCC} \cap \mathcal{R}^{(k)}$.

[15] Otherwise, the tester rejects upfront, since no N-vertex graph can be $k(N)$-regular.

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.

Considering the complementary graphs, this establishes a two-sided error POT for the set of graphs consisting of t isolated cliques of equal size. We mention that a more general result that refers to t isolated cliques having densities that fits one out of a constant number of fixed sequences is proved in Proposition 6.4 of our technical report [GS12].[16]

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 $(V_1, ..., V_t)$ 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 δ-close to G, and so G' is 2δ-close to $\mathcal{R}^{(k)}$. Let $I_1, ..., I_t$ be the partition of G' to t independent sets such that there is a complete bipartite graph between each two I_j's. Since G' is 2δ-close to $\mathcal{R}^{(k)}$, 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 $\sum_{i \in [t]} (x_i - (1/t))^2 = (\sum_{i \in [t]} x_i^2) - 1/t \leq 4\delta$, which implies $|x_i - (1/t)| \leq \sqrt{4\delta}$ for every $i \in [t]$. By moving at most $O(\sqrt{\delta} \cdot N)$ vertices between the various I_i's and modifying the edges accordingly, we conclude that G' is $O(\sqrt{\delta})$-close to Π' (and the same holds for G). ■

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 $S_0, \ldots, S_{\ell-1}$, vertices $u, v \in \bigcup_{i \in \{0,1,...,\ell-1\}} S_i$ are connected by an edge if and only if for some $i \in \{0, 1, ..., \ell - 1\}$ 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 $\mathcal{S}\mathsf{C}_\ell\mathcal{C}$. As is shown in our technical report [GS12], for every $\ell \leq 3$, there is a dichotomy in the behavior of the set $\mathcal{S}\mathsf{C}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$: For some integers t and $k(N) = 2N/t\ell$, the sets $\mathcal{S}\mathsf{C}_\ell\mathcal{C}$ and $\mathcal{R}^{(k)}$ satisfy the "intersection" condition of Theorem 3.2, whereas for the remaining values of t the said condition is not satisfied. Specifically, the condition is satisfied for every $t \in \mathbb{N}$ when ℓ is a multiple of four, and is satisfied by at most three values of t (i.e., $t \in \{1, 2, 3\}$) otherwise. For details see Propositions 3.4 and 3.5 in [GS12]. We stress that the fact that for some ℓ and k the set $\mathcal{S}\mathsf{C}_\ell\mathcal{C}$ does not satisfy the "intersection" condition (w.r.t $\mathcal{R}^{(k)}$) does not mean that $\mathcal{S}\mathsf{C}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$ has no POT. In fact, we ask whether $\mathcal{S}\mathsf{C}_\ell\mathcal{C} \cap \mathcal{R}^{(k)}$ always has a POT. More generally:

Open Problem 3.4 (on POTs for $\mathcal{R}^{(k)} \cap \Pi$): *Let Π be a graph property that has a one-sided error POT and $k : \mathbb{N} \to \mathbb{N}$. Does it always hold that $\Pi \cap \mathcal{R}^{(k)}$ has a two-sided error POT?*

[16] That is, for a fixed sequence $((\rho_1^{(1)}, ..., \rho_t^{(1)}), ..., (\rho_1^{(k)}, ..., \rho_t^{(k)}))$, the set consists all graphs consisting of t isolated cliques such that for some $i \in [k]$ and for every $j \in [t]$ the j^{th} clique has density $\rho_j^{(i)}$.

3.3 The set of regular graphs

We consider the set $\mathcal{REG}$ of regular graphs. Note that this set strictly contains the sets 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 subsets.) Clearly, $\mathcal{REG}$ has no one-sided error POT. In contrast, we show that it has a two-sided error POT.

Theorem 3.5 (a POT for $\mathcal{REG}$): *The set $\mathcal{REG}$ 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 densities of the neighbor sets of the two chosen vertices.[17] We show first that the task of testing that two Boolean distributions are equal can be performed in a proximity oblivious manner (and will return to Theorem 3.5 later).

Proposition 3.6 (a POT for $\mathcal{EQ}$): *Let $\mathcal{EQ} = \{(P,Q) : \Pr[P\!=\!1] = \Pr[Q\!=\!1]\}$ be the set of pairs of equal Boolean distributions, and let the distance of a pair (P,Q) from $\mathcal{EQ}$ be defined as*

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

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 recipe 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

$$\alpha_{(i,j)} = \begin{cases} c - 2\delta & \text{if } (i,j) \in \{(0,2),(2,0)\} \\ c + \delta & \text{if } (i,j) = (1,1) \\ c & \text{otherwise} \end{cases} \tag{7}$$

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

[17] This reduction is analogous to the proof of Theorem 3.1 in which we check that the degree of a vertex equals a predetermined value $k(N) \in \mathbb{N}$ that is *a priori* known to the tester.

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

$$\Pr[T \text{ accepts } (P,Q)] = \sum_{i,j\in\{0,1,2\}} \alpha_{(i,j)} \cdot \binom{2}{i} p^i (1-p)^{2-i} \cdot \binom{2}{j} q^j (1-q)^{2-j}. \tag{8}$$

Note that almost all the $\alpha_{(i,j)}$'s in Eq. (7) 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. (8), we get

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

The proposition follows. ■

Proof of Theorem 3.5. 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 $f_{v_1} : [N] \to \{0,1\}$ and $f_{v_2} : [N] \to \{0,1\}$ such that $f_{v_i}(w) = 1$ if and only if $\{v_i, w\} \in E$.

2. Invoke the POT of Proposition 3.6 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.

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). Using Proposition 3.6, it follows that every non-regular graph is accepted with probability at most $c - \Omega(N^{-3})$. This is because such a graph contains at least $N-1$ pairs of vertices of different degrees, and if any such pair is chosen in Step 1, then Step 2 accepts with probability at most $c - \Omega(N^{-2})$. (Note that this argument as well as the rest relies on the fact that Step 2 accepts with probability at most c, regardless of the choice made in Step 1.)

Suppose now that a (non-regular) graph $G = ([N], E)$ is accepted with probability exactly $c - \epsilon$, where $\epsilon = \Omega(N^{-3})$. Then, by an averaging argument, there are at least $(1 - \epsilon^{1/3}) \cdot N^2$ pairs of vertices (v_1, v_2) such that when such a vertex-pair is chosen in Step 1, then Step 2 accepts with probability at least $c - \epsilon^{2/3}$. Thus, there exists a vertex $v_1 \in [N]$ and a subset $V_2 \subseteq [N]$ of size at least $(1 - \epsilon^{1/3})N$ 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 $c - \epsilon^{2/3}$. Proposition 3.6 implies that for every such vertex v_2 it holds that $|\Pr[X_{v_1}=1] - \Pr[X_{v_2}=1]| = O(\sqrt{\epsilon^{2/3}})$, and therefore $|d_G(v_2) - d_G(v_1)| = O(\epsilon^{1/3} \cdot N)$.

Let $K = d_G(v_1)$ be the degree of v_1. Then, using the fact that $|V_2| \geq (1 - \epsilon^{1/3})N$ and that for every $v_2 \in V_2$ it holds that $|d_G(v_2) - K| = O(\epsilon^{1/3}N)$, we have

$$\sum_{v \in [N]} |d_G(v) - K| = \sum_{v \in V_2} |d_G(v) - K| + \sum_{v \in [N] \setminus V_2} |d_G(v) - K| = O(\epsilon^{1/3} \cdot N^2).$$

Assuming that KN is even and applying Claim 3.1.1, we infer that the graph G is $O(\epsilon^{1/3})$-close to being K-regular. If KN is odd, then we redefine $K = d_G(v_1) + 1$ (or $K = d_G(v_1) - 1$), and upper bound $\sum_{v \in [N]} |d_G(v) - K|$ by $O(\epsilon^{1/3} \cdot N^2) + N = O(\epsilon^{1/3} \cdot N^2)$, using $\epsilon = \Omega(N^{-3})$. Applying Claim 3.1.1, we infer that G is $O(\epsilon^{1/3})$-close to being K-regular. This completes the proof of Theorem 3.5. ■

Remark 3.7 (on POTs for $\Pi \cap \mathcal{REG}$): A result analogous to Theorem 3.2 can be proved in the current context. That is, *$\Pi \cap \mathcal{REG}$ has a two-sided error POT if Π has a one-sided error POT and there exists a monotone function $F : (0,1] \to (0,1]$ such that every graph that is δ-close to both Π and $\mathcal{REG}$ is $F(\delta)$-close to $\Pi \cap \mathcal{REG}$.*

We note that the four queries made by the POT for $\mathcal{EQ}$ (i.e., the POT asserted in Proposition 3.6) are in some sense the minimum number of queries possible.

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

Indeed, as pointed out by Johan Håstad, 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.[18]

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, 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

$$A(p,q) \stackrel{\text{def}}{=} \Pr[T \text{ accepts } (P,Q)] = \sum_{i \in \{0,1,\ldots,t\}} \binom{t}{i} p^i (1-p)^{t-i} \cdot \left(q\alpha_{(i,1)} + (1-q)\alpha_{(i,0)}\right). \tag{9}$$

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

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

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

[18] This 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 of our technical report [GS12]) 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.

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 $A(p, p) \geq c$ must hold).

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 $B(p) \geq c$ and $|D(p)| \leq 1$), which violates the requirement $A(p, q) < c$.

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 Observation 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. ■

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. We shall denote the number of vertices in this graph by n, and stress that n *is a constant* since we refer to a fixed graph. In contrast, the number of vertices in the tested graph will be denoted N, and so N is a variable.

Fixing any n-vertex graph H, denote by $\rho_H(G)$ the density of H as an induced 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 set of H-free graphs. Alon *et al.* [AFKS] showed that, for some monotone function $F_n : (0, 1] \to (0, 1]$ if G is δ-far from the set 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).[19]

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 rejects if and only if the induced subgraph is isomorphic to H. This POT accepts any 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-1} is $(\delta - 2(i-1)/N)$-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 (by an averaging argument) 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. Omitting from G_{i-1} all edges incident at v_i, we obtain a graph G_i that is ϵ-close to G_{i-1}, where $\epsilon = (N-1)/\binom{N}{2} = 2/N$. 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 $v_1, ..., v_{i-1}$, since the latter vertices

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

are isolated in G_{i-1} (whereas H contains no isolated vertices). Also note that the copies of H counted in the $i^{\rm th}$ iteration also occur in G, since they contain no vertex pair on which G_{i-1} differs from G. Thus, after $t \stackrel{\rm def}{=} \lfloor \delta N/2 \rfloor$ iterations, we obtain a graph $G_t \notin \Pi_{H,\tau}$ that contains at least $\tau \cdot \binom{N}{n}$ 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 $v_1, ..., v_{i-1}$ (this time because each vertex in $v_1, ..., v_{i-1}$ is connected to all vertices in G_{i-1}), and we can proceed as before. ■

Generalization. We now consider, for any fixed $n \in \mathbb{N}$, graph properties that are each parameterized by a sequence of weights $\overline{w} = (w_H)_{H:V(H)=[n]}$ and by $b \in [0,1]$, where $w_H \in [0,1]$ for each n-vertex graph H (with vertex set $V(H)$ that equals $[n]$). The corresponding graph property is denoted $\Pi_{n,\overline{w},b}$, and a graph G is in $\Pi_{n,\overline{w},b}$ if and only if $\sum_{H:V(H)=[n]} w_H \cdot \rho_H(G) \leq b$.

Note that the case of $b = 0$ corresponds to $\mathcal{F}$-freeness for $\mathcal{F} = \{H : w_H > 0\}$, and hence has one-sided error POTs. More generally, if for every n-vertex graph H it holds that $w_H \geq b$, then $\Pi_{n,\overline{w},b}$ equals the set of $\mathcal{F}$-free graphs, where $\mathcal{F} = \{H : V(H) = [n] \wedge w_H > b\}$. Another interesting case is where $w_{H_1} = 1$ for a unique graph H_1 and $w_H = 0$ otherwise (i.e., for every $H \neq H_1$): In this case the property $\Pi_{|V(H_1)|,\overline{w},b}$ corresponds to having an H_1-density that does not exceed b (i.e., in this case $G \in \Pi_{|V(H_1)|,\overline{w},b}$ if and only if $\rho_{H_1}(G) \leq b$, which is the case considered in Theorem 3.9).

In the rest of this section, we shall discard the case of a uniform sequence $\overline{w}$ (i.e., $w_H = w$ for some w and all H's), since in this case the property is trivial.[20] We conjecture that, for any $b > 0$ and $\overline{w} = (w_H)_{H:V(H)=[n]}$, the property $\Pi_{n,\overline{w},b}$ has a two-sided error POT, but we are only able to establish it for the following special cases.

Theorem 3.10 (Theorem 3.9, generalized): *Let $b \in (0,1]$ and $\overline{w} = (w_H)_{H:V(H)=[n]}$. 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 $\Pi_{n,\overline{w},b}$ it holds that $\sum_H w_H \cdot \rho_H(G) > b + (dn/3) \cdot \delta$, where $d \stackrel{\rm def}{=} b - \max_{H:w_H<b}\{w_H\}$ and 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.*

It follows that in these cases $\Pi = \Pi_{n,\overline{w},b}$ has a POT: This POT inspects a random sample of n vertices and accepts with probability $1 - w_H$ if the induced subgraph is isomorphic to H. Hence, it accepts any graph in Π with probability at least $1 - b$, whereas it accepts any graph G that is δ-far from Π with probability at most $1 - b - (dn/3) \cdot \delta$ (if $\delta > 6/|V(G)|$, and with probability at most $1 - b - (\delta/6)^n$ otherwise).

[20]This is the case $\sum_H w_H \cdot \rho_H(G) = w$ for every graph G, since $\sum_H \rho_H(G) = 1$.

Proof: Letting $b' = \max_{H:w_H<b}\{w_H\}$, we mimic the proof of Theorem 3.9, focusing on the case that $\{H : w_H \geq b\}$ contains only graphs with no isolated vertices. The invariant that we maintain is that G_{i-1} is $(\delta-2(i-1)/N)$-far from $\Pi = \Pi_{n,\overline{w},b}$, which implies that $\sum_{H:w_H\geq b}(w_H - b')\cdot\rho_H(G_{i-1}) > b - b' = d$. In the $i^{\rm th}$ iteration, we pick a vertex v_i that contributes the most to the latter sum, and omit all edges incident at it, obtaining G_i. We note that the contribution of v_i to the latter sum is at least $d \cdot \binom{N-1}{n-1}/\binom{N}{n}$ units. Therefore, by omitting all edges incident at v_i we reduce the sum $\sum_{H:V(H)=[n]} w_H \cdot \rho_H(G_{i-1})$ by at least $d \cdot \binom{N-1}{n-1}/\binom{N}{n} = dn/N$, i.e.,

$$\sum_{H:V(H)=[n]} w_H \cdot \rho_H(G_{i-1}) - \sum_{H:V(H)=[n]} w_H \cdot \rho_H(G_i) \geq \frac{dn}{N}.$$

After $t = \lfloor \delta N/2 \rfloor$ iterations we obtain a graph $G_t \notin \Pi_{n,\overline{w},b}$, which implies $\sum_{H:V(H)=[n]} w_H \cdot \rho_H(G_t) > b$. Hence, we obtain

$$\sum_{H:V(H)=[n]} w_H \cdot \rho_H(G) \quad \geq \quad \sum_{H:V(H)=[n]} w_H \cdot \rho_H(G_t) + \lfloor \delta N/2 \rfloor \cdot \frac{dn}{N},$$

which is at least $b + (\delta n/3) \cdot d$, provided that $\delta > 6/|V(G)|$. ■

Needless to say, a begging open problem is whether POTs exist for arbitrary properties of the form $\Pi_{n,\overline{w},b}$. That is.

Open Problem 3.11 (POT for every $\Pi_{n,\overline{w},b}$): *For every $b \in (0,1]$ and $\overline{w} = (w_H)_{H:V(H)=[n]}$, does it always hold that $\Pi_{n,\overline{w},b}$ has a two-sided error POT?*

We conjecture that the answer is positive, although we do not even know whether $\Pi_{n,\overline{w},b}$ is testable within query complexity that only depends on the proximity parameter.

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

Recall that we conjectured that, for any $b > 0$ and $\overline{w} = (w_H)_{H:V(H)=[n]}$, the property $\Pi_{n,\overline{w},b}$ has a two-sided error POT (see Open Problem 3.11), but we are only able to establish it in special cases (see Theorem 3.10). On the other hand, we shall 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 $\overline{w} = (w_H)$ (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 also on N (and is not upper bounded by a function of n). For example, for every $m(N) \in \{0,1,...,\binom{N}{2}\}$, we may consider the property of N-vertex graphs having at most $m(N)$ edges.[21]

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

1. A "global" notion of distance: The graph G is δ-far from $\Pi_{n,\overline{w},b}$; that is, for every $G' \in \Pi_{n,\overline{w},b}$ it holds that G and G' differ on at least δ fraction of vertex pairs.

2. A "local" notion of distance: 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 $G \notin \Pi_{n,\overline{w},b}$ 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, although we do have such bounds in special cases (e.g., either $b = 0$ or $|\{H : w_H > 0|\}| = 1$, see Section 3.4).[22] In light of this state of affairs, a first step towards a characterization is provided by the following result.

Theorem 3.12 (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 monotone function $F : (0,1] \to (0,1]$ such that for every N there exists a sequence $\overline{w} = (w_H)_{H:V(H)=[n]}$ that satisfies the following two conditions:*

1. *Π_N equals the set of N-vertex graphs in $\Pi_{n,\overline{w},b}$.*

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 3.10 presents cases in which this condition holds. Note that while one direction of Theorem 3.12 is quite obvious (i.e., properties that correspond to such sequences of $\Pi_{n,\overline{w},b}$'s have a POT), the opposite direction requires a proof (i.e., having a POT implies a correspondence to such sequences of $\Pi_{n,\overline{w},b}$'s).

Proof: The proof of the "only if" direction (i.e., having a POT implies the existence of suitable $\Pi_{n,\overline{w},b}$'s) 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).[23] Let c be the acceptance threshold of the original POT (i.e., $c = \min_{G\in\Pi}\{\Pr[\mathrm{Test}^G(|V(G)|) = 1]\}$). 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 $\Pi_{n,\overline{w},b}$ for $w_H = 1 - \alpha_H$ and $b = 1 - c$. That is, this $\overline{w}$ 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 $\overline{w}$ 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.

[21] This 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.

[22] A more general result is presented in Theorem 3.10.

[23] We 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).

Suppose, on the other hand, that for some n, b and F, it holds that for every N there exists a sequence $\overline{w} = (w_H)_{H:V(H)=[n]}$ that satisfies the two conditions (i.e., (i) Π_N equals the set of N-vertex graphs in $\Pi_{n,\overline{w},b}$, 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 $\varrho = F$. ■

Remark 3.13 (obliviousness of N): *The proof of Theorem 3.12 implies that if the final decision of the POT is oblivious of N* (i.e., the derived α_H's are independent of N)*, then so is the sequence of w_H's, and vice versa.*

Discussion. As admitted upfront, Theorem 3.12 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 $\Pi_{\overline{w},b}$ implies $\sum_H w_H \cdot \rho_H(\cdot) \geq b + F(\delta)$). By Theorem 3.10 this condition is satisfied by many properties of the form $\Pi_{n,\overline{w},b}$. In contrast, we observe that not every property $\Pi_{n,\overline{w},b}$ satisfies the second condition of Theorem 3.12. Specifically, we show the following.

Proposition 3.14 (violating the second condition of Theorem 3.12): *There exists $b \in (0,1)$, $n = O(1)$ and $\overline{w} = (w_H)_{H:V(H)=[n]}$ such that for every N there exists an N-vertex graph G that is $\Omega(1)$-far from $\Pi_{n,\overline{w},b}$ and yet $\sum_{H:V(H)=[n]} w_H \cdot \rho_H(G) = b + O(1/N)$.*

Note that this does not mean that $\Pi_{n,\overline{w},b}$ does not have a POT, since such a possible POT may use an alternative characterization of the same property (i.e., $\Pi_{n,\overline{w},b}$ may equal $\Pi_{n,\overline{w}',b'}$ such that the former violates the second condition of Theorem 3.12 whereas the latter satisfies this very condition).[24] In fact, that is the case of the property used in the following proof (i.e., it does have a POT, which uses an alternative sequence $\overline{w}'$).

Proof: We shall derive the claimed $b \in (0,1)$, $n = O(1)$ and $\overline{w} = (w_H)_{H:V(H)=[n]}$ by considering a (weak) tester (for some $\Pi \cap \mathcal{R}^{(k)}$) that emerges from the proof of Theorem 3.2 and applying the translation presented in the proof of Theorem 3.12 to this tester. Indeed, the transformations presented in both proofs are intended to be used to derive positive results, when some conditions hold, but the transformations yield what we need (although in our case the corresponding conditions are violated).

Let $\Pi = \mathcal{BCC} \cup \mathcal{CC}$, where $\mathcal{CC}$ is the set of all graphs that consist of a collection of isolated cliques and $\mathcal{BCC}$ is the set of all graphs that that consist of a collection of isolated bicliques.[25] Recall that, 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 (for even N) the graph G_N that consists of two $N/2$-vertex cliques is 0.49-far from $\mathcal{R}^{(k)} \cap \Pi$. On the other hand, G_N is in $\mathcal{CC} \subset \Pi$ and is $1/N$-close to $\mathcal{R}^{(k)}$ (by virtue of adding a perfect matching between the two cliques). Recall that $\Pi \cap \mathcal{R}^{(k)}$ does have a POT; but this is irrelevant to our argument (but it implies that $\Pi \cap \mathcal{R}^{(k)}$ equals some $\Pi_{n,\overline{w}',b'}$ that satisfies the second condition of Theorem 3.12).

[24] Needless to say, the corresponding set $\{H : w_H \geq b\}$ contains both a graph with an isolated vertex and a graph in which some vertex neighbors all other vertices.

[25] Indeed, Π was used in Section 3.2 as an example violating the "intersection" condition made in Theorem 3.2.

Next, consider the tester for $\Pi \cap \mathcal{R}^{(k)}$ described in the proof of Theorem 3.2: This tester selects a random sample of $n = O(1)$ vertices, and inspects 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 specific tester accepts with probability $\alpha'_{H'}$ if $H'' \in \mathtt{AS}$ and rejects otherwise, where $(\alpha'_{H'})_{H'}$ denotes the sequence of probabilities used by the POT of $\mathcal{R}^{(k)}$ and $\mathtt{AS}$ denotes the set of all possible induced subgraphs of graphs in Π. 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 (for $\mathcal{R}^{(k)}$) with exactly the same probability).[26] Denoting by $\mathtt{H}(H', H'')$ the set of all n-vertex graphs H such that the subgraph induced by the first three (resp., last $n-3$) vertices of H equals H' (resp., H''), we observe that for every graph G it holds that $\sum_{H \in \mathtt{H}(H',H'')} \rho_H(G) = \rho_{H'}(G) \cdot \rho_{H''}(G) \pm O(n/N)$, where the error term is due to the probability that a random sample of three vertices in $V(G)$ intersects a random sample of $n-3$ vertices (in G).[27] Denoting by α_H the probability that the test for $\Pi \cap \mathcal{R}^{(k)}$ accepts a graph when seeing the induced subgraph H, recall that, for every $H \in \mathtt{H}(H', H'')$, it holds that $\alpha_H = \alpha'_{H'}$ if $H'' \in \mathtt{AS}$ and $\alpha_H = 0$ otherwise. Then, for every N-vertex graph G, we have

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

Recall that $n = O(1)$ and thus $O(2^{n^2}/N) = O(1/N)$. Then, for the N-vertex graph G_N asserted upfront, we have $\sum_{H:V(H)=[n]} \alpha_H \cdot \rho_H(G_N) = c' - O(1/N)$, because $G_N \in \Pi$ implies $\sum_{H'' \in \mathtt{AS}} \rho_{H'}(G_N) = 1$ whereas the fact that G_N is $1/N$-close to $\mathcal{R}^{(k)}$ implies $\sum_{H':V(H')=[3]} \alpha'_{H'} \rho_{H'}(G_N) = c' - O(1/N)$ (since for every G' that is ϵ-close to G'' and for every 3-vertex H' it holds that $|\rho_{H'}(G') - \rho_{H'}(G'')| \leq 3\epsilon$). Finally, using the same translation as in the proof of Theorem 3.12 (i.e., $b = 1 - c'$ and $w_H = 1 - \alpha_H$), we conclude that although G_N is $\Omega(1)$-far from $\mathcal{R}^{(k)} \cap \Pi = \Pi_{n,\overline{w},b}$ it holds that $\sum_H w_H \cdot \rho_H(G_N) = b + O(1/N)$. The claim follows. ■

Failure of a natural conjecture. Proposition 3.14 indicates that some graph properties can be captured by sequences of w_H's that do not satisfy the second condition of Theorem 3.12 (i.e., that being δ-far from $\Pi_{n,\overline{w},b}$ implies $\sum_H w_H \cdot \rho_H(\cdot) \geq b + F(\delta)$). But this does not mean that these graph properties have no POTs, since it may be the case that these properties can be captured by alternative sequences of w_H's that do satisfy the second condition of Theorem 3.12. A *natural conjecture* would thus state that any graph property that satisfy the first condition of Theorem 3.12 has a POT. Note that this conjecture is stronger statement than the one offered in Open Problem 3.11, which refers to properties of the form $\Pi_{n,\overline{w},b}$ (whereas the conjecture refers to properties $\cup_{N \in \mathbb{N}} \Pi_N$ such that for some $b \in (0,1), n \in \mathbb{N}$ and every $N \in \mathbb{N}$ the set Π_N equals the set of all N-vertex

[26] Recall that $\sum_{H'} \rho_{H'}(G) \cdot \alpha'_{H'}$ represents the probability that this POT (for $\mathcal{R}^{(k)}$) accepts the graph G.

[27] Note that $\rho_{H'}(G) \cdot \rho_{H''}(G)$ represents the probability that such independent samples yield the views (H', H''), whereas $\sum_{H \in \mathtt{H}(H',H'')} \rho_H(G)$ represents the probability that a random sample of n vertices yields these views.

graphs in $\Pi_{n,\overline{w},b}$). A counterexample to the aforementioned conjecture is provided Section 3.6, where we establish the following result.

Proposition 3.15 (the first condition of Theorem 3.12 does not suffice for a POT): *There exist a graph property $\Pi = \cup_N \Pi_N$ such that for some fixed $n \in \mathbb{N}$ and for every $N \in \mathbb{N}$ there exists a sequence $\overline{w} = (w_H)_{H:V(H)=[n]}$ such that Π_N is equal to the set of all N-vertex graphs in $\Pi_{n,\overline{w},b}$, but Π does not have a POT.*

We note that Proposition 3.14 is not subsumed by Proposition 3.15, since the parameters $\overline{w}$ used in the former are independent of N whereas in the latter $\overline{w}$ depend on N.

3.6 Impossibility results

In this section we present two impossibility results; that is, we show that certain graph properties have no two-sided error POTs. The first result refers to Bipartitness and merely illustrates a type of impossibility results that are easy to obtain. The second result is a proof of Proposition 3.15, which is more complicated.

Natural impossibility results that are easy to obtain. It is easy to derive impossibility results regarding general POTs by considering two distributions on N-vertex graphs 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.[28]

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[29] 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:

Proposition 3.16 (an example): *Bipartiteness has no two-sided error POT.*

Proving Proposition 3.15. The rest of this section is devoted to the proof of Proposition 3.15: We show a natural graph property $\Pi = \cup_N \Pi_N$ such that (1) for some fixed $n \in \mathbb{N}$ and for every $N \in \mathbb{N}$ there exists a sequence $\overline{w} = (w_H)_{H:V(H)=[n]}$ such that Π_N is equal to the set of all N-vertex graphs in $\Pi_{n,\overline{w},b}$, but (2) Π does not have a POT. Specifically, we consider subsets 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. We start with the following definition.

[28] The foregoing method directly establishes the non-existence of a two-sided error POT. Alternatively, one may use this method to show that the second condition in Theorem 3.12 is not satisfied. Indeed, using Theorem 3.12 allows to replace (1) by (1') the two distributions have the same densities of various induced subgraphs of constant size.

[29] The N-vertex graph G is a balanced blow-up of an n-vertex graph H, if G consists of n independent sets, called clouds, each of size N/n such that the i^{th} and j^{th} clouds are connected by a complete bipartite graph if and only if the i^{th} vertex of H is connected to the j^{th} vertex of H. We stress that if the latter vertices are not connected in H, then there would be no edges between the i^{th} and j^{th} clouds of G.

Definition 3.17 (graphs consisting of two isolated cliques): *Denote by $\mathcal{CC}^{\leq 2}$ the set of all graphs that consist of at most two isolated cliques. For $0 \leq \alpha \leq \beta \leq \frac{1}{2}$, define $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ to be the set of all graphs that consist of at most two isolated cliques such that the density of the smaller clique is between α and β.*

We establish Proposition 3.15 by proving two results regarding the properties $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$. In Proposition 3.18 we show that all sets $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ fit the framework of Theorem 3.12. Namely, for some constant $n \in \mathbb{N}$ and for every $N \in \mathbb{N}$ there exists a sequence $\overline{w} = (w_H)_{H:V(H)=[n]}$ such an N-vertex graph is in $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ if and only if it is in $\Pi_{n,\overline{w},b}$. Next, in Proposition 3.19 we show that for any constants $0 \leq \alpha < \beta \leq \frac{1}{2}$ such that $\beta - \alpha < \frac{1}{2}$ the set $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ does not have a two-sided error POT.

Writing $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ as $\Pi_{n,\overline{w},b}$. We start by showing for each $N \in \mathbb{N}$ a sequence $\overline{w} = (w_H)_{H:V(H)=[n]}$ such that the set of N-vertex graphs in $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ is equal to the set of N-vertex graphs in $\Pi_{n,\overline{w},b}$.

Proposition 3.18 ($\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ versus $\Pi_{n,\overline{w},b}$): *Fix $0 \leq \alpha \leq \beta \leq \frac{1}{2}$. Then, there exists constants $n \in \mathbb{N}$ and $b \in (0,1)$ such that for every $N \in \mathbb{N}$ there exists a sequence $\overline{w} = (w_H)_{H:V(H)=[n]}$ such that the set of N-vertex graph in $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ equals the set of N-vertex graphs in $\Pi_{n,\overline{w},b}$.*

Actually, n and b are independent of α and β.

Proof: As in the proof of Proposition 3.14, we derive the sequence $\overline{w}$ by considering a weak tester for $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$, and applying the transformation provided in the proof of Theorem 3.12. The tester consists of invoking two tests: (1) an one-sided error POT for $\mathcal{CC}^{\leq 2}$, and (2) a weak tester for edge density (see details below). We rely on the fact that the one-sided error POT of [GR09b] makes three queries (to the subgraph induced by three random vertices), and rejects any N-vertex graph that is not in $\mathcal{CC}^{\leq 2}$ with probability at least $1/\binom{N}{3}$. For (2) we use a weak tester that accepts any N-vertex graph with probability $0.5 \pm N^{-3}$ such that an N-vertex graph G in $\mathcal{CC}^{\leq 2}$ is accepted with probability at least 0.5 if and only if G is in $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$. Specifically, this tester checks the edge density of the graph, accepting each N-vertex graphs with probability $0.5 \pm N^{-3}$ such that a graph is accepted with probability at least 0.5 if and only if the number of its edges lies in the interval $[A, B]$ such that $A = \binom{\lceil \alpha N \rceil}{2} + \binom{N - \lceil \alpha N \rceil}{2}$ and $B = \binom{\lfloor \beta N \rfloor}{2} + \binom{N - \lfloor \beta N \rfloor}{2}$. (Such a tester is obtained by invoking a two-query POT for $D_{\alpha,\beta}$ with probability N^{-3}, and ruling by a fair coin toss otherwise.) Note that the combined weak tester accepts any N-vertex graph in $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ with probability at least 0.5 (since the one-sided POT always accepts such a graph), whereas it accepts N-vertex graphs that are not in $\mathcal{CC}^{\leq 2}$ with probability at most $(1 - 1/\binom{N}{3}) \cdot (0.5 + N^{-3}) < 0.5$ and accepts N-vertex graphs that are in $\mathcal{CC}^{\leq 2} \setminus \mathcal{CC}^{\leq 2}_{\alpha,\beta}$ with probability smaller than 0.5 (since this is what the tester (2) does). ■

Impossibility results for subclasses of $\mathcal{CC}^{\leq 2}$. Next, we prove that for any constants $0 \leq \alpha < \beta \leq \frac{1}{2}$ such that $\beta - \alpha < 0.5$ the set $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ does not have a two-sided error POT. The argument uses the fact (established in Theorem 3.12) that it suffices 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 $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ with respect to some threshold c (i.e., $\mathcal{T}$ accepts G with probability at least c if and only if $G \in \mathcal{CC}^{\leq 2}_{\alpha,\beta}$), then there exist infinitely many graphs G that are $\Omega(1)$-far from $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ such that $\mathcal{T}$ accepts G with probability $c - O(1/|V(G)|^2)$. It follows that $\mathcal{T}$ cannot be a POT.

Proposition 3.19 (sets $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ 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 set $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ does not have a two-sided error POT.*

Proof: We start with a proof sketch, leaving some technical details to later. We assume towards the contradiction that there is a constant query tester $\mathcal{T}$ for $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ with threshold probability c. Then (similarly to the proof of Theorem 3.12), we may assume that, for some constant t, the tester $\mathcal{T}$ reads a subgraph of G induced by a uniformly distributed set of t vertices, and accepts a view $H = ([t], E_H)$ with probability $p_H \in [0,1]$ (i.e., $\mathcal{T}$ accepts with probability p_H when the induced subgraph of G is isomorphic to H). Hence, the probability that $\mathcal{T}$ accepts G can be written as $\sum_H p_H \cdot \rho_H(G)$, where the sum is over all t-vertex graphs, $\rho_H(G)$ denotes the density of H as an induced subgraph of G, and $p_H \in [0,1]$ for each t-vertex graph H.

We first claim that all N-vertex graphs $G \in \mathcal{CC}^{\leq 2}_{\alpha,\beta}$ are accepted with probability that is at most $c + O(N^{-2})$. This follows from the fact that every graph $G \in \mathcal{CC}^{\leq 2}_{\alpha,\beta}$ is $O(N^{-2})$-close to a graph G' that is not in $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$, 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 $\sum_H p_H \cdot \rho_H(G') < c$ (since $G' \notin \mathcal{CC}^{\leq 2}_{\alpha,\beta}$) implies $\sum_H p_H \cdot \rho_H(G) = c + O(N^{-2})$.

In the second step we shall claim that since all graphs $G \in \mathcal{CC}^{\leq 2}_{\alpha,\beta}$ 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 $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$, 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 $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$, 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 $O(N^{-2})$, which implies that $\mathcal{T}$ is not a POT for $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$.

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, since the question of whether H is isomorphic to a subgraph of G_ρ that is induced by a specific t-vertex sample S depends only on the size of the intersection of S with the smallest clique. 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 $\mathtt{P} : [0, \frac{1}{2}] \to \mathbb{R}$. Recall that (by the first step) $\mathtt{P}$ is almost constant on the interval $[\alpha, \beta]$. We claim that this implies that $\mathtt{P}$ is almost constant also on the entire interval $[0, \frac{1}{2}]$, where the "almost" in the conclusion depends on $\deg(\mathtt{P})$, on the ratio between the length of the interval $\beta - \alpha$, and on the "almost" parameter in the hypothesis. Specifically, using the first step by which $\mathtt{P}(\rho)$ is $O(1/N^2)$-close to c for every $\rho \in [\alpha, \beta]$, we infer that $\mathtt{P}(\rho)$ 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.[30] Since, for every $\rho \in [0, \frac{1}{2}]$, we have $\Pr[\mathcal{T} \text{ accepts } G_\rho] = \mathtt{P}(\rho)$ 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 $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$. The detailed argument is given next.

[30] This claim follows almost immediately from Claim 3.19.1 (below), which in turn follows by polynomial interpolation. Specifically, apply Claim 3.19.1 to the polynomial $\mathtt{P}(\cdot) - c$, using $\rho_i = \alpha + (\beta - \alpha) \cdot (i-1)/t$ (for $i = 1, ..., t+1$).

Assume towards contradiction that $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ has a POT $\mathcal{T}$ with threshold probability c. Then, as explained in the proof sketch, we may assume that for some constant t and for every $N \in \mathbb{N}$ there exists a sequence $(p_H)_{H:V(H)=[t]}$ 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

$$\Pr[\mathcal{T} \text{ accepts } G] = \sum_{H:|V(H)|=t} p_H \cdot \rho_H(G), \tag{10}$$

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 any N-vertex graph $G \in \mathcal{CC}^{\leq 2}_{\alpha,\beta}$ we can drop an internal edge of the larger clique to obtain a graph G' that does not belong to $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$. Hence, the graphs G and G' are $\binom{N}{2}^{-1}$-close. Therefore, the densities $\rho_H(G)$ and $\rho_H(G')$ differ by at most $\binom{t}{2}/\binom{N}{2} < (t/N)^2$ for each H. We conclude that $\mathcal{T}$ accepts the graph G and G' with almost the same probability. That is:

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

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

Since any graph $G' \notin \mathcal{CC}^{\leq 2}_{\alpha,\beta}$ must be accepted with probability smaller than c, we conclude that any N-vertex graph $G \in \mathcal{CC}^{\leq 2}_{\alpha,\beta}$ 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:

$$c \leq \Pr[\mathcal{T} \text{ accepts } G] \leq c + rt^2 \cdot N^{-2} \qquad \text{for every } G \in \mathcal{CC}^{\leq 2}_{\alpha,\beta}. \tag{11}$$

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 sketch, since α and β are constants, there is a graph in $\mathcal{CC}^{\leq 2}$ that is $\Omega(1)$-far from $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$. 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 $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$.

In light of the above, we now focus on the behavior of $\mathcal{T}$ only on 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 sketch, 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. (10), the probability that $\mathcal{T}$ accepts the input graph G_ρ is also a polynomial of degree t. Consider a polynomial $\mathtt{P} : [0, \frac{1}{2}] \to \mathbb{R}$ defined as

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

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. (11), the polynomial $\mathtt{P}$ satisfies the following condition:

$$\mathtt{P}(\rho) \in [0, \frac{rt^2}{N^2}] \quad \text{for all } \rho \in [\alpha, \beta] \text{ that satisfy } \rho N \in \mathbb{N}. \tag{12}$$

It follows from the next claim that a polynomial $\mathtt{P}(\rho)$ that satisfies Eq. (12) cannot deviate from zero by more than $O(N^{-2})$ also on the larger interval $[0, \frac{1}{2}]$.

Claim 3.19.1 *Let* $\mathtt{P} : [0, \frac{1}{2}] \to \mathbb{R}$ *be a polynomial of degree at most* t*. Assume that for some* $\epsilon, \delta > 0$ *there are* $t+1$ *points* $\rho_1, \ldots, \rho_{t+1} \in [0, \frac{1}{2}]$ *such that* $|\rho_i - \rho_j| \geq \delta$ *for all* $i \neq j \in [t+1]$*, and* $|\mathtt{P}(\rho_i)| \leq \epsilon$ *for all* $i \in [t+1]$*. Then, for every* $x \in [0, \frac{1}{2}]$*, it holds that* $|\mathtt{P}(x)| < \frac{t+1}{2^t \delta^t} \cdot \epsilon$*.*

The proof of Claim 3.19.1 is presented in Appendix A.2. Using it, we complete the proof of Proposition 3.19 as follows. 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 $\alpha \leq \rho_1 < \rho_2 \ldots < \rho_{t+1} \leq \beta$ that satisfy $\rho_i N \in \mathbb{N}$ for all $i \in [t+1]$ and $|\rho_{i+1} - \rho_i| \geq \delta - \frac{2}{N} > \frac{\delta}{2}$ for all $i \in [d]$.[31] Then, by applying Claim 3.19.1 (and recalling that $\mathtt{P}(\rho) \in [0, \frac{rt^2}{N^2}]$ for all $\rho \in [\alpha, \beta]$ that satisfy $\rho N \in \mathbb{N}$), we conclude that

$$|\mathtt{P}(\rho)| \leq \frac{t+1}{\delta^t} \cdot \frac{rt^2}{N^2} = (t+1) \cdot \left(\frac{t}{\beta - \alpha}\right)^t \cdot rt^2 \cdot N^{-2} \qquad \text{for all } \rho \in [0, \tfrac{1}{2}] \ . \tag{13}$$

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

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

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 $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$, we conclude that $\mathcal{T}$ is not a POT for $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$. ■

Extension to smaller intervals (i.e., $\beta(N) = \alpha(N) + N^{-o(1)}$). The proof of Proposition 3.19 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 3.20 below.

POTs for the case of tiny intervals or tiny complements. As hinted above, *if the length of* $\beta(N) - \alpha(N)$ *is either smaller than* $N^{-\Omega(1)}$ *or larger than* $0.5 - N^{-\Omega(1)}$*, then the corresponding property has a POT*. The following results are proved in our technical report [GS12], and are stated here for the sake of providing a context for the result proved in Proposition 3.19. We mention that both results are proved by using Theorem 5.1. (In both results, the notation $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ is extended to functions $\alpha, \beta : \mathbb{N} \to [0, \frac{1}{2}]$ such that $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ contains the N-vertex graph G if and only if $G \in \mathcal{CC}^{\leq 2}$ and the density of its smaller clique is in the interval $[\alpha(N), \beta(N)]$.)

Proposition 3.20 (POT for $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ when $\beta - \alpha$ is tiny; see [GS12, Prop. 6.8]): *Let* $\alpha, \beta : \mathbb{N} \to [0, \frac{1}{2}]$ *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 set* $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ *has a POT that makes* $O(1/d)$ *queries and has detection probability* $\varrho(\epsilon) = \Omega(\epsilon^{O(1/d)})$*.*

Proposition 3.21 (POT for $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ when $\frac{1}{2} - (\beta - \alpha)$ is tiny; see [GS12, Prop. 6.10]): *Let* $\alpha, \beta : \mathbb{N} \to [0, \frac{1}{2}]$ *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 set* $\mathcal{CC}^{\leq 2}_{\alpha,\beta}$ *has a POT that makes six queries and has detection probability* $\varrho(\epsilon) = \Omega(\epsilon^{O(1/d)})$*.*

[31] This 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 $i \neq j \in [d+1]$. 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 $\frac{2}{N}$.

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 $g : [N] \times [d] \to \{0, 1, ..., N\}$ 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:

Proposition 4.1 (an impossibility result): *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 as a fraction of $dN/2$. (Note that a Boolean function $f : [N] \times [d] \to \{0, 1\}$ can be defined such that $f(v, i) = 1$ if and only if $g(v, i) \in [N]$.)[32] 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 $\overline{\tau} = (\tau_1, ..., \tau_t) \in [0, 1]^t$, such that $\tau_1 \leq \tau_2 < \tau_3 \leq \tau_4 < \cdots \leq \tau_t$, we considered (in Section 2) the set of distributions, denoted $\mathcal{D}_{\overline{\tau}}$, 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 set of bounded-degree graphs will consist of graphs that contain a fraction of isolated vertices that corresponds to a distribution in $\mathcal{D}_{\overline{\tau}}$. That is, $\mathcal{G}_{\overline{\tau}}$ 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 $\mathcal{D}_{\overline{\tau}}$.

Theorem 4.2 (POT for $\mathcal{G}_{\overline{\tau}}$): *For every $\overline{\tau} = (\tau_1, ..., \tau_t)$, the property $\mathcal{G}_{\overline{\tau}}$ 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 $u_1, ..., u_t$, and explore their immediate neighborhood. That is, for each i determine whether or not u_i is isolated in G.

2. Let $j \in \{0, 1, ..., t\}$ denote the number of isolated vertices seen in Step 1. Then, the tester accepts with probability α_j, where $(\alpha_0, \alpha_1, ..., \alpha_t)$ is the sequence of probabilities used by the POT that is guaranteed by Theorem 2.4 (i.e., the tester for $\mathcal{D}_{\overline{\tau}}$).

Let c be the threshold probability associated with the tester of Theorem 2.4. Then, each graph $G \in \mathcal{G}_{\overline{\tau}}$ is accepted with probability at least c. On the other hand, we shall show that if G is ϵ-far from being in $\mathcal{G}_{\overline{\tau}}$, then the fraction of isolated vertices in G is $(\epsilon/4)$-far from $\mathcal{D}_{\overline{\tau}}$, and the theorem follows. Actually, the validity of this claim presupposes that all the thresholds in $\overline{\tau}$ are multiples of $1/N$, and we shall defer this issue to the end of the proof.

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

We shall prove the counterpositive (i.e., if the fraction of isolated vertices in G is δ-close to $\mathcal{D}_{\overline{\tau}}$, then G is 4δ-close to $\mathcal{G}_{\overline{\tau}}$). Suppose that G is an N-vertex graph with M isolated vertices such that there exists $p \in \mathcal{D}_{\overline{\tau}}$ that satisfies $|p - (M/N)| \le \delta$. We shall show how to gradually decrease (resp., increase) the number of isolated vertices, while incurring at most $2d$ edges modifications per each unit of change in the number of isolated vertices.

- If $M > pN \ge 0$ (and $M < N$)[33], then we may decrement the number of isolated vertices by connecting any isolated vertex v to some non-isolated vertex. Specifically, 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 w of degree d and omit one of the current edges of w (while noting that the other end-point of this edge is also of degree d). Indeed, in this case decreasing the number of isolated vertices by one unit incurs at most two edge modifications.
- 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. Specifically, we disconnect v from all its neighbors and connect all neighbors that became isolated (by adding a path through them). Hence, increasing the number of isolated vertices by one unit incurs at most $2d - 1$ edge modifications.

Note that the foregoing argument presupposes that pN is an integer, which is indeed the case when all the thresholds in $\overline{\tau}$ 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 with 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 $p' \in [\tau_{2i-1}, \tau_{2i}] \cap \{j/N : j = 0, 1, ..., N\}$ (while increasing δ by at most $1/N$, which is fine since it suffices to establish an upper bound of $4(\delta + (1/N))$). In the second case, we should actually modify the algorithm and omit the pair (τ_{2i-1}, τ_{2i}) from $\overline{\tau}$, because in this case no N-vertex graph can have τ_{2i-1} isolated vertices (since $\tau_{2i-1}N \notin \mathbb{N}$). That is, the algorithm will refer to a modified $\overline{\tau}$ 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).[34] ■

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 $\overline{\tau}$ as above, we may consider the property $\mathcal{G}_{\mathcal{H},\overline{\tau}}$ that consist of N-vertex graphs G such that the fraction $\#_{\mathcal{H}}(G)/N$ is in $\mathcal{D}_{\overline{\tau}}$. (Indeed, $\mathcal{G}_{\overline{\tau}}$ 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 $\mathcal{G}_{\mathcal{H},\overline{\tau}}$): *For every $\mathcal{H}$ and every $\overline{\tau} = (\tau_1, ..., \tau_t)$ such that $\tau_t \le 1/s(\mathcal{H})$ the property $\mathcal{G}_{\mathcal{H},\overline{\tau}}$ has a two-sided error POT.*

[33] If $M = N$, then $pN \le N - 2$ must hold, since no graph can contain a single non-isolated vertex. In this case, connecting a pair of isolated vertices is fine, since the number of isolated vertices in the resulting graph will not be smaller than pN.

[34] Indeed, this means that the algorithm may use up to $2^{t/2}$ different sequences $\overline{\tau}$, 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).

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 $u_1, ..., u_t$, 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 $(\alpha_0, \alpha_1, ..., \alpha_t)$ is the sequence of probabilities used by the POT that is guaranteed by Theorem 2.4 (i.e., the tester for $\mathcal{D}_{\overline{T}}$).

Let c be the threshold probability associated with the tester of Theorem 2.4. Then, each graph $G \in \mathcal{G}_{\mathcal{H},\overline{T}}$ 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 $\mathcal{G}_{\mathcal{H},\overline{T}}$, then $\#_{\mathcal{H}}(G)/N$ is $\Omega(\epsilon)$-far from $\mathcal{D}_{\overline{T}}$, 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 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 $1 \in \mathcal{D}_{\mathcal{T}}$ 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 $s(\mathcal{H}) = 1$. ■

5 Classes of Non-binary Distributions

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

$$\Delta^{(r)} = \{(q_1, \ldots, q_r) \in [0,1]^r : \sum_{i \in [t]} q_i = 1\}. \tag{14}$$

Similarly, a set 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 class of distributions that have a POT

The following result asserts that a set 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 set. 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 set of distributions $\overline{q} = (q_1, \ldots, q_r)$ 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 $\mathtt{P} : \Delta^{(r)} \to \mathbb{R}$ such that for every distribution $\overline{q} = (q_1, \ldots, q_r) \in \Delta^{(r)}$ it holds*

$$\mathtt{P}(q_1, \ldots, q_r) \geq 0 \Longleftrightarrow \overline{q} \in \Pi. \tag{15}$$

If the total degree of $\mathtt{P}$ is t, then Π has a two-sided error POT $\mathcal{T}_\Pi$ that makes t queries and has polynomial detection probability $\varrho(\epsilon) = \Omega(\epsilon^C)$, where $C < t^{O(r)}$.[35] *Moreover, the acceptance*

[35]The constant in the $\Omega()$ notation depends on $\mathtt{P}$, while the $O()$ notation hides some absolute constant.

probability of $\mathcal{T}_\Pi$ when testing $\overline{q} \in \Delta^{(r)}$ can be written as

$$\Pr[\mathcal{T}_\Pi \textit{ accepts } \overline{q}] = \frac{1}{2} + \delta \cdot \mathtt{P}(q_1, \ldots, q_r) \tag{16}$$

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

Proof: The "POT implies polynomial" direction is proved by using the independence of samples of the given distribution. Consider a POT $\mathcal{T}_\Pi$ for Π, which makes t sampling queries and accepts each distribution in Π with probability at least c. When testing $\overline{q} = (q_1, \ldots, q_r)$, for every possible view $\overline{v} = (v_1, \ldots, v_t) \in [r]^t$, the probability of seeing this view is $\prod_{i=1}^t q_{v_i}$. Denoting by $\alpha_{\overline{v}}$ the probability that the tester accepts the view $\overline{v} = (v_1, \ldots, v_t)$, we have

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

Define a polynomial $\mathtt{P}$ to be

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

Then, by definition of the tester, $\mathtt{P}$ satisfies Eq. (15).

For the other direction (i.e., "polynomial implies POT"), let $\mathtt{P} : \Delta^{(r)} \to \mathbb{R}$ be a polynomial of degree t. We show that the set

$$\Pi = \{(q_1, \ldots, q_r) \in \Delta^{(r)} : \mathtt{P}(q_1, \ldots, q_r) \geq 0\} \tag{17}$$

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

In order to simplify the proof, we shall slightly modify $\mathtt{P}$, while making sure that the modifications of $\mathtt{P}$ does not affect Π in Eq. (17). Specifically, we multiply each monomial of degree $d < t$ (of $\mathtt{P}$) by $(\sum_{i\in[r]} q_i)^{t-d}$. This does not change the value of $\mathtt{P}$ in $\Delta^{(r)}$, and hence does not affect Π.[36] Henceforth we shall assume that $\mathtt{P}$ is a homogeneous polynomial of degree t, and therefore can be written as

$$\mathtt{P}(q_1, \ldots, q_r) = \sum_{\overline{v}\in[r]^t} \alpha_{\overline{v}} \prod_{i=1}^t q_{v_i} \tag{18}$$

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

Assume that Π is non trivial. This implies that not all coefficients $\alpha_{\overline{v}}$ are zeros. Given Eq. (18), we define a POT $\mathcal{T}_\Pi$ for Π as follows. The tester makes t queries to a given distribution, gets t samples, denoted by $\overline{v} = (v_1, \ldots, v_t)$, and accepts with probability

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

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

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

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

and hence, by Eq. (18), the equality above becomes

$$\Pr[\mathcal{T}_\Pi \text{ accepts } \overline{q}] = \frac{1}{2} + \delta \cdot \mathtt{P}(q_1, \ldots, q_r). \tag{19}$$

Next, we analyze the acceptance probability in Eq. (19). If $\overline{q} \in \Pi$, then, by Eq. (17), we have $\mathtt{P}(q_1, \ldots, q_r) \geq 0$, and therefore

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

Assume $\overline{q}$ is ϵ-far from Π. Then, in particular $\overline{q} \notin \Pi$, and hence $\mathtt{P}(q_1, \ldots, q_r) < 0$. Thus, using Eq. (19), we have $\Pr[\mathcal{T}_\Pi \text{ accepts } \overline{q}] < \frac{1}{2}$. In order to prove that $\mathcal{T}_\Pi$ is a POT, we need to show that $\Pr[\mathcal{T}_\Pi \text{ accepts } \overline{q}]$ 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 $\mathtt{P} : \Delta^{(r)} \to \mathbb{R}$ be a polynomial, and let*

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

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

$$\mathrm{dist}(\overline{q}, \Pi) = \inf\{\frac{1}{2} \sum_{i \in [r]} |q_i - p_i| : (p_1, \ldots, p_r) \in \Pi\} > \epsilon.$$

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

By applying Theorem 5.2 on Eq. (19), we conclude that if $\overline{q} \in \Delta^{(r)}$ is ϵ-far from Π, then $\Pr[\mathcal{T}_\Pi \text{ accepts } \overline{q}] < \frac{1}{2} - \Omega(\epsilon^C)$, where $C < \deg(\mathtt{P})^{O(r)}$. 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. In Section 5.2 we use this tool to show that class of distributions that have a POT is closed under taking *disjoint* unions. We also mention that this tool has been used to present positive and negative results for some concrete properties of interest. Specifically, in Section 5.3 of our technical report [GS12] we utilize Theorem 5.1 towards presenting POTs for several concrete properties, whereas in Section 5.4 of [GS12] we use it towards proving that some other properties have no POTs. Finally, we mention that a straightforward application of Theorem 5.1 yields the following.

Corollary 5.3 (testing distance to a target distribution): *For a fixed $r, p \in \mathbb{N}$, $\delta \in [0,1]$ and $v = (v_1, ..., v_r) \in \Delta^{(r)}$, let $\Pi_{p,v,\delta}$ be the set of distributions that are at distance at most δ from v in L_p-norm; that is,*

$$\Pi_{p,v,\delta} = \left\{ (q_1, ..., q_r) \in \Delta^{(r)} : \left(\sum_{i \in [r]} (q_i - v_i)^r \right)^{1/r} \leq \delta \right\}.$$

Then, for any $p \geq 2$, the property $\Pi_{p,v,\delta}$ has a two-sided error POT. On the other hand, for either $p = 1$ or $p = \infty$, there exists δ and v such that $\Pi_{p,v,\delta}$ has no two-sided error POT.

Proof: The case of $p \geq 2$ is obvious, since the property is characterized by a degree p polynomial. On the other hand, for either $p = 1$ or $p = \infty$, the region $\Pi_{p,v,\delta}$ is a polyhedra, and for sufficiently small $\delta > 0$ it is strictly contained in $\Delta^{(r)}$. Clearly, this region cannot be characterized by the non-negative values of a polynomial, because its boundary cannot equal the set of zeros of a polynomial (e.g., its boundary contains a line segment, whereas any polynomial that is zero on a line segment must be zero on the entire line). ■

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 in general: Indeed, in Proposition 5.13 of our technical report [GS12], 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* sets of distributions have two-sided error POTs, then so does their union.

Corollary 5.4 (closure under disjoint union): *Let Π and Π' be two disjoint sets of distributions with domain $[r]$, and suppose that both Π and Π' have a two-sided error POT. Then, 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 $\mathtt{P} : \Delta^{(r)} \to \mathbb{R}$, such that $\Pi = \{\overline{q} \in \Delta^{(r)} : \mathtt{P}(\overline{q}) \geq 0\}$. Similarly, there is a polynomial $\mathtt{P}' : \Delta^{(r)} \to \mathbb{R}$, such that $\Pi' = \{\overline{q} \in \Delta^{(r)} : \mathtt{P}'(\overline{q}) \geq 0\}$. Define a polynomial $\mathtt{P}_{\text{union}} : \Delta^{(r)} \to \mathbb{R}$ to be

$$\mathtt{P}_{\text{union}}(\overline{q}) = -\mathtt{P}(\overline{q}) \cdot \mathtt{P}'(\overline{q}).$$

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

By applying Corollary 5.4 repeatedly, it follows that if for each $i \in [k]$ the set Π_i (of distributions with domain $[r]$) has a two-sided error POT that makes t_i queries then the disjoint union $\Pi = \cup_{i \in [k]} \Pi_i$ has a two-sided error POT that makes $\sum_{i \in [k]} t_i$ queries. Unfortunately, this does not give an explicit bound on the detection probability of the resulting POT. Such a bound is provided by Proposition 5.5 of our technical report [GS12], which asserts that the detection probability of the resulting POT is $\Omega(\min\{\varrho_i : i \in [k]\})$, where ϱ_i is the detection probability of the i^{th} POT.

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 $\Pi = \{\overline{q} \in \Delta^{(r)} : \mathtt{P}(\overline{q}) \geq 0\}$ for some polynomial $\mathtt{P} : \Delta^{(r)} \to \mathbb{R}$, and thus is a closed[37] 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 closure[38] of the complement, denoted by $\text{cl}(\Delta^{(r)} \setminus \Pi)$, has a POT. In Appendix A.3 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.

[37] A 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 $\overline{q} \in B$ there exists an $\epsilon > 0$ such that every $\overline{q'} \in \Delta^{(r)}$ that is at distance at most ϵ from $\overline{q}$ is actually in B.

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

Acknowledgments

We are grateful to Dana Ron for collaboration in early stages of this research. We also wish to thank Johan Håstad and the anonymous referees for their useful comments.

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.

[GS12] O. Goldreich and I. Shinkar. Two-Sided Error Proximity Oblivious Testing. *ECCC*, TR12-021, 2012. (See Revision 4, 2014.)

[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`

[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.

Appendices

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 K, N and $KN/2$ are natural numbers and $\sum_{v\in[N]} |d_G(v) - K| \le \epsilon' \cdot N^2$, 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 $v \in H \stackrel{\rm def}{=} \{u : d_G(u) > K\}$. Since $\sum_{v\in H}(d_G(v) - K) \le \epsilon' N^2$, we obtain a graph G' that is $2\epsilon'$-close to G such that $d_{G'}(v) \le K$ holds for each vertex v, because every omitted edge reduces $\sum_{v\in H}(d_G(v) - K)$ by at least one unit. Furthermore, $\sum_{v\in[N]} |d_{G'}(v) - K| \le \epsilon' \cdot N^2$, 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 $d_{G''}(v) \le K$ 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 $|\Gamma(v)| \le K - 2$ must hold. We shall show that there exists two vertices u and 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 $u \neq v$)[39], 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 $|\Gamma(v_i)| \le K - 1$ 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 $|\Gamma(v_1) \cup \{v_1, v_2\}| \le K < N$. 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 $|\Gamma(v_2) \cup \{v_2\} \setminus \{v_1\}| \le K - 1$). Thus, there exists $u_2 \in \Gamma(u_1) \setminus (\Gamma(v_2) \cup \{v_2\})$.

[39] Recall 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 $d_{G''}(v) \le K$ for all v). Since in each step we modified three edges (and there are at most $\epsilon' N^2/2$ steps), 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''). ■

A.2 Proof of Claim 3.19.1

The following technical claim quantifies the assertion that if a degree t polynomial is close to zero on $t+1$ distinct points, then it is close to zero on a line segment that contains these points and is not much longer than the distance between two closest points.

Claim 3.19.1, restated: *Let* $\mathtt{P} : [0, \frac{1}{2}] \to \mathbb{R}$ *be a polynomial of degree at most* t*, and* $\epsilon, \delta > 0$*. Let* $\rho_1, \ldots, \rho_{t+1} \in [0, \frac{1}{2}]$ *such that* $|\rho_i - \rho_j| \ge \delta$ *for all* $i \ne j \in [t+1]$*, and suppose that* $|\mathtt{P}(\rho_i)| \le \epsilon$ *for each* $i \in [t+1]$*. Then, for every* $x \in [0, \frac{1}{2}]$*, it holds that* $|\mathtt{P}(x)| < \frac{t+1}{2^t \delta^t} \cdot \epsilon$*.*

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

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

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

$$\begin{aligned} |\mathtt{P}(x)| &\le \sum_{i\in[t+1]} \left(\prod_{j\ne i} \frac{|x - \rho_j|}{|\rho_i - \rho_j|} \right) \cdot |\mathtt{P}(\rho_i)| \\ &\le \sum_{i\in[t+1]} \left(\prod_{j\ne i} \frac{1/2}{\delta} \right) \cdot \epsilon \\ &= (t+1) \cdot \left(\frac{1}{2\delta} \right)^t \cdot \epsilon \end{aligned}$$

where the second inequality uses $x, \rho_j \in [0, \frac{1}{2}]$, $|\mathtt{P}(\rho_i)| \le \epsilon$ and $|\rho_i - \rho_j| \ge \delta$ (for every $i \ne j$). ■

A.3 Sets of distributions having a POT are not closed under complement

Following the remark in the end of Section 5.2, we describe a set of ternary distributions $\Pi \subseteq \Delta^{(3)}$ that has a POT, while $\mathrm{cl}(\Delta^{(3)} \setminus \Pi)$ does not have one[40]. We start with the following claim.

Claim A.5 *Let* $\mathcal{D} = \{(x, y) \in [0, 1]^2 : x + y \le 1\}$ *be a subset of* $\mathbb{R}^2$*. For* $\alpha \in (0, 1)$ *let* $A = \{(x, y) \in \mathcal{D} : \mathtt{P}(x, y) \ge 0\}$*, where* $\mathtt{P} : \mathcal{D} \to \mathbb{R}$ *is the polynomial* $\mathtt{P}(x, y) = y^2 - (x - \alpha) \cdot x^2$*. Then, there is no real polynomial* $\mathtt{Q}$ *such that* $\mathrm{cl}(\mathcal{D} \setminus A) = \{(x, y) \in \mathcal{D} : \mathtt{Q}(x, y) \ge 0\}$*, where* $\mathrm{cl}(\mathcal{D} \setminus A)$ *is the closure of the complement*[41] *of* A *in* $\mathcal{D}$*.*

[40] By $\mathrm{cl}(\Delta^{(3)} \setminus \Pi)$ 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 $\frac{1}{2}(|q_1 - q'_1| + |q_2 - q'_2| + |q_3 - q'_3|) < \epsilon$.

[41] By $\mathrm{cl}(\mathcal{D} \setminus A)$ 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 $\frac{1}{2}(|x - x'| + |y - y'|) < \epsilon$.

Proof: Note first that A can be written as

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

where

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

In particular the boundary of A within the interior of $\mathcal{D}$ is $\partial A = \left\{\left(x, x\sqrt{(x-\alpha)}\right) : x \in (\alpha, \beta)\right\}$, where β is the solution to the equation $x + x\sqrt{(x-\alpha)} = 1$. Note that ∂A can also be written as

$$\partial A = \left\{\left(x^2, x^2 \cdot (x-\alpha)\right) : x \in (\sqrt{\alpha}, \sqrt{\beta})\right\}.$$

Assume towards contradiction that there is a polynomial $\mathtt{Q}$ that satisfies the condition stated in the claim, namely $\mathrm{cl}(\mathcal{D} \setminus A) = \{(x,y) \in \mathcal{D} : \mathtt{Q}(x,y) \geq 0\}$. Then, in particular (1) $\mathtt{Q}$ must be zero on ∂A, and (2) for any $(x,y) \in A \setminus \partial A$, it must be the case that $\mathtt{Q}(x,y) < 0$. 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 $\mathtt{Q}$ be a polynomial that vanishes of ∂A. Note that the polynomial $\mathtt{P}$ is irreducible[42], and the two polynomials $\mathtt{P}$ and $\mathtt{Q}$ agree on the curve $\partial A = \left\{\left(x^2, x^2 \cdot (x-\alpha)\right) : x \in (\sqrt{\alpha}, \sqrt{\beta})\right\}$, 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 $\mathtt{R}$, such that $\mathtt{P} = \mathtt{R} \cdot \mathtt{P}'$ and $\mathtt{Q} = \mathtt{R} \cdot \mathtt{Q}'$, for some polynomials $\mathtt{P}'$ and $\mathtt{Q}'$. However, since $\mathtt{P}$ is irreducible, we conclude that $\mathtt{P}'$ is some constant polynomial and $\mathtt{R} = c\mathtt{P}$ for some non-zero constant $c \in \mathbb{R}$, and thus $\mathtt{Q}$ can be written as $\mathtt{Q} = c \cdot \mathtt{P} \cdot \mathtt{Q}'$. Therefore, since $\mathtt{P}$ vanishes at $(0,0)$ it follows that $\mathtt{Q}$ also vanishes and $(0,0)$. The claim follows. ∎

Using Claim A.5 we exhibit a property of ternary distributions Π that has a POT, while $\mathrm{cl}(\Delta^{(3)} \setminus \Pi)$ does not have one.

Proposition A.6 *Let $\alpha \in (0,1)$ and let $\mathtt{P}(x,y) = y^2 - (x-\alpha) \cdot x^2$ be as in Claim A.5. Define $\Pi \subseteq \Delta^{(3)}$ to be*

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

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

Proof: Clearly, by Theorem 5.1, Π has a two-sided error POT. In order to prove that $\Pi' := \mathrm{cl}(\Delta^{(3)} \setminus \Pi)$ does not have a two-sided error POT, it is enough to show that there is no polynomial $\mathtt{P}' : \Delta^{(3)} \to \mathbb{R}$, that satisfies $\Pi' = \{(q_1, q_2, q_3) \in \Delta^{(3)} : \mathtt{P}'(q_1, q_2, q_3) \geq 0\}$, which follows easily from Claim A.5. Details follow.

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

$$\mathtt{Q}(x,y) = \mathtt{P}'(x, y, 1-x-y),$$

where $\mathcal{D} = \{(x,y) \in [0,1]^2 : x + y \leq 1\}$, as in Claim A.5. Note that $(x, y, 1-x-y) \in \Delta^{(3)}$ for all $(x,y) \in \mathcal{D}$, and thus $\mathtt{Q}$ is well defined.

Let $A = \{(x,y) \in \mathcal{D} : \mathtt{P}(x,y) \geq 0\}$. We claim that $\mathrm{cl}(\mathcal{D} \setminus A) = \{(x,y) \in \mathcal{D} : \mathtt{Q}(x,y) \geq 0\}$. By Claim A.5, such polynomial does not exists, thus contradicting the assumption that Π' has a POT.

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

In order to prove that $\mathrm{cl}(\mathcal{D} \setminus A) = \{(x, y) \in \mathcal{D} : \mathtt{Q}(x, y) \geq 0\}$ recall that $\mathtt{Q}(x, y) \geq 0$ if and only if $\mathtt{P}'(x, y, 1 - x - y) \geq 0$, which by the assumption (towards contradiction) on $\mathtt{P}'$ is equivalent to $(x, y, 1 - x - y) \in \Pi' = \mathrm{cl}(\Delta^{(3)} \setminus \Pi)$. This, by definition of $\mathtt{P}$ is equivalent to $\mathtt{P}(x, y) \leq 0$, and hence to the assertion $(x, y) \in \mathrm{cl}(\mathcal{D} \setminus A)$. Therefore, for all $(x, y) \in \mathcal{D}$ we have $\mathtt{Q}(x, y) \geq 0$ if and only if $(x, y) \in \mathrm{cl}(\mathcal{D} \setminus A)$, and the proposition follows. ■