Sketching Distances in Monotone Graph Classes

Louis Esperet
Univ. Grenoble Alpes, CNRS, Laboratoire G-SCOP, Grenoble, France

Nathaniel Harms
University of Waterloo, Canada

Andrey Kupavskii
Univ. Grenoble Alpes, CNRS, Laboratoire G-SCOP, Grenoble, France
Moscow Institute of Physics and Technology, Russia
Huawei R&D Moscow, Russia

Abstract

We study the problems of adjacency sketching, small-distance sketching, and approximate distance threshold (ADT) sketching for monotone classes of graphs. The algorithmic problem is to assign random sketches to the vertices of any graph G in the class, so that adjacency, exact distance thresholds, or approximate distance thresholds of two vertices u, v can be decided (with probability at least $2/3$) from the sketches of u and v, by a decoder that does not know the graph. The goal is to determine when sketches of *constant size* exist.

Our main results are that, for monotone classes of graphs: constant-size adjacency sketches exist if and only if the class has bounded arboricity; constant-size small-distance sketches exist if and only if the class has bounded expansion; constant-size ADT sketches imply that the class has bounded expansion; any class of constant expansion (i.e. any proper minor closed class) has a constant-size ADT sketch; and a class may have arbitrarily small expansion without admitting a constant-size ADT sketch.

2012 ACM Subject Classification Theory of computation → Randomness, geometry and discrete structures; Theory of computation → Graph algorithms analysis; Theory of computation → Distributed algorithms; Theory of computation → Sketching and sampling

Keywords and phrases adjacency labelling, informative labelling, distance sketching, adjacency sketching, communication complexity

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2022.18

Category RANDOM

Related Version *Full Version*: https://arxiv.org/abs/2202.09253

Funding *Louis Esperet*: supported by the French ANR Projects GATO (ANR-16-CE40-0009-01), GrR (ANR-18-CE40-0032), TWIN-WIDTH (ANR-21-CE48-0014-01), and by LabEx PERSYVAL-lab (ANR-11-LABX-0025).

Acknowledgements We thank Gwenaël Joret for many helpful discussions. We thank Viktor Zamaraev for leading us to Corollary 3.6 and carefully proofreading our manuscript. We thank Alexandr Andoni for a helpful discussion and for sharing with us the manuscript [10]. We thank Renato Ferreira Pinto Jr. and Sebastian Wild for comments on the presentation of this article.

1 Introduction

A common type of problem, with many theoretical and practical uses in computer science, is to assign short *labels* to each of n elements of a space, so that certain "local" information can be deduced from the labels. The Boolean hypercube graph of size $n = 2^d$, with vertex set $\{0,1\}^d$ and edges (x, y) where $x, y \in \{0,1\}^d$ differ on exactly 1 coordinate, has the trivial but useful property that one can assign to each vertex $x \in \{0,1\}^d$ a label of $d = \log n$ bits, so that any function $f(x, y)$ on vertex pairs can be decided given only the labels of x and y. Sometimes, the label size can be improved drastically by allowing *randomized* labels, which we refer to as *sketches*. For example:

1. Adjacency in the hypercube can be computed (with probability at least 2/3) from sketches of constant size (which follows from the Hamming distance communication protocol [42]);
2. Distinguishing between $\mathsf{dist}(x, y) \leq r$ and $\mathsf{dist}(x, y) > r$ can be done with sketches of size depending only on r (which also follows from the Hamming distance protocol);
3. Distinguishing between $\mathsf{dist}(x, y) \leq r$ and $\mathsf{dist}(x, y) > \alpha r$ (for constant $\alpha > 1$) can be done with sketches of size independent of r and n [51].

We call these *adjacency sketches*, *small-distance sketches*, and *approximate distance threshold (ADT) sketches*, respectively (see Section 1.2 for formal definitions). It is natural to ask which classes of graphs, other than the hypercubes, admit similarly efficient sketches. Motivated by a connection between communication complexity, sketching, and graph labelling schemes, recent work [40] asked which *hereditary* classes of graphs admit constant-size adjacency sketches, and also gave some examples of constant-size (i.e. independent of the number of vertices) small-distance sketches, including for planar graphs, answering a question of [38]. Sketches for deciding $\mathsf{dist}(x, y) \leq r$ vs. $\mathsf{dist}(x, y) > \alpha r$ are well-studied, and characterizing the metrics which admit this type of sketch is a well-known open problem [62, 10, 46, 61], but little is known about the natural case of path-distance metrics in graphs.

We study the relationships between these three types of sketches for the important special case of *monotone* classes of graphs. A class of graphs is a set of (labelled[1]) graphs closed under isomorphism. It is *hereditary* if it is closed under taking induced subgraphs, and *monotone* if it is closed under taking subgraphs. Monotone graph classes are ubiquitous: typical examples include minor-closed classes, graphs avoiding some subgraph H, or graphs with bounded chromatic number.

In this paper, we completely determine the monotone graph classes which admit constant-size adjacency sketches and constant-size (i.e. independent of the number of vertices) small-distance sketches, and show that constant-size (i.e. independent of the number of vertices and the parameter r) ADT sketches imply the existence of constant-size small-distance sketches. We show that the classes which admit constant-size adjacency sketches are exactly the classes with bounded arboricity, and the classes which admit constant-size small-distance sketches are exactly the classes with bounded *expansion*[2]. Classes which admit constant-size ADT sketches must also have bounded expansion, and any class with constant expansion (i.e. any proper minor-closed class) has a constant-size ADT sketch, but on the other hand a class can have expansion growing arbitrarily slowly and yet does *not* admit a constant-size ADT sketch. We describe these results in more detail below.

1.1 Motivation & Prior Work

Labelling schemes and sketches are important primitives for distributed computing, streaming, communication, data structures for approximate nearest neighbors, and even classical algorithms (see e.g. [47, 30, 63, 60, 20], and [4, 44, 10, 61, 11] and references therein). As such, a great deal of research has been done on finding other spaces having nice sketching and labelling properties.

---

[1] Standard terminology is that a *labelled* n-vertex graph is one with vertex set $[n]$; not to be confused with *informative labelling schemes*.

[2] We mean bounded expansion in the sense of sparsity theory [57], which is distinct from expansion in the context of expander graphs.

One direction of research investigates the metric spaces which admit approximate distance threshold (ADT) sketches, of the third type described above, as defined in [62]. This is a well-known open problem in sublinear algorithms (see e.g. [10, 46, 61]). Here, n points $X \subseteq \mathcal{X}$ in a metric space $(\mathcal{X}, \mathsf{dist})$, should be assigned random sketches $\mathsf{sk} : X \to \{0,1\}^*$ such that $\mathsf{dist}(x,y) \leq r$ or $\mathsf{dist}(x,y) \geq \alpha r$ can be determined (with probability at least 2/3) from $\mathsf{sk}(x)$ and $\mathsf{sk}(y)$. The goal is to obtain sketches whose size depends only on α. This problem is fairly well-understood when the metric is a norm: there is a constant-size sketch for the ℓ_p (quasi-)norm, for any $0 < p \leq 2$ [44], so any metric that can be embedded into such an ℓ_p is sketchable; conversely, sketching a norm is equivalent to embedding it into $\ell_{1-\varepsilon}$ [11]. Outside of norms, the problem is less well-understood: there are sketchable metrics that are not embeddable into $\ell_{1-\varepsilon}$ [48].

Another direction of research investigates the classes $\mathcal{F}$ of graphs that admit (deterministic) labelling schemes for various functions, generally called *informative labelling schemes* [60]. The most well-studied labelling schemes are for adjacency, introduced in [47, 55]. The main open problem is to identify the hereditary classes of graphs that admit adjacency labelling schemes of size $O(\log n)$. A solution was suggested in [47] and later conjectured in [63], but recently refuted in a breakthrough of [41], leaving the problem wide open. *Randomized* adjacency labelling (i.e. *adjacency sketching*) was studied in [24, 38, 40]. It was observed in [38, 40] that a constant-size sketch implies an $O(\log n)$ labelling scheme, as desired in the above open problem, and it was further observed in [40] that the set of hereditary graph classes which admit constant-size adjacency sketches is equivalent to the set of Boolean-valued communication problems that admit constant-cost public-coin protocols, whose structure is unknown [37]. This raises the following question, which was the main motivation of [40]:

▶ **Question 1.** *Which hereditary classes of graphs admit constant-size adjacency sketches?*

Perhaps the next most commonly studied graph labelling problem is *distance labelling* [31], where the goal is to compute $\mathsf{dist}(x,y)$ from the labels (see e.g. [7, 9, 25, 32]). Intermediate between distance and adjacency labelling is the decision version of distance labelling: for given r, decide whether $\mathsf{dist}(x,y) \leq r$ from the labels. We call this *small-distance labelling*, following the terminology of [6, 29]. For $r = 1$, this coincides with adjacency labelling. The natural generalization of constant-size adjacency sketches is to ask for small-distance sketches whose size depends only on r; it was shown in [38] that such sketches exist for trees, and in [40] that they exist for any Cartesian product graphs and any *stable*[3] class of bounded twin-width (including, for example, planar graphs or any proper minor-closed class; see [27]).

▶ **Question 2.** *Which hereditary classes of graphs admit small-distance sketches whose size depends only on r?*

It is common to weaken distance labelling to *approximate* distance labelling [28], where the goal is to approximate $\mathsf{dist}(x,y)$ up to a constant factor (see e.g. [65, 1, 8]). The decision version is to distinguish, for a given r, between $\mathsf{dist}(x,y) \leq r$ and $\mathsf{dist}(x,y) > \alpha r$; we will call this problem *α-approximate distance threshold (ADT)* labelling and sketching. This is a similar formulation as the distance sketching problem mentioned above, with the n points from the metric space $\mathcal{X}$ being replaced with a size n graph from a class $\mathcal{F}$. Despite significant interest in distance sketching and labelling, the only prior work explicitly relating the two, or studying *randomized* ADT labelling, appears to be the unpublished manuscript [10] (although there is extensive literature on the related problem of embedding graph metrics into normed

[3] See [40] for a discussion of stability, which is not necessary for the current paper.

spaces [54, Chapter 15]; embedding planar graphs into ℓ_1 with constant distortion is a major open problem [35]). This raises the following question, which is a special case of the open problem of identifying sketchable metrics:

▶ **Question 3.** *Which classes of graphs admit constant-size ADT sketches?*

It holds by definition (see definitions below) that a small-distance sketchable class $\mathcal{F}$ is adjacency sketchable, but the relationships between other types of sketching are otherwise unclear, *a priori*. It seems reasonable to suspect that these three types of sketching require similar conditions on the graph class $\mathcal{F}$; so we ask:

▶ **Question 4.** *What is the relationship between adjacency, small-distance, and ADT sketching?*

1.2 Our Results

In this paper, we resolve Questions 1, 2, and 4 for *monotone* classes of graphs, and make progress towards Question 3. The sketches we obtain usually do not assume that the classes under consideration are monotone, but our lower bounds crucially rely on this assumption. We first formally define the main three types of *sketchability* that we are concerned with. We will generalize these definitions in Section 2.2. For a graph class $\mathcal{F}$, we say:

1. $\mathcal{F}$ admits an *adjacency sketch of size* $s(n)$ if there is a function $D : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}$ such that $\forall G \in \mathcal{F}$ with size n, there is a random function $\mathsf{sk} : V(G) \to \{0,1\}^{s(n)}$ satisfying
$$\forall x, y \in V(G) : \qquad \Pr[D(\mathsf{sk}(x), \mathsf{sk}(y)) = 1 \iff x, y \text{ are adjacent}] \geq 2/3 \,.$$
$\mathcal{F}$ is *adjacency sketchable* if it admits an adjacency sketch of constant size.
2. $\mathcal{F}$ admits a *small-distance sketch of size* $s(n, r)$ if for every $r \in \mathbb{N}$ there is a function $D_r : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}$ such that $\forall G \in \mathcal{F}$ with size n, there is a random function $\mathsf{sk} : V(G) \to \{0,1\}^{s(n,r)}$ satisfying
$$\forall x, y \in V(G) : \qquad \Pr[D_r(\mathsf{sk}(x), \mathsf{sk}(y)) = 1 \iff \mathsf{dist}_G(x, y) \leq r] \geq 2/3 \,.$$
$\mathcal{F}$ is *small-distance sketchable* if it admits a small-distance sketch of size independent of n.
3. For constant $\alpha > 1$, $\mathcal{F}$ admits an *α-ADT sketch of size* $s(n)$ if for every $r \in \mathbb{N}$ there is a function $D_r : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}$ such that $\forall G \in \mathcal{F}$ with size n, there is a random function $\mathsf{sk} : V(G) \to \{0,1\}^{s(n)}$ satisfying
$$\forall x, y \in V(G) : \qquad \mathsf{dist}(x, y) \leq r \implies \Pr[D_r(\mathsf{sk}(x), \mathsf{sk}(y)) = 1] \geq 2/3$$
$$\mathsf{dist}(x, y) > \alpha r \implies \Pr[D_r(\mathsf{sk}(x), \mathsf{sk}(y)) = 0] \geq 2/3 \,.$$
For a constant $\alpha > 1$, we say that $\mathcal{F}$ is *α-ADT sketchable* if $\mathcal{F}$ admits an α-ADT sketch with size independent of n. $\mathcal{F}$ is *ADT sketchable* if there is a constant $\alpha > 1$ such that $\mathcal{F}$ is α-ADT sketchable.

Our results imply the following hierarchy, which answers Question 4 for monotone classes of graphs. Let ADJ be the adjacency sketchable monotone graph classes, SD the small-distance sketchable monotone graph classes, and ADT the ADT sketchable monotone graph classes. Then

$$\mathsf{ADT} \subsetneq \mathsf{SD} \subsetneq \mathsf{ADJ} \,.$$

That $\mathsf{SD} \subseteq \mathsf{ADJ}$ follows by definition, and $\mathsf{SD} \neq \mathsf{ADJ}$ is witnessed by the arboricity-2 graphs (as observed in [38]). Our contribution to this hierarchy is $\mathsf{ADT} \subsetneq \mathsf{SD}$ (which does not necessarily hold for non-monotone classes), a complete characterization of the sets SD and ADJ, and some results towards a characterization of ADT.

1.2.1 Adjacency Sketching

We resolve Question 1 for monotone classes by showing that they are adjacency sketchable if and only if they have bounded arboricity. A graph G has arboricity k if its edges can be partitioned into k forests. A class $\mathcal{F}$ has arboricity k if all graphs $G \in \mathcal{F}$ have arboricity at most k. If there exists some constant k such that $\mathcal{F}$ has arboricity k, we say $\mathcal{F}$ has *bounded arboricity.*

▶ **Theorem 1.1.** *Let $\mathcal{F}$ be a monotone class of graphs. Then $\mathcal{F}$ is adjacency sketchable if and only if $\mathcal{F}$ has bounded arboricity.*

All proofs for adjacency sketching are in Section 3. Using standard random hashing and the adjacency labelling scheme of [47], it is easy to see that any class of bounded arboricity is adjacency sketchable; this was stated explicitly in [38, 40] (the latter giving slightly improved sketch size). We prove the converse for monotone classes (which does not hold for hereditary classes in general [40]). We use the probabilistic method to find a subgraph of small discrepancy in any class of unbounded arboricity, inspired by the recent proof of [36, 37] that refuted the main conjecture of [40], and we find that the subgraphs of the hypercube are an easier-to-define counterexample to the conjecture of [40]).

It is interesting that the hashing-based sketch uses randomization only to compute Equality subproblems; i.e. it can be simulated by a constant-cost *deterministic* communication protocol with access to a unit-cost Equality oracle. This type of sketch is called *equality-based* in [40]. Equality-based sketches imply some structural properties of the graph class, such as the strong Erdős-Hajnal property [37]. Recent work has studied the power of the Equality oracle and found that it does not capture the full power of randomization [15, 37, 40]; in particular, the Boolean hypercubes (and any Cartesian product graphs) are adjacency sketchable, but not with an equality-based sketch [37, 40]. Our result shows that Equality captures the power of randomization for sketching monotone classes of graphs. In fact, it is only necessary to compute a *disjunction* of equality checks, which we think of as the simplest possible type of sketch.

We remark that sketches (especially small-distance or ADT sketches) which compute a disjunction of equality checks can be used to obtain *locality-sensitive hashes*, a widely-used algorithmic tool introduced in [45]. Almost all of our positive results are of this type. See Remark 2.6.

1.2.2 Small-Distance Sketching

We answer Question 2 by proving that the monotone graph classes that are small-distance sketchable are exactly those with *bounded expansion* (as in [57]; see our Definition 2.2). Informally, bounded expansion means that the edge density of a graph increases only as a function of r when contracting subgraphs of radius r into a single vertex. Many graph classes of theoretical and practical importance have bounded expansion, including bounded-degree graphs, proper minor-closed graph classes, and graphs of bounded genus [57], along with many random graph models and real-world graphs [17].

To state our theorem, we briefly describe another type of sketch that generalizes small-distance sketching, called *first-order* sketching. A graph class $\mathcal{F}$ is *first-order sketchable* if any first-order (FO) formula $\phi(x, y)$ over the vertices and edge relation of the graph (with two free variables whose domain is the set of vertices) is sketchable (see Section 2.2). This type of sketch was introduced in [40] and generalizes small-distance sketching, along with (for example) testing whether vertices x, y belong to a subgraph isomorphic to some fixed graph H. We show that, for monotone graph classes, first-order sketchability is equivalent to small-distance sketchability. All proofs for small-distance sketching are in Section A.

▶ **Theorem 1.2.** *Let $\mathcal{F}$ be a monotone class of graphs. Then the following are equivalent:*
1. *$\mathcal{F}$ is small-distance sketchable;*
2. *$\mathcal{F}$ is first-order sketchable;*
3. *$\mathcal{F}$ has bounded expansion.*

The implications (3) $\implies$ (2) $\implies$ (1) do not require monotonicity. (2) $\implies$ (1) holds by definition. The proof of (3) $\implies$ (2) is straightforward, but relies on a structural result of [26] whose proof is highly technical. We actually get the stronger result that any *first-order transduction* of a class with bounded expansion is first-order sketchable, which improves the results of [40]. It was proved in [40], using structural results of [27], that any stable class of bounded twin-width is first-order sketchable. A stable class has bounded twin-width if and only if it is a transduction of a class of bounded sparse twin-width [27]. Every class of bounded sparse twin-width has bounded expansion, but the converse does not hold (e.g. for cubic graphs) [14], so our result generalizes the result of [40]. It essentially follows from using the structural results of [26] instead of [27].

Our proof of (1) $\implies$ (3) (Section A.4) requires our proof of Theorem 1.1 and some results in sparsity theory [50, 57]. We actually prove a stronger statement: for any monotone class $\mathcal{F}$, the existence of a sketch for deciding $\mathsf{dist}(x, y) \leq r$ vs. $\mathsf{dist}(x, y) > 5r - 1$, with size depending only on r, implies bounded expansion. Under a conjecture of Thomassen [64], we can replace the constant 5 with any arbitrarily large constant; see the remark after Conjecture A.17. Note that, even with a constant-factor gap between distance thresholds, this problem is distinct from ADT sketching, since the small-distance sketch size is allowed to depend on r. If we could replace the constant 5 with any arbitrarily large constant, this would immediately imply $\mathsf{ADT} \subseteq \mathsf{SD}$.

We also present a more direct proof of (3) $\implies$ (1), without going through first-order sketching, that allows for quantitative results. Going through first-order sketching (as was also done in [40]) proves the existence of a function $f(r)$ bounding the sketch size, without giving it explicitly. We obtain explicit bounds in terms of the *weak coloring number* [57], written as $\mathrm{wcol}_r(\mathcal{F})$ for any $r \in \mathbb{N}$ (Definition A.2). Using known bounds on the weak coloring number [66], we obtain the following corollary. As was the case for adjacency sketching, we observe that this proof (unlike the more general one for first-order sketching) produces sketches that only use randomization to compute a disjunction of EQUALITY checks, establishing that this extremely simple type of sketch suffices for monotone classes.

▶ **Corollary 1.3.** *Any graph class $\mathcal{F}$ with bounded $\mathrm{wcol}_r(\mathcal{F})$ admits a small-distance sketch of size $O(r + \mathrm{wcol}_r(\mathcal{F}) \log(\mathrm{wcol}_r(\mathcal{F})))$. In particular, planar graphs admit a small-distance sketch of size $O(r^3 \log r)$, and the class of K_t-minor-free graphs admits a small-distance sketch of size $O(r^{t-1} \log r)$. Furthermore, planar graphs admit a small-distance labelling scheme of size $O(r^3 \log n)$ and K_t-minor-free graphs admit a small-distance labelling scheme of size $O(r^{t-1} \log n)$.*

1.2.3 Approximate Distance Sketching

In light of Theorem 1.2, a reasonable question is whether ADT sketching for monotone classes is also determined by expansion. Our first result is that bounded expansion is necessary. All proofs on approximate distance sketching are omitted here due to space limitations, but can be found in the full version of the paper.

▶ **Theorem 1.4.** *If a monotone class $\mathcal{F}$ is ADT sketchable, then it has bounded expansion.*

Combined with Theorem 1.2, this proves $\mathsf{ADT} \subseteq \mathsf{SD}$. Our proof uses a recent and fairly involved result in extremal graph theory [53], along with the theory of sparsity [57], to show that an α-ADT sketch for a monotone class $\mathcal{F}$ of unbounded expansion could be used to get a constant-size sketch for deciding $\mathsf{dist}(x, y) \le 1$ vs. $\mathsf{dist}(x, y) > \alpha$ in arbitrary graphs, which (as we show) is a contradiction.

We are then concerned with the converse. We show that the class of max-degree 3 graphs, which has expansion exponential in r [56], is not ADT sketchable. After proving this theorem, we learned of an unpublished result [10] which proves a $\Theta(\log(n)/\alpha)$ bound for one-way communication of the α-ADT problem on degree-3 expander graphs. This could be used in place of our theorem to get the same qualitative (constant vs. non-constant) results, but not the quantitative bound: note that communication complexity cannot give sketching or labelling lower bounds better than $\Theta(\log n)$.

▶ **Theorem 1.5.** *For any $\alpha > 1$, any α-ADT sketch for the class of graphs with maximum degree 3 has size at least $\Omega(n^{\frac{1}{4\alpha}-\varepsilon})$, for any constant $\varepsilon > 0$.*

This establishes that $\mathsf{ADT} \neq \mathsf{SD}$ (and negatively answers open problem 2 of [2] about approximate distance labels for bounded-degree graphs, which [10] does not). But max-degree 3 graphs have exponential expansion. Smaller bounds on the expansion are associated with structural properties: for example, in monotone classes, polynomial expansion is equivalent to the existence of strongly sublinear separators [19]. One may then wonder if smaller bounds on the expansion suffice to guarantee ADT sketchability. We prove that this is not the case for two natural examples: subgraphs of the 3-dimensional grid (with polynomial expansion [57]), and subgraphs of the 2-dimensional grid with crosses (with linear expansion [18]) are not ADT sketchable. For this we require our Theorem 1.5.

▶ **Proposition 1.6.** *For the class of subgraphs of the 3-dimensional grid (the Cartesian product of 3 paths), and the class of subgraphs of the 2-dimensional grid (the strong product of 2 paths), an α-ADT sketch requires size at least $n^{\Omega(1/\alpha)}$.*

We strengthen this result by showing that one can obtain monotone classes of graphs with expansion that grows arbitrarily slowly, which are not ADT sketchable.

▶ **Theorem 1.7.** *For any function ρ tending to infinity, there exists a monotone class of expansion $r \mapsto \rho(r)$ that is not ADT sketchable. Moreover, for any $\varepsilon > 0$, there exists a monotone class $\mathcal{F}$ of expansion $r \mapsto O(r^\varepsilon)$, such that, if $\mathcal{F}$ admits an α-ADT sketch of size $s(n)$, then we must have $s(n) = n^{\Omega(1/\alpha)}$.*

We conclude with a brief discussion of upper bounds for ADT sketching. A number of concepts have been introduced in the literature that can be used to obtain ADT sketches, including sparse covers [12] and padded decompositions [49].

Using the sketches obtained from sparse covers, combined with results of [23] on sparse covers (based on [49, 22]), we obtain the following, which complements our Theorem 1.7; note that the graph classes with constant expansion are exactly the proper minor-closed classes [57].

▶ **Corollary 1.8.** *For any $t \ge 4$, the class of K_t-minor-free graphs has a $O(2^t)$-ADT sketch of size $O(t^2 \log t)$. The sketch is equality-based and has one-sided error. As a consequence, every monotone class of constant expansion is ADT sketchable.*

It is also relatively straightforward to obtain ADT sketches from padded decompositions, with an interesting difference. These sketches may not have one-sided error and, unlike all other positive examples of sketches in this paper, they may not be equality-based. On the other hand, they are extremely small. We can use constructions of padded decompositions due to [52, 3] to obtain the following remarkable corollary:

▶ **Corollary 1.9.** *For any $t \geq 4$, the class of K_t-minor-free graphs has an $O(t)$-ADT sketch of size 2. For $g \geq 0$, the class of graphs embeddable on a surface of Euler genus g has an $O(\log g)$-ADT sketch of size 2.*

1.3 Discussion & Open Problems

The main problem left open by this paper is Question 3 for monotone classes of graphs; we have shown that a constant bound on the expansion implies ADT sketchability, while arbitrarily small non-constant bounds do not, but this does not rule out a monotone, ADT sketchable class with non-constant expansion.

We have examples showing that ADT sketching does not imply small-distance sketching, in general. But our examples are not even hereditary. Is there a hereditary class that is ADT sketchable, but not small-distance or adjacency sketchable?

Our Theorem 1.2 shows that bounded expansion implies first-order sketchability, and that for monotone classes the converse also holds. We showed more generally that classes of *structurally bounded expansion* are first-order sketchable. To extend our study of sketchability beyond monotone classes, it would be interesting to investigate whether the converse of this statement holds: does first-order sketchability of a *hereditary* class imply structurally bounded expansion?

In the preprint of this paper, we asked whether the class of subgraphs of hypercubes is a counterexample to the Implicit Graph Conjecture (IGC), which asks for deterministic adjacency labels of size $O(\log n)$. This conjecture was refuted recently in [41] by a non-constructive argument, and it would be interesting to find a more natural class that refutes the conjecture. The *induced* subgraphs of hypercubes are adjacency sketchable and therefore admit adjacency labels of size $O(\log n)$ (see e.g. [40]), but our Corollary 3.6 shows that the subgraphs are not adjacency sketchable. Prior work (e.g. [16]) has not succeeded in finding labeling schemes of size $O(\log n)$ for this class. These observations made it plausible to us that efficient labeling schemes for this class do not exist. However, efficient adjacency labels for this class have since been found in [21]. A related question is whether we may characterize the monotone classes of graphs which admit adjacency labeling schemes of size $O(\log n)$.

We have focused on determining whether there *exists* a constant α such that a class is α-ADT sketchable. It is also of interest to obtain sketches for arbitrarily small $\alpha > 1$, with sketch size depending on α. One strategy is to embed the graph isometrically into ℓ_1, but this is not always the best option. We obtained a $(1+\varepsilon)$-ADT sketch for the class of forests with size $O\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)$, but this result appeared earlier in [10]; this sketch is more efficient than the one obtained by embedding the trees isometrically in ℓ_1. We remark that a class (monotone or not) that admits a $(1+\varepsilon)$-ADT sketch for $\varepsilon < 1$ must also admit an adjacency sketch.

Finally, we point out an interesting conjecture of [37], that all constant-cost public-coin communication problems contain a large monochromatic rectangle. In our terminology, using the equivalence between constant-cost communication and adjacency sketching from [40], this conjecture states that all adjacency sketchable graph classes have the strong Erdős-Hajnal property.

2 Preliminaries

2.1 Notation

Throughout the paper, log denotes the logarithm base 2, while ln denotes the natural logarithm.
We will write $\mathbf{1}[E]$ for the indicator variable for the event E, which takes value 1 if E is true.
Given a graph G, the length of a path P in G is the number of edges of P. Given two vertices $x, y \in V(G)$, we define $\mathsf{dist}_G(x, y)$ to be the infimum of the length of a path between x and y in G; we define $\mathsf{dist}_G(x, y) = \infty$ if there exists no path between x and y. Notice that $(V(G), \mathsf{dist}_G)$ is a metric space (with possibly infinite distances between pairs of vertices if G is disconnected).
The *girth* of a graph G is defined as the size of a shortest cycle in G (if G is acyclic, its girth is infinite).

2.2 Distance and First-Order Sketching

We will require more general notions of sketching than those introduced above. For a class $\mathcal{F}$ of graphs, we will say that a sequence $\{f_G\}_{G\in\mathcal{F}}$ of partial functions $f_G : V(G)\times V(G) \to \{0, 1, *\}$ is a *partial function f parameterized by graphs $G \in \mathcal{F}$*. We will write f to refer to this sequence.

For a graph class $\mathcal{F}$, we define an *f-sketch* for $\mathcal{F}$ as a decoder $D : \{0, 1\}^* \times \{0, 1\}^* \to \{0, 1\}$, such that for every $G \in \mathcal{F}$ the following holds. There is a probability distribution over functions $\mathsf{sk} : V(G) \to \{0, 1\}^*$, such that for all $x, y \in V(G)$,

$$f_G(x, y) \neq * \implies \Pr[D(\mathsf{sk}(x), \mathsf{sk}(y)) = f_G(x, y)] \geq 2/3\,.$$

We define the *size* of the sketch as

$$\max_{G\in\mathcal{F}_n} \sup_{\mathsf{sk}} \max_{x\in V(G)} |\mathsf{sk}(x)|\,,$$

where the supremum is over the set of functions $\mathsf{sk} : V(G) \to \{0, 1\}^*$ in the support of the distribution defined for G, and $|\mathsf{sk}(x)|$ is the number of bits of $\mathsf{sk}(x)$. We will say that a class $\mathcal{F}$ is *f-sketchable* if there exists an f-sketch for $\mathcal{F}$ with size that does not depend on the number of vertices n.

For a graph class $\mathcal{F}$, we also define an *f-labelling scheme* for $\mathcal{F}$ similar to above, except that for every $G \in \mathcal{F}$ there is a *deterministic* function $\ell : V(G) \to \{0, 1\}^*$ such that for all $x, y \in V(G)$,

$$f_G(x, y) \neq * \implies D(\ell(x), \ell(y)) = f_G(x, y)\,.$$

The following simple proposition (observed in [38, 40]) relates sketches to labelling schemes:

▶ **Proposition 2.1.** *If $\mathcal{F}$ admits an f-sketch of size $s(n)$, then it admits an f-labelling scheme of size $O(s(n) \log n)$.*

We now define certain important types of f-sketches. Let $\mathcal{F}$ be a class of graphs. For any $r_1 \leq r_2$, a *distance-(r_1, r_2) sketch* for $\mathcal{F}$ is an f-sketch, as defined above, when for any graph G we define the function

$$f_G(x, y) = \begin{cases} 1 & \text{if } \mathsf{dist}_G(x, y) \leq r_1 \\ 0 & \text{if } \mathsf{dist}_G(x, y) > r_2 \\ * & \text{otherwise.} \end{cases}$$

The size of such a sketch may depend on r_1, r_2, the number of vertices n, or other graph parameters.

Recall the definitions of *adjacency sketchable*, *small-distance sketchable*, and *ADT sketchable*. It is clear that:

1. A class $\mathcal{F}$ is *adjacency sketchable* if it is distance-$(1,1)$ sketchable;
2. A class $\mathcal{F}$ is *small-distance sketchable* if for every $r \geq 1$ it is distance-(r,r) sketchable.
3. A class $\mathcal{F}$ is α*-ADT sketchable* if for every $r \geq 1$ it is distance-$(r, \alpha r)$ sketchable, and furthermore the size of the sketch does not depend on r.

Following [40], we will also define *FO-sketchable* classes, for which we require some terminology (see e.g. [59] for more on the following terminology). A *relational vocabulary* Σ is a set of relation symbols, with each $R \in \Sigma$ having an *arity* $\text{arity}(R) \in \mathbb{N} \setminus \{0\}$. A Σ-structure $\mathcal{A}$ consists of a *domain* A, and for each relation symbol $R \in \Sigma$ an *interpretation* $R^{\mathcal{A}} \subseteq A^{\text{arity}(R)}$, which is a relation. Fix a countably infinite set X of *variables*. *Atomic formulas of vocabulary* Σ are of the form

- $x = y$ for $x, y \in X$; or,
- $R(x_1, \ldots, x_r)$ for $x_1, \ldots, x_r \in X$, $R \in \Sigma$ and $r = \text{arity}(R)$, which evaluates to true when $(x_1, \ldots, x_r) \in R$.

First-order (FO) formulas of vocabulary Σ are inductively defined as either atomic formulas, or a formula of the form $\neg\phi, \phi \wedge \psi, \phi \vee \psi$, or $\exists x.\phi$ or $\forall x.\psi$, where ϕ and ψ are each FO formulas. A *free variable* of a formula ϕ is one which is not bound by a quantifier. We will write $\phi(x_1, x_2, \ldots, x_k)$ to show that the free variables of ϕ are $x_1, \ldots, x_k \in X$. For a value $u \in A$, we write $\phi[u/x]$ for the formula obtained by substituting the constant u for the free variable x.

Let $\phi(x, y)$ be any formula with two free variables and relational vocabulary $\Sigma = \{E', R_1, \ldots, R_k\}$ where E' is symmetric of arity 2 and each R_i is unary (i.e. of arity 1). We will say that a graph class $\mathcal{F}$ is ϕ*-sketchable* if it is f-sketchable for any f chosen as follows. For any graph $G = (V, E)$, we choose any Σ-structure with domain V where E is the interpretation of the symbol E'. Then set $f_G(u, v) = 1$ if and only if $\phi(u/x, v/y)$ evaluates to true.

We remark that for any graph G, there are many ways to choose a Σ-structure with domain V with E being the interpretation of E'. To be first-order sketchable, a class $\mathcal{F}$ must be f-sketchable for *every* such choice of functions f_G. A concrete example is that, for any $r \in \mathbb{N}$, we can choose the formula

$$\phi(x,y) = \exists u_1, u_2, \ldots, u_{r-1} : (E'(x,u_1) \vee x = u_1) \wedge (E'(u_1,u_2) \vee u_1 = u_2) \wedge \cdots \wedge (E'(u_r, y) \vee u_r = y)\,,$$

which evaluates to true if and only if $\mathsf{dist}_G(x, y) \leq r$.

2.3 Bounded expansion

Here we introduce the notion of expansion from sparsity theory, as discussed in [57].

▶ **Definition 2.2** (Bounded Expansion). *Given a graph G and an integer $r \geq 0$, a* depth-r minor *of G is a graph obtained by contracting pairwise disjoint connected subgraphs of radius at most r in a subgraph of G. For any function f, we say that a class of graphs $\mathcal{G}$ has* expansion *at most f if any depth-r minor of a graph of $\mathcal{G}$ has average degree at most $f(r)$ (see [57] for more details on this notion). We say that a class $\mathcal{G}$ has* bounded expansion *if there is a function f such that $\mathcal{G}$ has expansion at most f.*

Note that, for example, every proper minor-closed family has constant expansion.

2.4 Equality-Based Labelling Schemes and Sketches

An equality-based labelling scheme is one which assigns to each vertex a deterministic label, comprising a data structure of size s that holds k "equality codes", which can be used only for checking equality. These labelling schemes: 1) capture the constant-cost randomized communication protocols that can be simulated by a constant-cost *deterministic* communication protocol with access to an EQUALITY oracle (as studied in e.g. [15, 13, 37, 40]); and 2) capture a common type of adjacency labels, including those of [47] for bounded arboricity graphs (see [40] for others).

One might formalize these schemes in a few ways; we slightly adapt the definition from [40]. This definition is intended to simplify notation rather than optimize label size, since we care mainly about constant vs. non-constant.

▶ **Definition 2.3** (Equality-Based Labeling Scheme). *Let $\mathcal{F}$ be a class of graphs and let $f : \mathbb{N} \times \mathbb{N} \times \mathcal{F} \to \{0, 1, *\}$ be a partial function. An (s, t, k)-equality-based f-labelling scheme for $\mathcal{F}$ is an algorithm D, called a* decoder*, which satisfies the following. For every $G \in \mathcal{F}$ with vertex set $[n]$ and every $x \in [n]$, there is a sequence of the form*

$$\ell_G(x) = [(p_1(x) \mid \vec{q}_1(x)), (p_2(x) \mid \vec{q}_2(x)), \ldots, (p_t(x) \mid \vec{q}_t(x))]\,,$$

where the vectors $p_i(x) \in \{0, 1\}^*$ are called the prefixes*, and the entries of the vectors $\vec{q}_i(x) \in \mathbb{N}^*$ are called* equality codes *(which we may assume are positive integers). We must have $\sum_{i=1}^t |p_i(x)| \le s$ and $\sum_{i=1}^t |\vec{q}_i(x)| \le k$ (recall that given a vector v of binary numbers or integers, $|v|$ denotes the number of entries of v). We insist on the fact that k bounds the total number of equality codes associated with any vertex x, but not necessarily the total number of bits needed to store these codes (see Example 2.4 below, where $k = 2$ but storing the codes would require $2 \log n$ bits per vertex). On inputs $\ell_G(x), \ell_G(y)$, the algorithm D chooses a function $D_{p(x),p(y)}$, where $p(x) = (p_1(x), \ldots, p_t(x))$, and outputs*

$$D_{p(x),p(y)}(Q_{x,y})\,,$$

where

$$Q_{x,y}(i_1, i_2, j_1, j_2) = \mathbf{1}[(\vec{q}_{i_1}(x))_{j_1} = (\vec{q}_{i_2}(x))_{j_2}] \tag{1}$$

is the set of equality values for every pair of equality codes. It is required that, for every $G \in \mathcal{F}$ and $x, y \in V(G)$,

$$f(x, y, G) \ne * \implies D_{p(x),p(y)}(Q_{x,y}) = f(x, y, G)\,.$$

*We make the further distinction of calling a labelling scheme (s, k)-*disjunctive *if it is an (s, t, k)-equality-based labelling scheme, where each function $D_{p(x),p(y)}$ is simply a disjunction over a subset of values $Q_{x,y}(i_1, i_2, j_1, j_2)$.*

When an element $(p_i(x) \mid \vec{q}_i(x))$ in an equality-based label has $p_i(x)$ of size 0, we will write $(- \mid \vec{q}_i(x))$; similarly, we write $(p_1(x) \mid -)$ when $\vec{q}_i(x)$ is empty.

▶ **Example 2.4.** The adjacency labelling scheme of [47] for forests can be written as an equality-based labelling scheme. For each x in an n-vertex forest G with arbitrarily rooted trees, which we assume has vertex set $[n]$, we assign the label $\ell_G(x) = [(- \mid (x, p(x)))]$ where $p(x)$ is the parent of x if it has one, or 0 otherwise. Here $\vec{q}_1(x) = (x, p(x)) \in \mathbb{N}^2$. The decoder simply outputs the disjunction of $p(x) = y$ or $p(y) = x$, so in fact this is a $(0, 1, 2)$-disjunctive labeling scheme.

An equality-based labelling scheme is easily transformed into a standard deterministic labelling scheme or a sketch. The following simple proposition was observed in [40]. We sketch the proof for the sake of clarity.

▶ **Proposition 2.5.** *Let $\mathcal{F}$ be a class of graphs and $f : \mathbb{N} \times \mathbb{N} \times \mathcal{F} \to \{0, 1, *\}$ be a partial function. If there is an (s, t, k)-equality-based f-labelling scheme for $\mathcal{F}$ then there is an f-sketch for $\mathcal{F}$ of size at most $O(s + t + k \log k)$. If the scheme is disjunctive, the sketch has* one-sided error*: when $f(x, y, G) = 1$, the sketch will produce the wrong output with probability 0.*

Proof sketch. Choose a random function $\xi : \mathbb{N} \to [w]$ for $w = 3k^2$. For any vertex x of a graph G, replace each vector $\vec{q}_i(x) = (q_{i,1}(x), \ldots, q_{i,m}(x))$ with $(\xi(q_{i,1}(x)), \ldots, \xi(q_{i,m}(x)))$. We have replaced each of the (at most) k equality codes $(\vec{q}_i(x))_j$ with $\xi((q_i(x))_j)$, using $k \log w = O(k \log k)$ bits in total. The sketch has size $O(s+t+k \log k)$ since we must include each $p_i(x)$ (using s bits in total), the $O(k \log k)$ bits for the equality codes, and $O(t)$ bits to encode the symbols (|).

For two vertices x, y, write $Q^{\xi}_{x,y}(i_1, i_2, j_1, j_2) = \mathbf{1}[\xi((\vec{q}_{i_1}(x))_{i_2}) = \xi((\vec{q}_{j_1}(y))_{j_2})]$. Since there are at most k equality codes in each label, there are at most k^2 equality comparisons. By the union bound, the probability that any of these comparisons have

$$\mathbf{1}[\xi((\vec{q}_{i_1}(x))_{i_2}) = \xi((\vec{q}_{j_1}(y))_{j_2})] \neq \mathbf{1}[(\vec{q}_{i_1}(x))_{i_2} = (\vec{q}_{j_1}(y))_{j_2}]$$

is at most $k^2 \cdot (1/w) = 1/3$, so with probability at least $2/3$ all of the comparisons made by the decoder have the correct value, so the decoder will be correct. Note that when $(\vec{q}_{i_1}(x))_{i_2} = (\vec{q}_{j_1}(y))_{j_2}$, the random values under ξ will be equal with certainty. We conclude from this that disjunctive schemes will produce sketches with one-sided error. ◀

▶ Remark 2.6. Disjunctive labelling schemes with $s = 0$ (i.e. the p values are empty) can be transformed into *locality-sensitive hashes (LSH)* [45]. A $(r_1, r_2, \gamma_1, \gamma_2)$-LSH must map any two points x, y with $\mathsf{dist}(x, y) \leq r_1$ to the same hash value with probability at least γ_1, and map any two points x, y with $\mathsf{dist}(x, y) > r_2$ to the same hash value with probability at most γ_2, where $r_1 < r_2$ and $\gamma_1 > \gamma_2$. By boosting the success probability of each Equality check in the disjunction, and then sampling a uniformly random term from the disjunction, one obtains an LSH with distance parameters that depend on the original sketch. All of the equality-based sketches presented in this paper, except the first-order sketches, are of this form.

3 Adjacency Sketching

In this section, we prove Theorem 1.1, and include the additional equivalent statement that $\mathcal{F}$ admits a constant-size *disjunctive* adjacency sketch. We think of disjunctive sketches as the simplest possible use of randomization in a sketch, with the theorem establishing that the simplest possible sketches are sufficient for monotone classes.

▶ **Theorem 3.1.** *Let $\mathcal{F}$ be a monotone class of graphs. Then the following are equivalent:*
1. *$\mathcal{F}$ is adjacency sketchable.*
2. *$\mathcal{F}$ admits a constant-size disjunctive adjacency labelling scheme.*
3. *$\mathcal{F}$ has bounded arboricity.*

A disjunctive labelling scheme for graphs of arboricity k can be obtained from the adjacency labelling scheme of [47], as in Example 2.4. This leads to a sketch of size $O(k \log k)$ by Proposition 2.5, which was improved slightly in [40]:

▶ **Proposition 3.2** ([40])**.** *Let $\mathcal{F}$ be any class with arboricity at most k. Then $\mathcal{F}$ admits a $(0, 1, k+1)$-disjunctive adjacency labelling scheme, and an adjacency sketch of size $O(k)$.*

Therefore, to prove Theorem 3.1, it suffices to prove (1) $\implies$ (3), which we will prove by contrapositive. This proof will use the notion of discrepancy from communication complexity. Our proof is inspired by the recent proof of Hambardzumyan, Hatami, & Hatami [37], which refuted the main conjecture of [40]. Our proof also leads to another, more natural counterexample to the conjecture of [40]: the class of subgraphs of the hypercube.

Consider a graph $G = (V, E)$, let $f : V \times V \to \{0, 1, *\}$ be a partial function, and let μ be a probability distribution over $V \times V$ that is supported on pairs (x, y) which satisfy $f(x, y) \neq *$. Let $X, Y \subseteq V$. Then we define the *discrepancy* of $R = X \times Y$ as

$$\mathrm{Disc}_{\mu,f}(G, R) = \left| \Pr_{\mu}[(x, y) \in R \cap f^{-1}(1)] - \Pr_{\mu}[(x, y) \in R \cap f^{-1}(0)] \right| ,$$

where (x, y) is drawn from μ. The discrepancy of G under μ is defined as

$$\mathrm{Disc}_{\mu,f}(G) = \max_{R} \mathrm{Disc}_{\mu,f}(G, R) ,$$

where the maximum is over all sets $R = X \times Y$ with $X, Y \subseteq V$. The following lemma is essentially a restatement of a standard lower-bound technique in communication complexity.

▶ **Lemma 3.3.** *Let $G = (V, E)$ be any graph on n vertices, let $\mathcal{F}$ be any class of graphs containing G, and let f be a partial function parameterized by graphs in $\mathcal{F}$. Let μ be any probability distribution over $V \times V$ supported on a subset of $\{(x, y) : f_G(x, y) \neq *\}$. Then any f-sketch for $\mathcal{F}_n$ has size at least $\frac{1}{2} \log \frac{1}{3\mathrm{Disc}_{\mu,f}(G)}$.*

A *spanning subgraph* of a graph $G = (V, E)$ is a subgraph of G with vertex set V. Our next lemma will give a lower bound on the adjacency sketch size for the class $\mathcal{G}$ of spanning sugraphs of a graph G of minimum degree d. We will actually prove the lower bound for a weaker type of adjacency sketch, which is only required to be correct on pairs (x, y) that were originally edges in G. This stronger statement is not necessary for the current section, but will be used in the proof of Theorem A.16.

For a graph $G = (V, E)$ and the class $\mathcal{G}$ of spanning subgraphs of G, and any subgraph $H \in \mathcal{G}$, we will define the partial function $\mathrm{adj}^E_H : V \times V \to \{0, 1, *\}$ as

$$\mathrm{adj}^E_H(x, y) = \begin{cases} \mathrm{adj}_H(x, y) & \text{if } (x, y) \in E \\ * & \text{otherwise.} \end{cases}$$

In the remainder of this section, we view adj^E as the function $(\mathrm{adj}^E_H)_{H \in \mathcal{G}}$ parametrized by $H \in \mathcal{G}$. In particular, an adj^E-sketch for $\mathcal{G}$ computes the partial function adj^E_H for each $H \in \mathcal{G}$.

We show by the probabilistic method that there is a distribution μ and a subgraph of G with discrepancy $O(1/\sqrt{d})$ with respect to μ. We will require the standard Chernoff bound for the binomial distribution with parameters n and $\frac{1}{2}$ (see Corollary A.1.2 in [5]): for any $t > 0$,

$$\Pr(|\mathrm{Bin}(n, \tfrac{1}{2}) - \tfrac{n}{2}| > t) < 2 \exp(-2t^2/n).$$

▶ **Lemma 3.4.** *Let $G = (V, E)$ be a graph of minimum degree d, and let $\mathcal{G}$ be the class of spanning subgraphs of G. Then any adj^E-sketch for $\mathcal{G}$ requires size at least $\Omega(\log d)$.*

Proof. Let $\mathbf{H}$ be a random spanning subgraph of G obtained by including each edge of G independently with probability $1/2$. Note that $\mathbf{H} \in \mathcal{G}$ with probability 1. Let $m = |E|$ and let μ be the probability distribution over $V \times V$ such that for every $(x, y) \in V \times V$, we have $\mu((x, y)) = 1/m$ if $(x, y) \in E$, and $\mu((x, y)) = 0$ otherwise (so that μ is uniform over the edges of G). For simplicity, write Disc_μ for $\mathrm{Disc}_{\mu,f}$ where $f = \mathrm{adj}^E$. We will prove that $\mathrm{Disc}_\mu(\mathbf{H})$ is small, with nonzero probability over $\mathbf{H}$.

Consider a set $R = X \times Y$ with $X, Y \subseteq V$, and let $k \leq m$ be the number of edges (x, y) of G with $(x, y) \in R$. Let H be any subgraph of G with $|E(H) \cap R| = \ell \leq k$. Then

$$\begin{aligned}\mathrm{Disc}_\mu(H, R) &= \left| \Pr_\mu[(x, y) \in E(H) \cap R] - \Pr_\mu[(x, y) \in R \setminus E(H)] \right| \\ &= \left| \frac{\ell}{m} - \frac{k - \ell}{m} \right| = \frac{|2\ell - k|}{m} .\end{aligned}$$

For fixed $R = X \times Y$, it then holds that $\mathrm{Disc}_\mu(\mathbf{H}, R)$ is a random variable $\frac{|2\ell - k|}{m}$, where $\ell \sim \mathrm{Bin}(k, 1/2)$. Then, by the Chernoff bound, we have for any $\varepsilon > 0$ that

$$\Pr[\mathrm{Disc}_\mu(\mathbf{H}, R) > \varepsilon] = \Pr\left[\left|\mathrm{Bin}\left(k, \tfrac{1}{2}\right) - \tfrac{k}{2}\right| > \tfrac{1}{2}\varepsilon m\right] \leq 2\exp\left(-\tfrac{\varepsilon^2 m^2}{2k}\right) \leq 2\exp\left(-\varepsilon^2 m/2\right) ,$$

where the last inequality is due to $k \leq m$. There are at most 2^{2n} sets $R = X \times Y \subseteq V \times V$, so by the union bound,

$$\begin{aligned}\Pr\left[\exists R = X \times Y \subseteq V \times V : \mathrm{Disc}_\mu(\mathbf{H}, R) > \varepsilon\right] &\leq 2^{2n+1} \exp\left(-\varepsilon^2 m/2\right) \\ &= \exp\left((2n + 1)\ln(2) - \varepsilon^2 m/2\right) .\end{aligned}$$

Now, since G has minimum degree d, we have $m \geq dn/2$. Setting $\varepsilon = \Omega\left(\frac{1}{\sqrt{d}}\right)$ with a sufficiently large implicit multiplicative constant, we get an upper bound on this probability of

$$\exp\left((2n + 1)\ln(2) - \varepsilon^2 m/2\right) \leq \exp\left((2n + 1)\ln(2) - \varepsilon^2 dn/4\right) < 1 .$$

Therefore there exists a subgraph H with $\mathrm{Disc}_\mu(H) = O(1/\sqrt{d})$. Applying Lemma 3.3, we see that any adj^E-sketch for $\mathcal{G}$ must have size at least $\Omega(\log(\sqrt{d})) = \Omega(\log d)$. ◀

We may now complete the proof of Theorem 3.1. We aim to prove (1) $\implies$ (3), which we will prove by contrapositive: i.e. that any class of unbounded arboricity has non-constant adjacency sketch size.

▶ **Lemma 3.5.** *Let $\mathcal{F}$ be any monotone class of graphs with unbounded arboricity. Then $\mathcal{F}$ does not admit a constant-size adjacency sketch.*

Proof. It is well-known that the degeneracy of a graph is within factor 2 of the arboricity, so the degeneracy of $\mathcal{F}$ must also be unbounded. Then for any integer $d \in \mathbb{N}$, there is a graph $G \in \mathcal{F}$ with degeneracy at least d. By definition, G contains a subgraph H of minimum degree at least d. Let $\mathcal{G}$ be the class of spanning subgraphs of G. Since $\mathcal{F}$ is monotone, we have $\mathcal{G} \subseteq \mathcal{F}$. Then by Lemma 3.4, any adjacency sketch for $\mathcal{G}$ must have size at least $\Omega(\log d)$. Then for any integer d, we obtain a lower bound of $\Omega(\log d)$ on the size of an adjacency sketch for $\mathcal{F}$; it follows that any adjacency sketch for $\mathcal{F}$ is of non-constant size. ◀

As a consequence, we obtain the following counterexample to the main conjecture of [40]. We remind the reader that the conjecture was already refuted in [36], using an interesting construction of a graph class that was originally used to establish a "proof barrier" in communication complexity [37]. Our counterexample, the subgraphs of the hypercube, is more easily defined. The following bound on the number of subgraphs of the hypercube was observed by Viktor Zamaraev (personal communication). See [40] for a definition of *stable*.

▶ **Corollary 3.6.** *Let $\mathcal{F}$ be a class of subgraphs of the hypercube. Then:*

1. *$\mathcal{F}$ is stable, and there are at most $2^{O(n \log n)}$ graphs on n vertices in $\mathcal{F}$.*
2. *$\mathcal{F}$ is not adjacency sketchable.*

Proof. Since the d-dimensional hypercube of size $N = 2^d$ has minimum degree $d = \log N$, $\mathcal{F}$ has non-constant adjacency sketch size. To bound the number of n-vertex subgraphs of the hypercubes, we first observe that there are at most $2^{O(n \log n)}$ *induced* subgraphs of the hypercube on n vertices, which follows from the $O(\log n)$ adjacency labelling scheme for this class [38] (see a simpler exposition at [39]). It is known that any n-vertex induced subgraph of the hypercube has at most $O(n \log n)$ edges [33], so each induced subgraph admits at most $2^{O(n \log n)}$ spanning subgraphs. Therefore the number of n-vertex subgraphs of the hypercube is at most $2^{O(n \log n)} \cdot 2^{O(n \log n)} = 2^{O(n \log n)}$. Any monotone class of graphs which is not stable contains $K_{t,t}$, for every $t \in \mathbb{N}$, and therefore contains the class of all bipartite graphs. This does not hold for $\mathcal{F}$ (or indeed for any class of factorial speed), so $\mathcal{F}$ must be stable. ◀

References

1 Ittai Abraham, Shiri Chechik, and Cyril Gavoille. Fully dynamic approximate distance oracles for planar graphs via forbidden-set distance labels. In *Proceedings of the forty-fourth annual ACM Symposium on Theory of Computing (STOC 2012)*, pages 1199–1218, 2012.

2 Ittai Abraham and Cyril Gavoille. Object location using path separators. In *Proceedings of the twenty-fifth annual ACM Symposium on Principles of Distributed Computing (PODC 2006)*, pages 188–197, 2006.

3 Ittai Abraham, Cyril Gavoille, Anupam Gupta, Ofer Neiman, and Kunal Talwar. Cops, robbers, and threatening skeletons: Padded decomposition for minor-free graphs. *SIAM Journal on Computing*, 48(3):1120–1145, 2019.

4 Noga Alon, Yossi Matias, and Mario Szegedy. The space complexity of approximating the frequency moments. *Journal of Computer and system sciences*, 58(1):137–147, 1999.

5 Noga Alon and Joel H Spencer. *The probabilistic method.* John Wiley & Sons, 2016.

6 Stephen Alstrup, Philip Bille, and Theis Rauhe. Labeling schemes for small distances in trees. *SIAM Journal on Discrete Mathematics*, 19(2):448–462, 2005.

7 Stephen Alstrup, Søren Dahlgaard, Mathias Bæk Tejs Knudsen, and Ely Porat. Sublinear distance labeling. In *24th Annual European Symposium on Algorithms (ESA 2016)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2016.

8 Stephen Alstrup, Cyril Gavoille, Esben Bistrup Halvorsen, and Holger Petersen. Simpler, faster and shorter labels for distances in graphs. In *Proceedings of the twenty-seventh annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2016)*, pages 338–350. SIAM, 2016.

9 Stephen Alstrup, Inge Li Gørtz, Esben Bistrup Halvorsen, and Ely Porat. Distance labeling schemes for trees. In *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2016.

10 Alexandr Andoni and Robert Krauthgamer. Distance estimation protocols for general metrics. `https://www.cs.columbia.edu/~andoni/papers/de.pdf`, 2008.

11 Alexandr Andoni, Robert Krauthgamer, and Ilya P. Razenshteyn. Sketching and embedding are equivalent for norms. *SIAM J. Comput.*, 47(3):890–916, 2018. `doi:10.1137/15M1017958`.

12 Baruch Awerbuch and David Peleg. Sparse partitions (extended abstract). In *31st Annual Symposium on Foundations of Computer Science, St. Louis, Missouri, USA, October 22-24, 1990, Volume II*, pages 503–513. IEEE Computer Society, 1990. `doi:10.1109/FSCS.1990.89571`.

13 Alexander R. Block, Simina Branzei, Hemanta K. Maji, Himanshi Mehta, Tamalika Mukherjee, and Hai H. Nguyen. P_4-free partition and cover numbers and application. Cryptology ePrint Archive, Report 2020/1605, 2020. URL: `https://ia.cr/2020/1605`.

14 Édouard Bonnet, Colin Geniet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width II: small classes. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA 2021)*, pages 1977–1996. SIAM, 2021.

15 Arkadev Chattopadhyay, Shachar Lovett, and Marc Vinyals. Equality alone does not simulate randomness. In *34th Computational Complexity Conference (CCC 2019)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2019.

16 Victor Chepoi, Arnaud Labourel, and Sébastien Ratel. On density of subgraphs of Cartesian products. *Journal of Graph Theory*, 93(1):64–87, 2020.

17 Erik D Demaine, Felix Reidl, Peter Rossmanith, Fernando Sánchez Villaamil, Somnath Sikdar, and Blair D Sullivan. Structural sparsity of complex networks: Bounded expansion in random models and real-world graphs. *arXiv preprint*, 2014. arXiv:1406.2587.

18 Zdeněk Dvorák. A note on sublinear separators and expansion. *European J. Combin.*, 93:103273, 2021. doi:10.1016/j.ejc.2020.103273.

19 Zdeněk Dvorák and Sergey Norin. Strongly sublinear separators and polynomial expansion. *SIAM J. Discret. Math.*, 30(2):1095–1101, 2016. doi:10.1137/15M1017569.

20 Talya Eden, Piotr Indyk, and Haike Xu. Embeddings and labeling schemes for A^*. In *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, volume 215, page 62. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2022.

21 Louis Esperet, Nathaniel Harms, and Viktor Zamaraev. Optimal adjacency labels for subgraphs of cartesian products. *arXiv preprint*, 2022. arXiv:2206.02872.

22 Jittat Fakcharoenphol and Kunal Talwar. An improved decomposition theorem for graphs excluding a fixed minor. In *Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques*, pages 36–46. Springer, 2003.

23 Arnold Filtser. Scattering and sparse partitions, and their applications. In *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2020.

24 Pierre Fraigniaud and Amos Korman. On randomized representations of graphs using short labels. In *Proceedings of the twenty-first annual Symposium on Parallelism in Algorithms and Architectures (SPAA 2009)*, pages 131–137, 2009.

25 Ofer Freedman, Paweł Gawrychowski, Patrick K Nicholson, and Oren Weimann. Optimal distance labeling schemes for trees. In *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC 2017)*, pages 185–194, 2017.

26 Jakub Gajarský, Stephan Kreutzer, Jaroslav Nešetřil, Patrice Ossona de Mendez, Michał Pilipczuk, Sebastian Siebertz, and Szymon Toruńczyk. First-order interpretations of bounded expansion classes. *ACM Trans. Comput. Logic*, 21(4), July 2020. doi:10.1145/3382093.

27 Jakub Gajarský, Michał Pilipczuk, and Szymon Toruńczyk. Stable graphs of bounded twin-width. *arXiv preprint*, 2021. arXiv:2107.03711.

28 Cyril Gavoille, Michal Katz, Nir A Katz, Christophe Paul, and David Peleg. Approximate distance labeling schemes. In *European Symposium on Algorithms (ESA 2001)*, pages 476–487. Springer, 2001.

29 Cyril Gavoille and Arnaud Labourel. On local representation of distances in trees. In *Proceedings of the twenty-sixth annual ACM Symposium on Principles of Distributed Computing (PODC 2007)*, pages 352–353, 2007.

30 Cyril Gavoille and David Peleg. Compact and localized distributed data structures. *Distributed Computing*, 16(2):111–120, 2003.

31 Cyril Gavoille, David Peleg, Stéphane Pérennes, and Ran Raz. Distance labeling in graphs. *Journal of Algorithms*, 53(1):85–112, 2004.

32 Paweł Gawrychowski and Przemysław Uznański. Better distance labeling for unweighted planar graphs. In *Workshop on Algorithms and Data Structures (WADS 2021)*, pages 428–441. Springer, 2021.

33 Ron L. Graham. On primitive graphs and optimal vertex assignments. *Annals of the New York academy of sciences*, 175(1):170–186, 1970.

34 Martin Grohe, Stephan Kreutzer, and Sebastian Siebertz. Deciding first-order properties of nowhere dense graphs. *Journal of the ACM (JACM)*, 64(3):1–32, 2017.

35 Anupam Gupta, Ilan Newman, Yuri Rabinovich, and Alistair Sinclair. Cuts, trees and ℓ_1-embeddings of graphs. *Combinatorica*, 24(2):233–269, 2004.

36 Lianna Hambardzumyan, Hamed Hatami, and Pooya Hatami. A counter-example to the probabilistic universal graph conjecture via randomized communication complexity, 2021. `arXiv:2111.10436`.

37 Lianna Hambardzumyan, Hamed Hatami, and Pooya Hatami. Dimension-free bounds and structural results in communication complexity. *Electronic Colloquium on Computational Complexity (ECCC)*, page 66, 2021. URL: `https://eccc.weizmann.ac.il/report/2021/066`.

38 Nathaniel Harms. Universal communication, universal graphs, and graph labeling. In *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*, volume 151, page 33. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2020.

39 Nathaniel Harms. Adjacency labeling and sketching for induced subgraphs of the hypercube. `https://cs.uwaterloo.ca/~nharms/downloads/hypercube_sketch.pdf`, 2022. URL: `https://cs.uwaterloo.ca/~nharms/downloads/hypercube_sketch.pdf`.

40 Nathaniel Harms, Sebastian Wild, and Viktor Zamaraev. Randomized communication and implicit graph representations. In *Proceedings of the 54th annual ACM Symposium on Theory of Computing (STOC)*, 2022.

41 Hamed Hatami and Pooya Hatami. The implicit graph conjecture is false. *arXiv preprint arXiv:2111.13198*, 2021.

42 Wei Huang, Yaoyun Shi, Shengyu Zhang, and Yufan Zhu. The communication complexity of the hamming distance problem. *Information Processing Letters*, 99(4):149–153, 2006.

43 Tony Huynh, Bojan Mohar, Robert Šámal, Carsten Thomassen, and David R Wood. Universality in minor-closed graph classes. *arXiv preprint*, 2021. `arXiv:2109.00327`.

44 Piotr Indyk. Stable distributions, pseudorandom generators, embeddings, and data stream computation. *Journal of the ACM (JACM)*, 53(3):307–323, 2006.

45 Piotr Indyk and Rajeev Motwani. Approximate nearest neighbors: towards removing the curse of dimensionality. In *Proceedings of the thirtieth annual ACM Symposium on Theory of Computing (STOC 1998)*, pages 604–613, 1998.

46 T.S. Jayram. Problem 25: Communication complexity and metric spaces. `https://sublinear.info/25`, 2009. URL: `https://sublinear.info/25`.

47 Sampath Kannan, Moni Naor, and Steven Rudich. Implicit representation of graphs. *SIAM Journal on Discrete Mathematics*, 5(4):596–603, 1992.

48 Subhash Khot and Assaf Naor. The Andoni–Krauthgamer–Razenshteyn characterization of sketchable norms fails for sketchable metrics. In *Proceedings of the thirtieth annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2019)*, pages 1814–1824. SIAM, 2019.

49 Philip Klein, Serge A Plotkin, and Satish Rao. Excluded minors, network decomposition, and multicommodity flow. In *Proceedings of the twenty-fifth annual ACM Symposium on Theory of Computing (STOC 1993)*, pages 682–690, 1993.

50 Daniela Kühn and Deryk Osthus. Every graph of sufficiently large average degree contains a C_4-free subgraph of large average degree. *Combinatorica*, 24(1):155–162, 2004.

51 Eyal Kushilevitz, Rafail Ostrovsky, and Yuval Rabani. Efficient search for approximate nearest neighbor in high dimensional spaces. *SIAM Journal on Computing*, 30(2):457–474, 2000.

52 James R Lee and Anastasios Sidiropoulos. Genus and the geometry of the cut graph. In *Proceedings of the twenty-first annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2010)*, pages 193–201. SIAM, 2010.

53 Hong Liu and Richard Montgomery. A solution to Erdős and Hajnal's odd cycle problem. *arXiv preprint*, 2020. `arXiv:2010.15802`.

54 Jiří Matoušek. *Lectures on discrete geometry*, volume 212. Springer Science & Business Media, 2013.

55 John H Muller. Local structure in graph classes, 1989.

56 Jaroslav Nešetřil and Patrice Ossona de Mendez. Grad and classes with bounded expansion II. algorithmic aspects. *European Journal of Combinatorics*, 29(3):777–791, 2008.

57 Jaroslav Nešetřil and Patrice Ossona de Mendez. *Sparsity: graphs, structures, and algorithms*, volume 28. Springer-Verlag, 2012.

58 Jaroslav Nešetřil and Patrice Ossona de Mendez. On low tree-depth decompositions. *Graphs and combinatorics*, 31(6):1941–1963, 2015.

59 Jaroslav Nešetřil, Patrice Ossona de Mendez, and Sebastian Siebertz. Structural properties of the first-order transduction quasiorder. In *30th EACSL Annual Conference on Computer Science Logic (CSL 2022)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2022.

60 David Peleg. Informative labeling schemes for graphs. *Theor. Comput. Sci.*, 340(3):577–593, 2005.

61 Ilya Razenshteyn. *High-dimensional similarity search and sketching: algorithms and hardness*. PhD thesis, Massachusetts Institute of Technology, 2017.

62 Michael Saks and Xiaodong Sun. Space lower bounds for distance approximation in the data stream model. In *Proceedings of the thirty-fourth annual ACM Symposium on Theory of Computing (STOC 2002)*, pages 360–369, 2002.

63 Jeremy P. Spinrad. *Efficient graph representations*. American Mathematical Society, 2003.

64 Carsten Thomassen. Girth in graphs. *Journal of Combinatorial Theory, Series B*, 35(2):129–141, 1983.

65 Mikkel Thorup. Compact oracles for reachability and approximate distances in planar digraphs. *Journal of the ACM*, 51(6):993–1024, 2004.

66 Jan van den Heuvel, Patrice Ossona de Mendez, Daniel Quiroz, Roman Rabinovich, and Sebastian Siebertz. On the generalised colouring numbers of graphs that exclude a fixed minor. *European Journal of Combinatorics*, 66:129–144, 2017.

A Small-Distance Sketching

In this section we prove Theorem 1.2. As in Theorem 3.1, we refine the theorem by showing that the sketches are in fact disjunctive.

▶ **Theorem A.1.** *Let $\mathcal{F}$ be a monotone class of graphs. Then the following are equivalent:*

1. *$\mathcal{F}$ is small-distance sketchable.*
2. *For some function $f : \mathbb{N} \to \mathbb{N}$ and every $r \in \mathbb{N}$, $\mathcal{F}$ admits a disjunctive small-distance labelling scheme of size $f(r)$.*
3. *$\mathcal{F}$ is first-order sketchable.*
4. *$\mathcal{F}$ has bounded expansion.*

It holds by definition that (3) $\implies$ (1) and (2) $\implies$ (1), even without the assumption of monotonicity. We will prove (4) $\implies$ (3) and (4) $\implies$ (2) using different methods. We prove (4) $\implies$ (3) (again without the assumption of monotonicity) in Section A.2 using the structural result of [26]. This proof does not give explicit bounds on the sketch size. (4) $\implies$ (2) is proved in Section A.3 and gives explicit upper bounds on the sketch size. The final piece of the theorem, (1) $\implies$ (4), is proved in Section A.4.

A.1 Bounded expansion

▶ **Definition A.2** (Weakly r-reachable). *Given a total order $(V,<)$ on the vertex set V of a graph G and an integer $r \geq 0$, we say that a vertex $v \in V$ is* weakly r-reachable *from a vertex $u \in V$ if there is a path of length at most r connecting v to u in G, and such that for any vertex w on the path, $v \leq w$ (in words, v is the smallest vertex on the path with respect to $(V,<)$). For a graph G and an integer $r \geq 0$, we denote by* $\mathrm{wcol}_r(G)$ *the smallest integer k for which the vertex set of G has a total order $(V,<)$ such that for any vertex $u \in V$, at* *most k vertices are weakly r-reachable from u with respect to $(V,<)$. For a graph class $\mathcal{F}$, we write* $\mathrm{wcol}_r(\mathcal{F})$ *for the supremum of* $\mathrm{wcol}_r(G)$*, for $G \in \mathcal{F}$.*

▶ **Definition A.3** ((k,ℓ)-Subdivisions)**.** *For a graph G and two integers $0 \le k \le \ell$, a (k,ℓ)-subdivision of G is any graph obtained from G by subdividing each edge of G at least k times and at most ℓ times (i.e. we replace each edge of G by a path with at least k and an most ℓ internal vertices). A (k,k)-subdivision is also called a* k-subdivision *for simplicity;*

▶ **Definition A.4** (Depth-r Topological Minor)**.** *We say that H is a* depth-r topological minor *of a graph G if G contains a $(0,2r)$-subdivision of H as a subgraph. In the proof below it will be convenient to use the following equivalent definition of bounded expansion [57].*

▶ **Theorem A.5.** *For a class $\mathcal{F}$ of graphs, the following are equivalent:*
1. *$\mathcal{F}$ has bounded expansion.*
2. *There is a function $f : \mathbb{N} \to \mathbb{N}$ such that for any $r \in \mathbb{N}$,* $\mathrm{wcol}_r(\mathcal{F}) \le f(r)$*.*
3. *There is a function $f : \mathbb{N} \to \mathbb{N}$ such that for any $r \in \mathbb{N}$ and any $G \in \mathcal{F}$, any depth-r topological minor of G has average degree at most $f(r)$.*

We will also require the following standard fact about the expansion of monotone classes, which is a simple consequence of Theorem A.5 (see for instance [58]) combined with a result of Kühn & Osthus [50].

▶ **Corollary A.6.** *Let $\mathcal{F}$ be a monotone class of unbounded expansion. Then there is a constant $r \ge 0$, so that for any $d \ge 0$, $\mathcal{F}$ contains an r-subdivision of a bipartite graph of minimum degree at least d and girth at least 6.*

The proof is omitted here due to the space limitation.

A.2 Bounded Expansion Implies FO Labelling Schemes

To prove that any class of bounded expansion is first-order sketchable, we use the result of [26] that shows how to decompose any class of (structurally) bounded expansion into a number of graphs of bounded shrubdepth. We will require an adjacency sketch for classes of bounded shrubdepth, given below.

A.2.1 Adjacency Sketching for Bounded Shrubdepth

We must first define shrubdepth. A *connection model* for a graph G is a rooted tree T whose nodes are colored with a bounded number of colors such that:
- the vertices of G are the leaves of T; and
- for two vertices $u, v \in V(G)$, whether u and v are adjacent in G depends only on the colors of u and v in T, and the color of the lowest common ancestor of u and v in T_G.

To avoid ambiguity, we say G has *vertices* while T has *nodes*. Note that we can assume without loss of generality that all leaves are at the same distance from the root in T. A class $\mathcal{G}$ has *bounded shrubdepth* if there are some $d, k \in \mathbb{N}$ such that every $G \in \mathcal{G}$ has a connection model of depth d with colors in $[k]$ (we recall that the depth of a rooted tree T is the maximum number of vertices on a root-to leaf path in T).

▶ **Lemma A.7.** *Any class $\mathcal{G}$ of bounded shrub-depth admits a constant-size equality-based adjacency labelling scheme.*

Proof. Let d, k be such that any graph $G \in \mathcal{G}$ has a connection model T_G of depth d using color set $[k]$. We denote by $\varphi_G : [k]^3 \to \{0,1\}$ the function such that if u has color a, v has color b, and the lowest common ancestor of u and v has color c in T_G, then u and v are adjacent in G if and only if $\varphi_G(a,b,c) = 1$. For every node u of T_G, write $\chi(u)$ for the color of u in the connection model.

We now construct our equality-based labels for G. For any vertex x, let $t_1(x), t_2(x), \dots, t_d(x)$ be the leaf-to-root path for x, where $t_1(x) = x$ and $t_d(x)$ is the root of T_G. Then the label for x is the sequence $(\varphi_G \mid -), (\chi(t_1(x)) \mid t_1(x)), \dots, (\chi(t_d(x)) \mid t_d(x))$.

On inputs

$$(\varphi_G \mid -), (\chi(t_1(x)) \mid t_1(x)), \dots, (\chi(t_d(x)) \mid t_d(x)),$$
$$(\varphi_G \mid -), (\chi(t_1(y)) \mid t_1(y)), \dots, (\chi(t_d(y)) \mid t_d(y)),$$

the decoder operates as follows. It finds the smallest $i \in [d]$ such that $\mathbf{1}[t_i(x) = t_i(y)]$ and outputs $\varphi_G(\chi(t_1(x)), \chi(t_1(y)), \chi(t_i(x)))$.

The correctness of this labelling scheme follows from the fact that we will have $t_i(x) = t_i(y)$ if and only if the node $t_i(x) = t_i(y)$ is an ancestor of both x and y in T_G, so the smallest $i \in [d]$ such that $t_i(x) = t_i(y)$ identifies the lowest common ancestor of x and y in T_G. ◀

A.2.2 Structurally Bounded Expansion Implies First-Order Sketching

Following [26], we say that a class of graphs has *structurally bounded expansion* if it can be obtained from a class of bounded expansion by first-order (FO) transductions. We omit the precise definition of FO transductions in this paper, as they are not necessary to our discussion, and instead refer the reader to [26]. We just note that a particular case of FO transduction is the notion of *FO interpretation*, which is of specific interest to us. Consider an FO formula $\phi(x,y)$ with two free variables and relational vocabulary $\Sigma = \{F, R_1, \dots, R_k\}$ where F is symmetric of arity 2. We will say that a graph class $\mathcal{F}'$ is an FO interpretation of a graph class $\mathcal{F}$ with respect to ϕ if for any graph $G' = (V, E') \in \mathcal{F}'$ there is a graph $G = (V, E) \in \mathcal{F}$ and a Σ-structure with domain V where E is the interpretation of the symbol F, such that for any pair $u, v \in V$, $uv \in E'$ if and only if $\phi(u/x, v/y)$ evaluates to true. For instance, if $\phi(u/x, v/y)$ encodes the property $\mathsf{dist}_G(u,v) \le r$ for some fixed integer $r \ge 1$ (which can be written as an FO formula), then the corresponding FO interpretation of the class $\mathcal{F}$ is the class of all graph powers $\{G^r \mid G \in F\}$. FO transductions are slightly more involved, as it is allowed to consider a bounded number of copies of a graph before applying the formula, and then it is possible to delete vertices. We will use the following structural result for classes of structurally bounded expansion, proved in [26].

▶ **Theorem A.8** ([26]). *A class $\mathcal{G}$ of graphs has structurally bounded expansion if and only if the following condition holds. For every $p \in \mathbb{N}$, there is a constant $m = m(p)$ such that for every graph $G \in \mathcal{G}$, one can find a family $\mathcal{F}(G)$ of vertex subsets of G with $|\mathcal{F}(G)| \le m$ and the following properties:*
- *for every $X \subseteq V(G)$ with $|X| \le p$, there is $A \in \mathcal{F}(G)$ such that $X \subseteq A$; and*
- *the class $\{G[A] \mid G \in \mathcal{G}, A \in \mathcal{F}(G)\}$ of induced subgraphs has bounded shrubdepth.*

We directly deduce the following result.

▶ **Lemma A.9.** *Any class $\mathcal{G}$ of structurally bounded expansion admits a constant-size equality-based adjacency labelling scheme.*

Proof. Let m and $\mathcal{F}$ be given by applying Theorem A.8 to $\mathcal{G}$ with $p = 2$. By definition, for every graph $G \in \mathcal{G}$ and every pair of vertices $u, v \in V(G)$, there is a set $A \in \mathcal{F}(G)$ containing u and v. Moreover, $\mathcal{F}(G)$ contains at most m sets and the family $\mathcal{C}$ of all graphs $G[A]$, for

$G \in \mathcal{G}$, and $A \in \mathcal{F}(G)$, has bounded shrubdepth. It follows from Lemma A.7 that there is a constant-size equality-based adjacency labelling scheme for $\mathcal{C}$. We denote the decoder of this scheme by D, and the corresponding labels as $\ell'_{G[A]}$.

Consider some graph $G \in \mathcal{G}$, and let $\mathcal{F}(G) = \{A_1, \dots, A_m\}$. For each vertex x of G and $i \in [m]$, we write $a(x) = (a_1(x), \dots, a_m(x))$ where $a_i(x) = \mathbf{1}[x \in A_i]$. Then we define the label for x by taking the prefix $a(x)$ and appending the labels $\ell'_{G[A_i]}(x)$ for each induced subgraph $G[A_i] \in \mathcal{C}$ to which x belongs. Given the labels for vertices x and y, the decoder finds any $i \in [m]$ such that $a_i(x) = a_i(y) = 1$; and outputs $D'(\ell'_{G[A_i]}(x), \ell'_{G[A_i]}(y))$. Such a number $i \in [m]$ always exists due to Theorem A.8. The correctness of this labelling scheme follows from Theorem A.8 and Lemma A.7. ◀

Since FO-transductions compose (see e.g. [59]), sketching FO formulas in a class of structurally bounded expansion is equivalent to sketching adjacency in another class of structurally bounded expansion. We obtain the following direct corollary of Theorem A.9.

▶ **Corollary A.10.** *Any class $\mathcal{G}$ of structurally bounded expansion is first-order sketchable.*

As the property $\mathsf{dist}_G(x,y) \le r$ can be written as an FO formula, this directly implies that classes of bounded expansion are small-distance sketchable. However, this does not tell anything on the size of the sketches as a function of r, unlike the approach using weak coloring numbers described in the next section.

A.3 Bounded Expansion Implies Small-Distance Sketching

Recall the definition of weak reachability from Definition A.2. We give a quantitative bound on the small-distance sketch of any graph class $\mathcal{F}$ in terms of $\mathrm{wcol}_r(\mathcal{F})$. Recall from Theorem A.5 that any class with bounded expansion has $\mathrm{wcol}_r(\mathcal{F}) \le f(r)$ for some function $f(r)$; therefore we obtain the existence of small-distance sketches for any class of bounded expansion.

▶ **Theorem A.11.** *For any $r \in \mathbb{N}$, any class $\mathcal{F}$ has an $(0, r, \mathrm{wcol}_r(\mathcal{F}))$-disjunctive distance-$(r,r)$ labelling scheme.*

Proof. Let $G \in \mathcal{F}$, and consider a total order $(V, \prec)$ such that for any vertex $x \in V$, at most $\mathrm{wcol}_r(\mathcal{F})$ vertices are weakly r-reachable from v in G with respect to $(V, \prec)$. We say that vertex $y \in V$ has *x-rank k* if y is weakly k-reachable from x but not weakly $(k-1)$-reachable from x. For each vertex x and $k \in [r]$, write $S_k(x)$ for the set of vertices y with x-rank k.

We construct a disjunctive labelling scheme as follows. Each vertex x is assigned the label

$$(- \mid \vec{q}_1(x)), (- \mid \vec{q}_2(x)), \dots, (- \mid \vec{q}_{r'}(x))$$

where $r' \le r$ is the maximum number such that $S_{r'}(x) \ne \emptyset$, and the equality codes $\vec{q}_i(x)$ are names of vertices in the set $S_i(x)$. Each label contains at most $\mathrm{wcol}_r(G)$ equality codes, plus a constant number of bits per equality code and $O(r)$ bits to separate the elements of the list. Given labels for x and y, the decoder outputs 1 if and only if there exist $0 \le i, j \le r$ such that $i + j \le r$ and $S_i(x) \cap S_j(y) \ne \emptyset$, which can be checked using the equality codes in $\vec{q}_i(x)$ and $\vec{q}_j(y)$.

Suppose that $\mathsf{dist}_G(x,y) \le r$ and let $P \subseteq V(G)$ be a path of length $\mathsf{dist}_G(x,y)$. Let $z \in P$ be the minimal element of P with respect to $\prec$. Then z is weakly i-reachable from x and weakly j-reachable from y, for some values i, j such that $i + j \le r$. Then $z \in S_i(x) \cap S_j(y)$, so the decoder will output 1 given the labels for x and y. On the other hand, if the decoder outputs 1, then there are values i, j such that $i + j \le r$ and $S_i(x) \cap S_j(y) \ne \emptyset$. Let $z \in S_i(x) \cap S_j(y)$, so that z is weakly i-reachable from x and weakly j-reachable from y. Then $\mathsf{dist}_G(x,y) \le \mathsf{dist}_G(x,z) + \mathsf{dist}_G(z,y) \le i + j \le r$. ◀

We noticed after proving this result that a similar idea was used in [34, Lemma 6.10] to obtain sparse neighborhood covers in nowhere-dense classes.

We will need the following quantitative results for planar graphs and graphs avoiding some specific minor, due to [66].

▶ **Theorem A.12** ([66]). *For any planar graph G, and any integer $r \geq 0$,* $\mathrm{wcol}_r(G) \leq (2r+1)\binom{r+2}{2} = O(r^3)$.

▶ **Theorem A.13** ([66]). *For any integer $t \geq 3$, any graph G with no K_t-minor, and any integer $r \geq 0$,* $\mathrm{wcol}_r(G) \leq \binom{r+t-2}{t-2}(t-3)(2r+2) = O(r^{t-1})$.

In the proof of Theorem A.11, the equality codes are just the names of vertices; so we can use $\lceil \log n \rceil$ bits to encode each of the $\mathrm{wcol}_r(\mathcal{F})$ equality codes to obtain an adjacency label. Then, combined with Proposition 2.5, we obtain the following corollary:

▶ **Corollary A.14.** *If a class $\mathcal{F}$ has bounded expansion, then $\mathcal{F}$ has a small-distance sketch of size at most $O(r + \mathrm{wcol}_r(\mathcal{F}) \log(\mathrm{wcol}_r(\mathcal{F})))$. If $\mathcal{F}$ is the class of planar graphs, then the sketch has size $O(r^3 \log r)$ and if $\mathcal{F}$ is the class of K_t-minor free graphs for some fixed integer $t \geq 3$, then the sketch has size $O(r^{t-1} \log r)$. Furthermore, $\mathcal{F}$ admits a distance-(r,r) labelling scheme of size $O(r + \mathrm{wcol}_r(\mathcal{F}) \log n)$; if $\mathcal{F}$ is the class of planar graphs, then the scheme has size $O(r^3 \log n)$ and if $\mathcal{F}$ is the class of K_t-minor free graphs, then the scheme has size $O(r^{t-1} \log n)$.*

▶ Remark A.15. The fact that the sketch size is independent of the number of vertices in Corollary A.14 implies that the scheme actually works for infinite graphs. It was proved in [43] that for infinite graphs G, $\mathrm{wcol}_r(G)$ is the supremum of $\mathrm{wcol}_r(H)$ for all finite subgraphs H of G (this was actually proved explicitly for the strong coloring numbers instead of the weak coloring numbers, but the proof is the same). This shows that Theorems A.12 and A.13, and thus Corollary A.14, also hold for infinite graphs.

A.4 Small-Distance Sketching Implies Bounded Expansion

To complete the proof of Theorem A.1, we must show that any monotone class of graphs that is small-distance sketchable has bounded expansion, which we do by contrapositive. In fact, we will prove a stronger statement: even having a weaker $(r, 5r-1)$-distance sketch of size $f(r)$ implies bounded expansion.

▶ **Theorem A.16.** *Let $\mathcal{F}$ be a monotone class of graphs and assume that there is a function f such that for any $r \geq 1$, $\mathcal{F}$ has a $(r, 5r-1)$-distance sketch of size $f(r)$. Then $\mathcal{F}$ has bounded expansion.*

Proof. Assume for the sake of contradiction that $\mathcal{F}$ has unbounded expansion. By Corollary A.6, there is a constant k such that for every $d \geq 0$, $\mathcal{F}$ contains a k-subdivision of some bipartite graph $G = (V, E)$ of minimum degree at least d and girth at least 6. Let $\mathcal{G}$ be the class consisting of the graph G, together with all its spanning subgraphs. By monotonicity, $\mathcal{F}$ contains k-subdivisions of all the graphs of $\mathcal{G}$.

Recall the definition of the partial function adj^E parameterized by graphs $H \in \mathcal{G}$, from the discussion preceding Lemma 3.4. We will show that the $(k+1, 5(k+1)-1)$-distance sketch of size $f(k+1)$ for $\mathcal{F}$ can be used to obtain a adj^E-sketch for $\mathcal{G}$, which must have size $\Omega(\log d)$ due to Lemma 3.4. This is a contradiction since we must have $f(k) = \Omega(\log d)$ for arbitrarily large d, whereas $f(k+1)$ is a constant independent of d.

Let H be any spanning subgraph of G and let $H^{(k)}$ denote the k-subdivision of H. Consider two vertices $u, v \in V(H) \subseteq V(G)$ that are adjacent in G. Observe that $\mathsf{dist}_{H^{(k)}}(u,v) = (k+1)\mathsf{dist}_H(u,v)$, and thus if u, v are adjacent in H then $\mathsf{dist}_{H^{(k)}}(u,v) \leq k+1$. Assume now that u, v are non-adjacent in H. Since u, v are adjacent in G, G has girth at least 6, and H is a spanning subgraph of G, it follows that in this case $\mathsf{dist}_H(u,v) \geq 5$, and thus $\mathsf{dist}_{H^{(k)}}(u,v) \geq 5(k+1)$. Therefore, by using the same decoder as the $(k+1, 5(k+1)-1)$-distance sketch for $\mathcal{F}$, and using the random sketch sk defined for G, we obtain an adj_H^E-sketch for H. This gives an adj^E-sketch for $\mathcal{G}$ of size $f(k+1)$. ◀

In our proof of Theorem A.16 we have used Corollary A.6, which is based on the result of [50], stating that every graph of large minimum degree contains a bipartite subgraph of girth at least 6 and large minimum degree. The following stronger statement was conjectured by Thomassen [64].

▶ **Conjecture A.17** ([64])**.** *For every integer k, every graph of sufficiently large minimum degree contains a bipartite subgraph of girth at least k and large minimum degree.*

If Conjecture A.17 is true, it readily follows from our proof that the constant 5 in Theorem A.16 can be replaced by an arbitrarily large constant.