

# 1 Overview

We consider the notion of a *local-characterization* of an infinite family of unlabeled bounded-degree graphs. Such a local-characterization is defined in terms of a finite set of (marked) graphs yielding a generalized notion of subgraph-freeness, which extends the standard notions of induced and non-induced subgraph freeness.<sup>1</sup> Intuitively, this notion corresponds to forbidden neighbourhoods of constant distance; that is, we consider the set of graphs such that each vertex in them has a neighborhood that is not forbidden. In other words, a family of graphs  $\mathcal{G}$  is locally-characterized by a finite set of marked graphs  $\mathcal{F}$  if  $\mathcal{G}$  equals the set of  $\mathcal{F}$ -free graphs (i.e., a graph  $G$  is in  $\mathcal{G}$  if and only if for every  $F \in \mathcal{F}$  the graph  $G$  is  $F$ -free).

We stress that, for every  $\mathcal{F}$  as above, the set of  $\mathcal{F}$ -free graphs is a graph property; that is, it is closed under isomorphism. We also mention that this definition is related both to expressibility by first-order formula (cf. [1]) and to having a proximity-oblivious tester (cf. [5]). In light of this, one would expect that such graph properties would be testable within query complexity that depends only on the proximity parameter; however, as shown by Adler, Kohler, and Peng [1], this is not the case. The latter result is proved by constructing a locally-characterizable graph property that consists solely of expander graphs (and using the fact that such a property cannot be tested within query complexity that depends only on the proximity parameter [2]).

**Theorem 1.1** (a locally-characterizable property of expander graphs): *There exists a finite collection of marked graphs, denoted  $\mathcal{F}$ , such that the set of  $\mathcal{F}$ -free graphs is a set of expanders.*

The proof of Theorem 1.1 is pivoted at the Zig-Zag construction of Reingold, Vadhan, and Wigderson [7]. Recall that they presented a sequence of graphs,  $(G_i)_{i \in \mathbb{N}}$ , such that  $G_1 = H^2$  for some constant-size expander  $H$  (of degree  $d$  and second-eigenvalue  $d/4$ ) and  $G_i = G_{i-1}^2 \circledast H$ , where  $\circledast$  denotes the Zig-Zag product. Furthermore, each vertex in  $G_{i-1}$  is replaced in  $G_i$  by a cloud of vertices of the same size as  $H$ , and edges of  $G_{i-1}^2$  yield “connections” between the corresponding clouds in  $G_i$  (where the edges of the Zig-Zag product correspond to three-step walks on a graph that combine these connections with copies of  $H$  that are “placed” on each cloud). Loosely speaking, for every  $m$ , the proof of Theorem 1.1 (provided in [1]) identifies a graph that consists of the graphs  $G_1, \dots, G_m$  along with edges that connect each vertex of  $G_{i-1}$  to each vertex in the corresponding cloud in  $G_i$ . Additional gadgets are added to perform the Zig-Zag product; in fact, the construction is first presented in terms of a directed edge-colored graph, and gadgets are later used to yield an undirected graphs (with no colors).

The point is that, while it is inconceivable that we can provide a local-characterization of  $G_m$  itself, it is quite conceivable that we can provide a local-characterization of  $G_m$  in terms of  $G_{m-1}$ . Indeed, the simple relation between  $G_{m-1}$  and  $G_m$ , already capitalized by Reingold [6], is pivotal here too. Needless to say, materializing this outline requires a careful implementation.

**Application to property testing.** We say that a property tester is *size-oblivious* if its query complexity depends only on the proximity parameter. The main result of [2] states that every infinite property of bounded-degree graphs that has a size-oblivious tester must contain an infinite

---

<sup>1</sup>This notion, defined in [5], is reviewed in Section 2.2.

hyperfinite subproperty. Recalling that hyperfinite graphs are the “extreme opposite” of expander graphs, it follows that the property asserted in Theorem 1.1 does not have a size-oblivious tester. Hence, we get

**Corollary 1.2** (a locally-characterizable property that is not size-oblivious testable): *There exists a finite collection of marked graphs, denoted  $\mathcal{F}$ , such that the set of  $\mathcal{F}$ -free graphs does not have a size-oblivious tester. In particular, this set has no proximity-oblivious tester.*

Recall that a proximity-oblivious tester is an oracle machine of constant query-complexity that accepts each graph in the property with probability 1, and rejects each graph that is not in the property with probability that is related to the (relative) distance of the graph from the property (cf. [3, Def. 1.7]). Hence, a proximity-oblivious tester for a property yields a size-oblivious tester for that property (cf. [3, Thm. 1.9]).

**Organization.** After recalling the background (Section 2), we turn to the core of this survey (Section 3), where we provide more details on the proof of Theorem 1.1.

## 2 Background

While the contents of Section 2.1 is well-known and may be skipped, the notion of generalized subgraph freeness (reviewed in Section 2.2) is likely to be unfamiliar to most readers. For sake of self-containment, we also review the Zig-Zag construction (in Section 2.3).

### 2.1 Testing in the Bounded-Degree Graph Model

(This model was introduced in [4] and is reviewed in [3, Chap. 9].)

The bounded-degree graph model refers to a fixed (constant) degree bound, denoted  $d \geq 2$ . In this model, a graph  $G = (V, E)$  of maximum degree  $d$  is represented by the incidence function  $g : V \times [d] \rightarrow V \cup \{\perp\}$  such that  $g(v, j) = u \in V$  if  $u$  is the  $j^{\text{th}}$  neighbor of  $v$  and  $g(v, j) = \perp \notin V$  if  $v$  has less than  $j$  neighbors.<sup>2</sup> Distance between graphs is measured in terms of their foregoing representation; that is, as the fraction of (the number of) different array entries (over  $d \cdot |V|$ ).

The tester is given oracle access to the representation of the input graph (i.e., to the incidence function  $g$ ), where for simplicity we assume that  $V = [n]$  for  $n \in \mathbb{N}$ . In addition, the tester is also given a proximity parameter  $\epsilon$  and a size parameter (i.e.,  $n$ ). Recall that **graph properties** are sets of graphs that are closed under isomorphism.

**Definition 2.1** (property testing in the bounded-degree graph model): *For a fixed  $d \in \mathbb{N}$ , a tester for the graph property  $\Pi$  is a probabilistic oracle machine  $T$  that, on input a proximity parameter  $\epsilon > 0$  and size parameter  $n \in \mathbb{N}$ , and when given oracle access to an incidence function  $g : [n] \times [d] \rightarrow [n] \cup \{\perp\}$ , outputs a binary verdict that satisfies the following two conditions:*

1. *The tester accepts each graph  $G = ([n], E)$  in  $\Pi$  with probability at least  $2/3$ ; that is, for every  $g : [n] \times [d] \rightarrow [n] \cup \{\perp\}$  representing a graph in  $\Pi$  (and every  $\epsilon > 0$ ), it holds that  $\Pr[T^g(n, \epsilon) = 1] \geq 2/3$ .*

---

<sup>2</sup>For simplicity, we adopt the standard convention by which the neighbors of  $v$  appear in arbitrary order in the sequence  $(g(v, 1), \dots, g(v, \deg(v)))$ , where  $\deg(v) \stackrel{\text{def}}{=} |\{j \in [d] : g(v, j) \neq \perp\}|$ .

2. Given  $\epsilon > 0$  and oracle access to any graph  $G$  that is  $\epsilon$ -far from  $\Pi$ , the tester rejects with probability at least  $2/3$ ; that is, for every  $g : [n] \times [d] \rightarrow [n] \cup \{\perp\}$  that represents a graph that is  $\epsilon$ -far from  $\Pi$ , it holds that  $\Pr[T^{g'}(n, \epsilon) = 0] \geq 2/3$ , where the graph represented by  $g$  is  $\epsilon$ -far from  $\Pi$  if for every  $g' : [n] \times [d] \rightarrow [n] \cup \{\perp\}$  that represents a graph in  $\Pi$  it holds that  $|\{(v, j) \in V \times [d] : g(v, j) \neq g'(v, j)\}| > \epsilon \cdot dn$ .

The tester is said to have **one-sided error probability** if it always accepts graphs in  $\Pi$ ; that is, for every  $g : [n] \times [d] \rightarrow [n] \cup \{\perp\}$  representing a graph in  $\Pi$  (and every  $\epsilon > 0$ ), it holds that  $\Pr[T^g(n, \epsilon) = 1] = 1$ .

The query complexity of a tester for  $\Pi$  is a function (of the parameters  $d, n$  and  $\epsilon$ ) that represents the number of queries made by the tester on the worst-case  $n$ -vertex graph of maximum degree  $d$ , when given the proximity parameter  $\epsilon$ . Fixing  $d$ , we typically ignore its effect on the complexity (equiv., treat  $d$  as a hidden constant). Our focus here is on cases in which the query complexity depends only on the proximity parameter (i.e., size-oblivious query complexity).

## 2.2 Generalized Subgraph Freeness Properties

The notion of *generalized subgraph freeness*, introduced in [5], is aimed to capture what one can see by exploring a constant-radius neighborhood of a vertex in a graph that has some predetermined graph property. The issue is that some vertices are fully explored (i.e., the explorer sees all their neighbors), whereas for other vertices (at the boundary of the exploration) the explorer may only encounter them but not all their neighbors (since it has not traversed their incident edges). We actually consider the set of subgraphs that the explorer cannot encounter when exploring a graph that has the property, where these forbidden subgraphs are represented by *marked graphs*, which are graphs in which each vertex is marked either **full** or **semi-full** or **partial**. Intuitively, the marking **full** represent a vertex that is not at the boundary of the exploration, which means that all its incident edges were traversed. In contrast, vertices at the boundary are marked as **partial**, whereas the marking **semi-full** is inessential (and is included for sake of greater flexibility (see Footnote 3)).

**Definition 2.2** (marked graphs, embedding, and generalized subgraph freeness): *A marked graphs is a pair consisting of a graph and a marking of its vertices such that each vertex is marked either full or semi-full or partial. We say that a marked graph  $F = ([h], A)$  can be embedded in a graph  $G = ([N], E)$  if there exists a 1-1 mapping  $\phi : [h] \rightarrow [N]$  such that for every  $i \in [h]$  the following two conditions hold:*

1. If  $i$  is marked **full**, then  $\phi$  yields a bijection between the set of neighbors of  $i$  in  $F$  and the set of neighbors of  $\phi(i)$  in  $G$ . That is, in this case  $\Gamma_G(\phi(i)) = \phi(\Gamma_F(i))$ , where  $\Gamma_X(v)$  denotes the set of neighbors of  $v$  in the graph  $X$ , and  $\phi(S) = \{\phi(v) : v \in S\}$ .
2. If  $i$  is marked **semi-full**, then  $\phi$  yields a bijection between the set of neighbors of  $i$  in  $F$  and the set of neighbors of  $\phi(i)$  in the subgraph of  $G$  induced by  $\phi([h])$ . That is, in this case  $\Gamma_G(\phi(i)) \cap \phi([h]) = \phi(\Gamma_F(i))$ .
3. If  $i$  is marked **partial**, then  $\phi$  yields an injection of the set of neighbors of  $i$  in  $F$  to the set of neighbors of  $\phi(i)$  in  $G$ . That is, in this case  $\Gamma_G(\phi(i)) \supseteq \phi(\Gamma_F(i))$ .

The graph  $G$  is called  $F$ -free if  $F$  cannot be embedded in  $G$  (i.e., there is no embedding of  $F$  in  $G$  that satisfies the foregoing conditions). For a set of marked graphs  $\mathcal{F}$ , a graph  $G$  is called  $\mathcal{F}$ -free if for every  $F \in \mathcal{F}$  the graph  $G$  is  $F$ -free.

Indeed, the standard notion of (non-induced) subgraph freeness is a special case of generalized subgraph freeness, obtained by considering the corresponding marked graph in which all vertices are marked **partial**. Similarly, the notion of induced subgraph freeness is a special case of generalized subgraph freeness, obtained by considering the corresponding marked graph in which all vertices are marked **semi-full**.<sup>3</sup>

Marking vertices as **full** introduces a new type of constraint; specifically, this constraint mandates the non-existence of neighbors that are outside the embedding of the marked subgraph. For example, using vertices that are marked **full**, it is possible to disallow certain degrees in the graph (see Example 2.3). Thus, the generalized notion of subgraph freeness includes properties that are not hereditary (e.g., regular graphs), whereas induced and non-induced subgraph freeness are hereditary.

**Example 2.3** (disallowing certain degrees via generalized subgraph freeness): *For every  $d' \in \{0, 1, \dots, d\}$ , we can disallow vertices of degree  $d'$  by using a  $(d' + 1)$ -vertex marked graph with  $d'$  edges in which a single vertex is marked **full**, and is connected to  $d'$  vertices that are each marked **partial**.*

The foregoing example as well as the next one are actually used in the proof of Theorem 1.1. The following example refer to the case that we want to mandate that if the graph contains some fixed subgraph  $H'$  then it actually contains additional edges (i.e.,  $H \setminus H'$ ) on the same vertices.

**Example 2.4** (mandating some subgraph via generalized subgraph freeness): *Let  $H' = ([h], A')$  be a subgraph of  $H = ([h], A)$ , and suppose that we want to enforce that every induced subgraph of  $G$  that contains  $H'$  also contain  $H$ . This can be obtained by requiring  $G$  to be  $\mathcal{F}$ -free, where  $\mathcal{F}$  is the set of all marked  $h$ -vertex graphs that are consistent with  $H'$  but not with  $H$ . Specifically,  $F$  is in  $\mathcal{F}$  if  $F$  is embedded in any  $h$ -vertex graph that contains  $H'$  but not  $H$ .*

For sake of completeness, we present the following definition, which we actually use only in headings.

**Definition 2.5** (locally characterizable properties): *A graph property  $\Pi$  is called locally characterizable if there exists a finite set of marked graphs  $\mathcal{F}$  such that  $\Pi$  equals the set of  $\mathcal{F}$ -free graphs.*

(We mention that Definition [5, Def. 5.2] is more general; it allows a different set of marked graphs to be used for each graph size as long as there is a uniform bound on the size of all marked graphs used.)

### 2.3 The Zig-Zag Product

Given a (big)  $D$ -regular graph  $G = (V, E)$ , and a (small)  $d$ -regular graph  $H = ([D], F)$ , their Zig-Zag product, denoted  $G \otimes H$ , consists of the vertex set  $V \times [D]$ , which is partitioned to  $D$ -vertex clouds

---

<sup>3</sup>Indeed, **semi-full** marking can be avoided by emulating marked graphs by sets of mark graphs that avoid **semi-full** marking. This is analogous to the emulation of induced subgraph freeness by non-induced subgraph freeness.

such that the cloud that corresponds to vertex  $v \in V$  is the set of vertices  $C_v = \{(v, i) : i \in [D]\}$ , and edges that correspond to certain 3-step walks (as detailed next).

Actually, it is instructive to first consider the graph, denoted  $G \circledast H$ , in which copies of  $H$  are placed on the clouds (i.e., for every  $v \in V$  and  $\{i, j\} \in F$  we place the intra-cloud edge  $\{(v, i), (v, j)\}$ ), and edges of  $G$  connect the corresponding clouds by using corresponding edges; that is, if  $\{u, v\} \in E$  is the  $i^{\text{th}}$  (resp.,  $j^{\text{th}}$ ) edge incident at  $u$  (resp., at  $v$ ), then we place the inter-cloud edge  $\{(u, i), (v, j)\}$ . Note that each vertex in  $G \circledast H$  has  $d$  intra-cloud edges and a single inter-cloud edge. Now, the edges of  $G \circledcirc H$  correspond to 3-step walks in  $G \circledast H$  that start with an intra-cloud edge, then take the (only available) inter-cloud edge, and lastly take some intra-cloud edge; that is, such a generic walk has the form  $(v, i) \rightarrow (v, j) \rightarrow (w, k) \rightarrow (w, \ell)$ , where  $\{i, j\}, \{k, \ell\} \in F$  and  $\{(v, j), (w, k)\}$  is an inter-cloud edge in  $G \circledast H$  (i.e.,  $\{v, w\} \in E$  is the  $j^{\text{th}}$  edge incident at  $v$  and the  $k^{\text{th}}$  edge incident at  $w$ ).

We shall assume that both  $G$  and  $H$  are connected and are not bipartite. In that case, it is clear that the graph  $G \circledast H$  is also connected and non-bipartite, and it can be shown that also  $G \circledcirc H$  has these properties. The main technical result of [7] asserts that the convergence rate of a random walk on  $G \circledcirc H$  (a.k.a the relative second eigenvalue of the graph) can be upper-bounded in terms of the convergence rates of random walks on  $G$  and on  $H$ . A simple form of their bound asserts that  $\lambda(G \circledcirc H) \leq \lambda(G) + \lambda(H)$ , where  $\lambda(X)$  denotes the convergence rate of a random walk on the graph  $X$ . Using  $\lambda(H) \leq 1/4$ , it follows that if  $\lambda(G) \leq 1/2$ , then  $\lambda(G \circledcirc H) \leq 1/2$ .

### 3 More on the Proof of Theorem 1.1

We shall first present a construction of a directed graph with edge-colors such that the corresponding underlying graph is an expander. The main part of the construction will be edges that represents the edge-rotation functions of the graphs  $G_1, \dots, G_m$  in the Zig-Zag construction. Recall that the edge-rotation function of a graph extend its adjacency function such that the pair  $(u, \alpha)$  is mapped to the pair  $(v, \beta)$  if the  $\alpha^{\text{th}}$  outgoing edge of  $u$  equals the  $\beta^{\text{th}}$  incoming edge of  $v$ .

Recalling that  $H$  is a  $d$ -regular  $d^4$ -vertex graph and that  $G_1 = H^2$  and  $G_i = G_{i-1} \circledcirc H$  are  $d^2$ -regular  $d^{4i}$ -vertex graphs, for every  $\alpha, \beta \in [d^2]$ , we consider the edge set  $E_{\alpha, \beta}$  such that  $(u, v) \in E_{\alpha, \beta}$  if for some  $i$  there exists an edge in  $G_i$  that connects the  $\alpha^{\text{th}}$  port of vertex  $u$  to the  $\beta^{\text{th}}$  port of vertex  $v$ . Indeed,  $E_{\alpha, \beta}$  is viewed as a set of directed edges that are colored  $(\alpha, \beta)$ , and we postulate that  $(u, v) \in E_{\alpha, \beta}$  if and only if  $(v, u) \in E_{\beta, \alpha}$ . Letting  $E \stackrel{\text{def}}{=} \bigcup_{\alpha, \beta \in [d^2]} E_{\alpha, \beta}$ , we refer to  $(u, v) \in E$  as an  $E$ -edge (and to  $(u, v) \in E_{\alpha, \beta}$  as an  $E_{\alpha, \beta}$ -edge). We stress that the foregoing anti-parallel postulate is a very minimal one and far more substantial conditions will be postulated about the  $E$ -edges by using also other edges which will induce a layered directed acyclic graph (with  $G_i$  identified with the  $i^{\text{th}}$  layer). Indeed, the actual structure of the graphs  $G_1, \dots, G_m$  will be enforced by relating each  $G_i$  to  $G_{i-1}$ .

As a warm-up, suppose that we want to augment the graph with auxiliary (colored) edges that will capture 2-step walks on the original graph. In such a case, we introduce, for every  $\alpha, \beta, \gamma, \delta \in [d^2]$ , an edge set  $E'_{(\alpha, \gamma), (\beta, \delta)}$  such that  $(u, w) \in E'_{(\alpha, \gamma), (\beta, \delta)}$  if and only if there exists  $v$  such that  $(u, v) \in E_{\alpha, \beta}$  and  $(v, w) \in E_{\gamma, \delta}$ . (The point is that the latter is a local condition about the edge sets  $E'_{(\alpha, \gamma), (\beta, \delta)}$ ,  $E_{\alpha, \beta}$  and  $E_{\gamma, \delta}$ ; however, we will actually use  $E'$  only as a shorthand.)

As stated above, the structure of the graphs  $G_1, \dots, G_m$  is enforced by relating each  $G_i$  to  $G_{i-1}$ , where for sake of simplicity we define  $G_0$  to be the graph consisting of a single vertex. The first step in enforcing this relation is the association of vertices in  $G_{i-1}$  with clouds of vertices in  $G_i$  such

that each cloud contains  $d^4$  vertices that are identified (equiv., ordered) so that the  $d^4$  ports of each vertex in  $G_{i-1}^2$  are associated with distinct vertices of the corresponding cloud. This association is enforced by using edges that are directed from each vertex of  $G_{i-1}$  to the corresponding cloud of  $G_i$  such that these  $d^4$  edges are assigned different colors. Specifically, for each  $\sigma \in [d^4]$ , we introduce a set of directed edges, denoted  $P_\sigma$ , and postulate that each vertex has at most one outgoing  $P_\sigma$ -edge and at most one incoming  $P$ -edge, where  $P \stackrel{\text{def}}{=} \bigcup_{\sigma \in [d^4]} P_\sigma$ . Indeed,  $(u, v) \in P_\sigma$  implies that  $v$  is the  $\sigma^{\text{th}}$  vertex in the cloud associated with  $u$ , where  $u$  is the “parent” of  $v$  in the directed tree induced by  $P$ . Additional postulates are added to identify the vertices of  $G_0$  and  $G_m$ ; specifically:

1. We postulate that there exists a single vertex with no incoming  $P$ -edges; the graph  $G_0$  will consist of this vertex.  
 Actually, here we postulate that there exists at most one vertex with no incoming  $P$ -edges; the existence of such a vertex will follow from the other postulates (including those that refer to  $E$ -edges).<sup>4</sup>
2. We postulate that vertices either have no outgoing  $P$ -edges or have at least  $d^4$  outgoing  $P$ -edges. Combined with the first postulate, this implies that the  $P$ -edges form a  $d^4$ -ary directed tree such that all leaves are at the same distance from the root.
3. Intuitively, we postulate that all vertices that have no outgoing  $P$ -edges belong to the same  $G_i$ , and that  $i = m$ . Actually, we postulate that vertices that are connected by  $E$ -edges have the same number of outgoing  $P$ -edges (or rather the same number of outgoing  $P_\sigma$ -edges, for every  $\sigma \in [d^4]$ ). The fact that vertices with no outgoing  $P$ -edges are in  $G_m$  follows from the first postulate (i.e., for  $i \geq 1$ , the graph  $G_i$  cannot contain vertices with no incoming  $P$ -edges).

The aforementioned postulates as well as subsequent ones are enforced by forbidden neighborhoods of constant distance. Indeed, the forbidden neighborhoods correspond to directed and edge-colored versions of marked graphs, which are defined analogously to the Definition 2.2.

Next, we postulate that the  $E$ -edges between the  $d^4$  vertices that neighbor the single vertex of  $P$ -indegree 0 form a copy of  $H^2$ . Recalling that these vertices are identified by their incoming  $P$ -edges, we postulate that  $(u, v) \in E$  if and only if there exist  $\sigma, \tau \in [d^4]$  such that  $u$  (resp.,  $v$ ) has an incoming  $P_\sigma$ -edge (resp.,  $P_\tau$ -edge) from  $G_0$  and  $\{\sigma, \tau\}$  is an edge in  $H^2$ . Furthermore, in this case  $(u, v) \in E_{\alpha, \beta}$  if and only if the foregoing edge in  $H^2$  uses the  $\alpha^{\text{th}}$  port of  $u$  and the  $\beta^{\text{th}}$  port of  $v$ .

The main issue is relating the  $E$ -edges of  $G_i$  to those of  $G_{i-1}$ , for  $i > 1$ . We stress that  $i$  itself cannot and is not referred to in this enforcement. Instead, we refer to any  $(x, y) \in E$  such that  $x$  and  $y$  have outgoing  $P$ -edges and introduce conditions on the opposite endpoints of these  $P$ -edges; that is, we mandate  $E$ -edges among  $d$  of the  $P$ -neighbors of  $x$  (which reside in the cloud that replaces  $x$ ) and  $d$  vertices of the  $P$ -neighbors of  $y$  (which reside in the cloud that replaces  $y$ ). Specifically, for  $(j, k) \in [d]^2 \equiv [d^2]$ , we postulate that  $(u, v) \in E_{(j, k), (k, j)}$  if and only if there exist  $\sigma, \tau, \sigma', \tau' \in [d^4]$  and  $(u', v') \in E'_{\sigma', \tau'}$  such that

1.  $\{\sigma, \sigma'\}$  is an edge colored  $j$  in  $H$ .
2.  $\{\tau, \tau'\}$  is an edge colored  $k$  in  $H$ .

---

<sup>4</sup>We warn that this deduction is the most complex part of the proof of [1, Thm. 3.1].

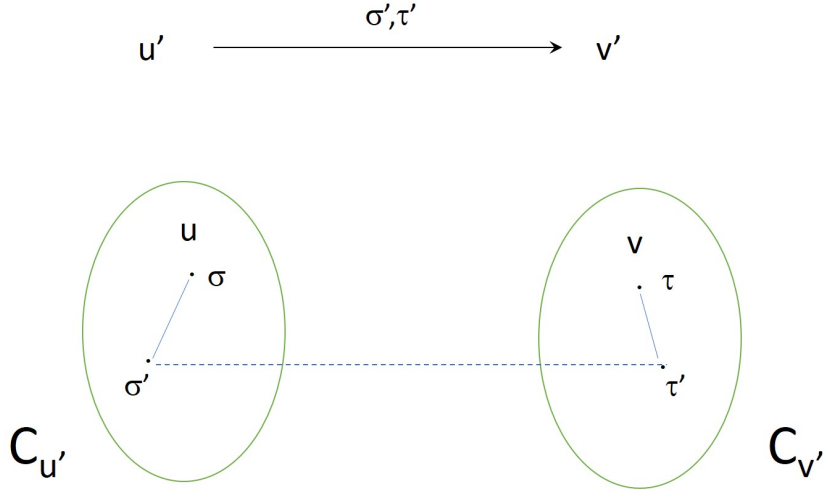


Figure 1: Vertex  $u$  (resp.,  $v$ ) is the  $\sigma^{\text{th}}$  (resp.,  $\tau^{\text{th}}$ ) vertex in the cloud  $C_{u'}$  (resp.,  $C_{v'}$ ) that replaces  $u'$  (resp.,  $v'$ ); these clouds are connected by an edge colored  $\sigma', \tau'$ .

3.  $(u', u) \in P_\sigma$  and  $(v', v) \in P_\tau$ .

(See Figure 1.) Intuitively, these conditions imply that, for some  $i$ , the  $u'$  and  $v'$  are connected in  $G_{i-1}^2$ , whereas  $u$  and  $v$  are vertices in the corresponding clouds of  $G_i$ . Furthermore,  $u$  (resp.,  $v$ ) is associated with the  $\sigma^{\text{th}}$  (resp.,  $\tau^{\text{th}}$ ) vertex of  $H$ , which in turn neighbor vertex  $\sigma'$  (resp.,  $\tau'$ ) of  $H$ . Moreover, vertices  $u'$  and  $v'$  are connected in  $G_{i-1}$  by a 2-path that uses the port  $\alpha \in [d^2]$  of  $u'$  and the port  $\beta \in [d^2]$  of the intermediate vertex such that  $\sigma' = \alpha\beta$  (whereas concatenating the ports used in the opposite direction yields  $\tau'$ ). Indeed, the 2-paths referred to here are the edges of  $E' \stackrel{\text{def}}{=} \bigcup_{\sigma', \tau' \in [d^4]} E'_{\sigma', \tau'}$ , which were defined in the warm-up.

The foregoing description suggests that the  $P$ -edges of a graph that satisfies the listed postulates form a  $d^4$ -ary directed tree such that all leaves are at the same distance from the root, and that the subgraph (of  $E$ -edges) induced by the vertices that are at distance  $i$  from the root equals  $G_i$ . This is indeed the case, but proving the former fact (which refers to the  $P$ -edges) requires using also the postulates that refer to the  $E$ -edges. This is core of the analysis provided in [1, Sec. 3.1]. Hence, we have

**Claim 3.1** (main claim): *For  $n = \sum_{i=0}^m d^{4i}$ , an  $n$ -vertex (unlabeled edge-colored) directed graph satisfies the foregoing conditions if and only if it consists of the graphs  $G_0, G_1, \dots, G_m$  (such that  $G_1 = H^2$  and  $G_i = G_{i-1}^2 \otimes H$ ) that are connected by  $P$ -edges as postulated above.*

Note that if  $n > 1$  is not of the foregoing form, then no  $n$ -vertex graph satisfies these conditions.

By construction, the foregoing  $n$ -vertex graph has constant degree. It is also instructive to observe that this graph is an expander, by using the combinatorial notion of expansion. This is the case because each of the  $G_i$ 's is an expander, whereas each vertex in  $G_{i-1}$  is connected to  $d^4$  different vertices in  $G_i$ . Hence, for every set of vertices  $S$  and each  $i$ , letting  $S_i$  denote the vertices of  $S$  that reside in  $G_i$ , either  $S_{i-1}$  contributes to the expansion inside  $G_{i-1}$  or  $S_{i-1}$  neighbors many vertices in  $S_i$  or  $|S_i| \gg |S_{i-1}|$ .

Lastly, we observe that the foregoing construction can be converted to the context of (simple) undirected graphs (with no edge-colors). This is done by replacing each color class and direction by a different asymmetric gadget such that the gadgets are non-isomorphic and their vertices can be distinguished from the original ones. In particular, we may use gadgets that contain vertices of higher degree than the degree of the original vertices. This yields a corresponding finite set of marked graphs, denoted  $\mathcal{F}$ , that satisfies the following

**Proposition 3.2** (a locally-characterizable property of expander graphs): *The set of  $\mathcal{F}$ -free graphs is an infinite set of expander graphs. Furthermore, this set contains a single unlabeled  $\Theta(d^{Am})$ -vertex graph for every  $m \in \mathbb{N}$ .*

We mention that, by using additional constraints, one can force these expanders to be regular graphs. In fact, this is done in [1, Sec. 3].

## References

- [1] Isolde Adler, Noleen Kohler and Pan Peng. On Testability of First-Order Properties in Bounded-Degree Graphs and Connections to Proximity-Oblivious Testing. In [arXiv:2304.03810](https://arxiv.org/abs/2304.03810) [cs.LO]. Combines preliminary versions that appeared in *SODA21* and *CCC21*.
- [2] Hendrik Fichtenberger, Pan Peng, and Christian Sohler. Every testable (infinite) property of bounded-degree graphs contains an infinite hyperfinite subproperty. In *30th SODA*, pages 714–726, 2019.
- [3] Oded Goldreich. *Introduction to Property Testing*. Cambridge University Press, 2017.
- [4] Oded Goldreich and Dana Ron. Property Testing in Bounded Degree Graphs. *Algorithmica*, Vol. 32 (2), pages 302–343, 2002.
- [5] Oded Goldreich and Dana Ron. On Proximity Oblivious Testing. *SIAM Journal on Computing*, Vol. 40, No. 2, pages 534–566, 2011.
- [6] Omer Reingold. Undirected ST-Connectivity in Log-Space. In *Journal of the ACM*, Vol. 55 (4), pages 17:1–17:24, 2008.
- [7] Omer Reingold, Salil Vadhan, and Avi Wigderson. Entropy Waves, the Zig-Zag Graph Product, and New Constant-Degree Expanders and Extractors. *Annals of Mathematics*, Vol. 155 (1), pages 157–187, 2001. Preliminary version in *41st FOCS*, pages 3–13, 2000.