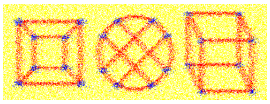
Journal of Graph Algorithms and Applications
http://jgaa.info/ vol. 27, no. 3, pp. 219–238 (2023)
DOI: 10.7155/jgaa.00621

Improved (In-)Approximability Bounds for d-Scattered Set

Ioannis Katsikarelis *Michael Lampis* *Vangelis Th. Paschos*

Université Paris-Dauphine, PSL Research University, CNRS, UMR 7243 LAMSADE, 75016 Paris, France

Submitted:	June 2020	Reviewed:	February 2021	Revised:	February 2021
Accepted:	April 2023	Final:	April 2023	Published:	May 2023
Article type:	Regular paper			Communicated by:	G. Even

Abstract. In the d-Scattered Set problem we are asked to select at least k vertices of a given graph, so that the distance between any pair is at least d. We study the problem's (in-)approximability and offer improvements and extensions of known results for Independent Set, of which it is a generalization. Specifically, we show:

- A lower bound of $\Delta^{\lfloor d/2 \rfloor - \epsilon}$ on the approximation ratio of any polynomial-time algorithm for graphs of maximum degree Δ and an improved upper bound of $O(\Delta^{\lfloor d/2 \rfloor})$ on the approximation ratio of any greedy scheme for this problem.
- A polynomial-time $2\sqrt{n}$-approximation for bipartite graphs and even values of d, that matches the known lower bound by considering the only remaining case.
- A lower bound on the complexity of any ρ-approximation algorithm of (roughly) $2^{\frac{n^{1-\epsilon}}{\rho d}}$ for even d and $2^{\frac{n^{1-\epsilon}}{\rho(d+\rho)}}$ for odd d (under the randomized ETH), complemented by ρ-approximation algorithms of running-times that (almost) match these bounds.

1 Introduction

In this paper we study the d-Scattered Set problem: given graph $G = (V, E)$, we are asked if there exists a set K of at least k *selections* from V, such that the shortest-path distance between any pair $v, u \in K$ is at least d. We focus on the corresponding optimization version where we seek to maximize the number of selections. The problem can already be seen to be hard as it generalizes Independent Set (for $d = 2$) and thus the optimal k cannot be approximated to

Work partially supported by ANR projects ANR-21-CE48-0022 (S-EX-AP-PE-AL) and ANR-18-CE40-0025-01 (ASSK).

E-mail addresses: ioannis.katsikarelis@lamsade.dauphine.fr (Ioannis Katsikarelis) michail.lampis@lamsade.dauphine.fr (Michael Lampis) paschos@lamsade.dauphine.fr (Vangelis Th. Paschos)

$n^{1-\epsilon}$ in polynomial time [19] (under standard complexity assumptions). An alternative name is Distance-d Independent Set [12, 25].

The problem has been well-studied, also from the parameterized point of view [23, 26], while approximability in polynomial time has already been considered for bipartite, regular and degree-bounded graphs [13, 12], perhaps the natural candidate for the next intractability frontier. This paper aims to advance our understanding in this direction by providing the first lower bound on the approximation ratio of any polynomial-time algorithm as a function of the maximum degree of any vertex in the input graph, while also improving upon the known ratios to match this lower bound. On bipartite graphs, our aim is to complete the picture by considering the only remaining open case for this class, before turning our attention to super-polynomial running-times with the purpose of extending known upper/lower bounds for Independent Set.

Before moving on to describe our results in detail, we note that these may be dependent on the parity of our distance parameter d as being even or odd. Both our running-times and ratios can be affected by this peculiarity of the problem that, intuitively, arises due to the (non)existence of a *middle* vertex on a path of length d between two endpoints: if d is even then such a vertex can exist at equal distance $d/2$ from any number of vertices in the solution, while if d is odd there can be no vertex at equal distance from any pair of vertices in the solution. This idiosyncrasy can change the way in which both our algorithms and hardness constructions work and in some cases even entirely alters the problem's complexity (e.g. in the results of [12]).

Our contribution: Section 3 concerns itself with strictly polynomial running-times. We first show that there is no polynomial-time approximation algorithm for d-Scattered Set with ratio $\Delta^{\lfloor d/2 \rfloor - \epsilon}$ in graphs of maximum degree Δ. Our complexity assumption is NP$\not\subseteq$BPP due to our use as a starting point of a randomized construction for Independent Set by [8], that we then build upon to produce highly efficient (in terms of maximum vertex degree and diameter) instances of d-Scattered Set. This is the first lower bound that considers Δ and generalizes the known $\Delta^{1-\epsilon}$-inapproximability of Independent Set (see Theorem 5.2 of [8], restated here as Theorem 1, as well as [1]). Maximum vertex degree Δ plays an important role in the context of independence (e.g. [3, 11, 18]) and was specifically studied for d-Scattered Set in [13], where polynomial-time $O(\Delta^{d-1})$- and $O(\Delta^{d-2}/d)$-approximations are given. We improve upon these upper bounds by showing that any degree-based greedy approximation algorithm in fact achieves a ratio of $O(\Delta^{\lfloor d/2 \rfloor})$, also matching our lower bound.

We then turn our attention to bipartite graphs and show that d-Scattered Set can be approximated within a factor of $2\sqrt{n}$ in polynomial time also for even values of d, matching its known $n^{1/2-\epsilon}$-inapproximability from [12] and complementing the known $\sqrt{n}$-approximation for odd values of d from [17].

Section 4 follows this up by considering super-polynomial running-times, presenting first an exact exponential-time algorithm for d-Scattered Set of complexity $O^*((ed)^{\frac{2n}{d}})$ based on a straightforward upper bound on the size of any solution and then considering the inapproximability of the problem in the same complexity range. We show that no ρ-approximation algorithm can take time (roughly) $2^{\frac{n^{1-\epsilon}}{\rho d}}$ for even d and $2^{\frac{n^{1-\epsilon}}{\rho(d+\rho)}}$ for odd d, under the (randomized) ETH. This is complemented by (almost) matching ρ-approximation algorithms of running-times $O^*((e\rho d)^{\frac{2n}{\rho d}})$ for even d and $O^*((e\rho d)^{\frac{2n}{\rho(d+\rho)}})$ for odd d. We note that the current state-of-the-art PCPs are unable to distinguish between optimal running-times of the form $2^{n/\rho}$ and $\rho^{n/\rho}$ for ρ-approximation algorithms, due to the poly-logarithmic factor added by even the most efficient constructions and we thus do not focus on such factors differentiating our upper and lower bounds.

These results provide a complete characterization of the optimal relationship between the worst-case approximation ratio ρ achievable for d-Scattered Set by any algorithm, its running-time and the distance parameter d, *for any point in the trade-off curve*, in a similar manner as was done for Independent Set in [8, 10] (see also [5, 6]), by also considering the range of possible values for d. We observe that the distance parameter d acts as a scaling factor for the size of the instance, whereby the problem becomes easier when vertices are required to be much further apart, a feature counterbalanced by the chosen approximation ratio ρ, with small values guaranteeing the quality of the produced solutions, yet also negatively impacting on the exponent of the running-time.

We close the paper with a supplementary note on the treewidth of power graphs obtained through observations related to our previous results, while Section 5 provides some concluding remarks and discussion on open problems. Our results are also summarized in Table 1 below.

	Inapproximability	Approximation
Super-polynomial	$2^{\frac{n^{1-\epsilon}}{\rho d}}$ (8) $2^{\frac{n^{1-\epsilon}}{\rho(d+\rho)}}$ (9)	$O^*((e\rho d)^{\frac{2n}{\rho d}})$ (10) $O^*((e\rho d)^{\frac{2n}{\rho(d+\rho)}})$ (11)
Polynomial	$\Delta^{\lfloor d/2 \rfloor - \epsilon}$ (2)	$O(\Delta^{\lfloor d/2 \rfloor})$ (4)
Bipartite graphs	$n^{1/2-\epsilon}$ [12]	$2\sqrt{n}$ (6)

Table 1: A summary of our results (theorem numbers), for even/odd values of d.

Related work: Eto et al. ([13]) showed that on r-regular graphs the problem is APX-hard for $r, d \geq 3$, while also providing polynomial-time $O(r^{d-1})$-approximations. They also show a polynomial-time 2-approximation on cubic graphs and a polynomial-time approximation scheme (PTAS) for planar graphs and every fixed constant $d \geq 3$. For a class of graphs with at most a polynomial (in n) number of minimal separators, d-Scattered Set can be solved in polynomial time for even d, while it remains NP-hard on chordal graphs and any odd $d \geq 3$ [25]. For bipartite graphs, the problem is NP-hard to approximate within a factor of $n^{1/2-\epsilon}$ and W[1]-hard for any fixed $d \geq 3$. Further, for any odd $d \geq 3$, it remains NP-complete, inapproximable and W[1]-hard [12]. It is NP-hard even for planar bipartite graphs of maximum degree 3, yet a 1.875-approximation is available on cubic graphs [14]. Furthermore, [15] shows the problem admits an EPTAS on (apex)-minor-free graphs, based on the theory of bidimensionality, while on a related result [24] offers an $n^{O(\sqrt{n})}$-time algorithm for planar graphs, making use of Voronoi diagrams and based on ideas previously used to obtain geometric QPTASs. Finally, [23] presents tight upper/lower bounds on the structurally parameterized complexity of the problem, while [26] shows that it admits an almost linear kernel on every nowhere dense graph class.

2 Definitions and Preliminaries

We use standard graph-theoretic notation. For a graph $G = (V, E)$, we let $V(G) \coloneqq V$ and $E(G) \coloneqq E$, an edge $e \in E$ between $u, v \in V$ is denoted by $\{u, v\}$ and for a subset $X \subseteq V$, $G[X]$ denotes the graph induced by X. We let $dst_G(v, u)$ denote the shortest-path distance (i.e. the number of edges) from v to u in G. We may omit subscript G if it is clear from the context. The maximum distance between vertices is the *diameter* of the graph, while the minimum among all the maximum distances between a vertex to all other vertices (their *eccentricities*) is the graph's *radius*.

For a vertex v, we let $N_G^d(v)$ denote the (open) *d-neighborhood* of v in G, being the set of vertices at distance $\leq d$ from v in G (without v), while for a subset $U \subseteq V$, $N_G^d(U)$ denotes the union of the d-neighborhoods of vertices $u \in U$. In a graph G whose maximum degree is bounded by Δ, the size of the d-neighborhood of any vertex v is upper bounded by the well-known *Moore bound*: $|N_G^d(v)| \leq \Delta \sum_{i=0}^{d-1}(\Delta-1)^i$.

For Σ being a set of δ elements (the alphabet) and d a positive integer, the *Hamming graph* $H(d,\delta)$ has vertex set Σ^d, the set of ordered d-tuples of elements of Σ, each vertex thus corresponding to one sequence of length d of symbols from Σ (a *word*). Two vertices are adjacent if their words differ in precisely one coordinate (if their Hamming distance is one). The Hamming graph $H(d,\delta)$ can equivalently be seen as the Cartesian product of d complete graphs K_δ and is also related to Hypercube graphs, that are similarly defined for $\delta = 2$ (over a binary alphabet). A graph $H(d,\delta)$ constructed in this way has δ^d vertices, diameter equal to d and is $d(\delta-1)$-regular (see [7, Chapter 12]).

For an integer q, the *q-th power graph of* G, denoted by G^q, is defined as the graph obtained from G by adding to $E(G)$ all edges between vertices $v, u \in V(G)$ for which $dst_G(v,u) \leq q$. Furthermore, we let $OPT_d(G)$ denote the maximum size of a d-scattered set in G and $\alpha(G) = OPT_2(G)$ denote the size of the largest independent set.

We use $\log(n)$ to denote the base-2 logarithm of n, while $\log_{1+\delta}(n)$ is the logarithm base-$(1+\delta)$, for $\delta > 0$. Recall also that $\log_{1+\delta}(n) = \log(n)/\log(1+\delta)$. The functions $\lfloor x \rfloor$ and $\lceil x \rceil$, for $x \in \mathbb{R}$, denote the maximum integer that is not larger and the minimum integer that is not smaller than x, respectively.

The Exponential Time Hypothesis (ETH) [20, 21] implies that 3-SAT cannot be solved in time $2^{o(n)}$ on instances with n variables (a slightly weaker statement), while the definition can also refer to *randomized* algorithms. Finally, we recall here the following result by [8] that some of our reductions will be relying on (slightly paraphrased, see also [5]), that can furthermore be seen as implying the $\Delta^{1-\epsilon}$-inapproximability of Independent Set in polynomial time:

Theorem 1 ([8], Theorem 5.2) *For any sufficiently small $\epsilon > 0$ and any $r \leq N^{5+O(\epsilon)}$, there is a randomized polynomial-time reduction that builds from a formula ϕ of* SAT *on N variables a graph G of size $n = N^{1+\epsilon}r^{1+\epsilon}$ and maximum degree r, such that with high probability:*

- *If ϕ is satisfiable, then $\alpha(G) \geq N^{1+\epsilon}r$.*
- *If ϕ is not satisfiable, then $\alpha(G) \leq N^{1+\epsilon}r^{2\epsilon}$.*

3 Polynomial Time

We begin by focusing on the behaviour of the problem in the context of strictly polynomial-time approximation. We first examine graphs of bounded degree and provide a tight bound on the achievable approximation ratio, before turning to bipartite graphs in order to finalize the classification in terms of approximability by considering the only open remaining case (when d is even).

3.1 Inapproximability

We show that for sufficiently large Δ and any $\epsilon_1 > 0, d \geq 4$, the d-Scattered Set problem is inapproximable to $\Delta^{\lfloor d/2 \rfloor - \epsilon_1}$ on graphs of degree bounded by Δ, unless NP$\subseteq$BPP. Let us first summarize our reduction. Starting from an instance of Independent Set of bounded degree, we create

an instance of d-Scattered Set where the degree is (roughly) the $d/2$-th square root of that of the original instance. As we are able to maintain a direct correspondence of solutions in both instances, the $\Delta^{1-\epsilon'}$-inapproximability of Independent Set implies the $\Delta^{\lfloor d/2\rfloor-\epsilon_1}$-inapproximability of d-Scattered Set.

The technical part of our reduction involves preserving the adjacency between vertices of the original graph without increasing the maximum degree (too far) beyond $\Delta^{2/d}$. We are able to construct a regular tree as a gadget for each vertex and let the edges of the leaves (their total number being equal to Δ) represent the edges of the original graph. To ensure that our gadget has some useful properties (i.e. small diameter), we overlay a number of extra edges on each level of the tree (i.e. between vertices at equal distance from the root), only sacrificing a small increase in maximum degree.[1] Our complexity assumption is NP$\not\subseteq$BPP, since for the $\Delta^{1-\epsilon'}$-inapproximability of IS we use the randomized reduction from SAT of [8] (Theorem 1 above). In particular, we will prove the following theorem:

Theorem 2 *For sufficiently large Δ and any $d \geq 4, \epsilon \in (0, \lfloor d/2\rfloor)$, there is no polynomial-time approximation algorithm for d-Scattered Set with ratio $\Delta^{\lfloor d/2\rfloor-\epsilon}$ for graphs of maximum degree Δ, unless NP$\subseteq$BPP.*

Construction: Let $\delta = \left\lceil \sqrt[\lfloor d/2\rfloor]{\Delta} \right\rceil$. Given $\epsilon_1 \in (0, \lfloor d/2\rfloor)$ and an instance of Independent Set $G = (V, E)$, where the degree of any vertex is bounded by Δ, we will construct an instance $G' = (V', E')$ of d-Scattered Set, where $OPT_2(G) = OPT_d(G')$, while the degree is bounded by $2\lfloor \frac{d}{2}\rfloor(\delta-1)+2$ for odd d and by $\frac{d}{2}(\delta-1)+2$ for even d. We thus will assume Δ is sufficiently large for $\epsilon_1 > \frac{\lfloor \frac{d}{2}\rfloor^2}{\log_d(\Delta)+\lfloor \frac{d}{2}\rfloor}$, for reasons that become apparent in the subsequent proof of Theorem 2.

Our construction for G' builds a gadget $T(v)$ for each vertex $v \in V$. For even d, each gadget $T(v)$ is composed of a $(\delta+1)$-regular tree of height $d/2-1$ and we refer to vertices of $T(v)$ at distance exactly i from the root t_v as being in the i-th *height-level* of $T(v)$, letting each such subset be denoted by $T_i(v)$. That is, every vertex of $T_i(v)$ has one neighbor in $T_{i-1}(v)$ (its parent) and δ neighbors in $T_{i+1}(v)$ (its children). We will also add extra edges on each height-level, described below. For odd values of d, the difference is in the height of each tree being $\lfloor d/2\rfloor$ instead of $d/2-1$.

Since for even d the number of leaves of $T(v)$ is $\delta^{d/2-1} \geq (\Delta^{2/d})^{d/2-1} = \Delta^{1-2/d}$ and each such leaf also has $\delta \geq \Delta^{2/d}$ edges, the number of edges leading outside each gadget is $\delta^{d/2} \geq \Delta$ and we let each of them correspond to one edge of the original vertex v in G, meaning there is an edge between a leaf x_v of $T(v)$ and a leaf y_u of $T(u)$, if $\{v, u\} \in E$. For odd d, the number of leaves is $\left(\left\lceil \sqrt[\lfloor d/2\rfloor]{\Delta} \right\rceil\right)^{\lfloor d/2\rfloor} \geq \Delta$ and we let each leaf correspond to an edge of the original vertex v in G, i.e. we *identify* two such leaves x_v, y_u of two gadgets $T(v), T(u)$, if $\{v, u\} \in E$ in G. In this way, the gadgets $T(v), T(u)$ share a common "leaf" of degree 2 (thus far), that is at distance $\lfloor d/2\rfloor$ from both roots $t_v \in T(v), t_u \in T(u)$. See Figure 1 for an illustration.

Next, in order to make the diameter of our gadgets at most equal to their height, we will add a number of edges between the vertices $T_i(v)$ of each height-level i of every gadget $T(v)$, for every $v \in V$. These will be the edges of the Hamming graph $H(i, \delta)$ for each height-level $i \leq d/2-1$ for even d and each height-level $i \leq \lfloor d/2\rfloor$ for odd d. The number of vertices of $T_i(v)$ at every height-level i is already equal to δ^i and thus a one-to-one correspondence between them and the words of length i with symbols from an alphabet of size δ is straightforward. The edges we add will

[1] The original version of this paper [22] made use of a randomized technique from [4] to obtain degree-diameter efficiency, but we give here an improved, deterministic gadget.

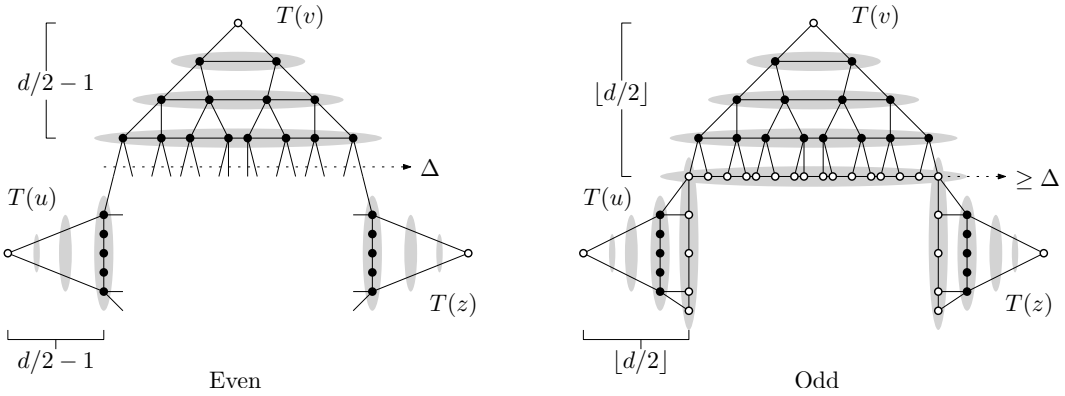

Figure 1: Our constructions for an example subgraph consisting of a path on three vertices (u, v, z) and even/odd d. Ellipses in grey designate the overlaid edges on each height-level.

connect vertices $x_i, y_i \in T_i(v)$ if the corresponding words differ in precisely one coordinate (their Hamming distance is 1). In this way, the edges already constructed above as part of the tree will now be seen to connect vertices x_i, x_{i+1} between Hamming graphs $H(i, \delta)$ and $H(i+1, \delta)$ such that the words corresponding to x_i, x_{i+1} are of lengths that differ by 1 and the word of x_i (the parent) is a prefix of the word of x_{i+1} (the child). This concludes our construction, while Figure 2 illustrates a small example.

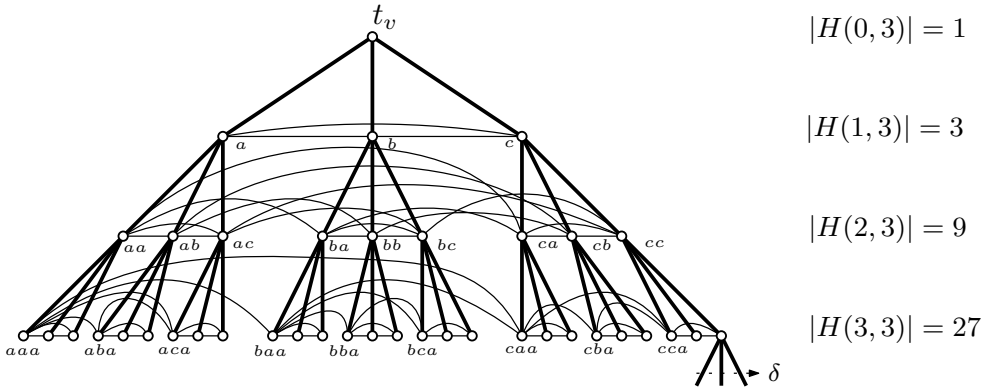

Figure 2: The upper height-levels of an example gadget $T(v)$, where the edges of the tree are shown in bold. The words of vertices for alphabet $\Sigma = \{a, b, c\}$ are shown for height-levels $i = 1, 2$, along with their edges, but only for the first vertices of each group in level $i = 3$. Observe that every $T_i(v)$ is composed of δ copies of $T_{i-1}(v)$ with edges between copies according to one added coordinate.

We are now ready to argue about the maximum degree of any vertex in the instances built by our construction.

Lemma 1 *The maximum degree of any vertex in G' is $\frac{d}{2}(\delta-1)+2$ for even d and $2\lfloor\frac{d}{2}\rfloor(\delta-1)+2$ for odd d.*

Proof: For even d, the degree of any vertex is bounded by the sum of the $\delta+1$ edges of the tree plus the number of edges added by the Hamming graph. Their maximum is $(d/2-1)(\delta-1)$ for height-level $i=d/2-1$, giving a total of $(d/2-1)(\delta-1)+\delta+1=\frac{d}{2}(\delta-1)+2$.

For odd d, the degree of all other vertices is strictly lower than that of the "shared" leaves between gadgets, since each leaf between two gadgets $T(v),T(u)$ (representing the edge $\{v,u\}$ of G) will belong to two Hamming graphs $H(\lfloor\frac{d}{2}\rfloor,\delta)$. Their degree will be $2\lfloor\frac{d}{2}\rfloor(\delta-1)+2$, giving the maximum, while the next highest degree (that of a parent of such a leaf) would be $(\lfloor\frac{d}{2}\rfloor-1)(\delta-1)+\delta+1$. □

We then bound the diameter of our gadgets in order to guarantee that the solutions in our reduction will be well-formed.

Lemma 2 *The diameter of each gadget $T(v)$ is $d/2-1$ for even d and $\lfloor d/2\rfloor$ for odd d.*

Proof: By construction, the diameter of a Hamming graph $H(i,\delta)$ is equal to i. Since at each height-level i, no pair of vertices is at distance $>i$ with $i\leq d/2-1$ for even d and $i\leq\lfloor d/2\rfloor$ for odd d, the distance between any vertex x at some height-level i_x to another vertex y at height-level $i_y>i_x$ will be at most i_x from x to the root of the subtree of $T(v)$ (at level i_x) that contains y. From there to y it will be at most $d/2-1-i_x$ for even d and at most $\lfloor d/2\rfloor-i_x$ for odd d. Furthermore, the distance from the root of $T(v)$ to a leaf is exactly $d/2-1$ for even d and exactly $\lfloor d/2\rfloor$ for odd d. □

We finalize our argument with a series of lemmas leading to the proof of Theorem 2, that detail the behaviour of solutions that can form in our construction, relative to the independence of vertices in the original graph.

Lemma 3 *No d-Scattered Set in G' can contain a vertex from gadget $T(v)$ and a vertex from gadget $T(u)$, if $\{u,v\}\in E$.*

Proof: Since $\{u,v\}\in E$, there is an edge $\{x_v,y_u\}\in E'$ between a leaf $x_v\in T(v)$ and $y_u\in T(u)$ for even d, while for odd d the leaf x belongs to both $T(v),T(u)$ and is at distance $\lfloor d/2\rfloor$ from each of their roots. Thus for even d the maximum distance from any vertex of $T(v)$ to $y_u\in T(u)$ is $d/2-1+1=d/2$, by Lemma 2, and for odd d this is $\lfloor d/2\rfloor$. Since, by the same lemma, the diameter of $T(u)$ is $d/2-1$ for even d and $\lfloor d/2\rfloor$ for odd d, there is no vertex of $T(u)$ that can be in any d-Scattered Set along with any vertex of $T(v)$, as the maximum distance is $\leq d/2+d/2-1=d-1$ for even d and $\leq\lfloor d/2\rfloor+\lfloor d/2\rfloor=d-1$ for odd d. □

Lemma 4 *If $\{u,v\}\notin E$, then the distance between the root t_v of $T(v)$ and the root t_u of $T(u)$ is at least d.*

Proof: Since $\{u,v\}\notin E$, then there is no edge between any pair of leaves x_u of $T(u)$ and y_v of $T(v)$ for even d. Thus the shortest possible distance between any such pair of leaves is 2 for even d, through a third leaf z_w of another gadget $T(w)$ corresponding to a vertex w adjacent to both u and v in G. The distance from t_v to any leaf of $T(v)$ is $d/2-1$ and the distance from t_u to any leaf of $T(u)$ is also $d/2-1$. Thus the distance from t_u to t_v must be at least $d/2-1+d/2-1+2=d$.

For odd d, there is no shared leaf x between the two gadgets, i.e. at distance $\lfloor d/2\rfloor$ from both roots. Thus the distance between two leaves $x_u\in T(u)$ and $y_v\in T(v)$ is at least 1, if each of these is shared with a third gadget $T(w)$ corresponding to a vertex w that is adjacent to both u and v in G. The distance from t_v to any leaf of $T(v)$ is $\lfloor d/2\rfloor$ and the distance from t_u to any leaf of $T(u)$ is also $\lfloor d/2\rfloor$. Thus the distance from t_u to t_v is at least $2\lfloor d/2\rfloor+1=d$. □

Lemma 5 *For any independent set S in G, there is a d-*Scattered Set *K in G', with $|S| = |K|$.*

Proof: Given an independent set S in G, we let K include the root vertex $t_v \in T(v)$ for each $v \in S$. Clearly, $|S| = |K|$. Since S is independent, there is no edge $\{u, v\}$ between any pair $u, v \in S$ and thus, by Lemma 4, vertices t_v and t_u are at distance at least d. □

Lemma 6 *For any d-*Scattered Set *K in G', there is an independent set S in G, with $|K| = |S|$.*

Proof: Given a d-Scattered Set K in G, we know there is at most one vertex from each gadget $T(v)$ in K, since its diameter is $d/2-1$ for even d and $\lfloor d/2 \rfloor$ for odd d, by Lemma 2. Furthermore, for any two vertices $x, y \in K$, we know by Lemma 3 that if $x \in T(u)$ and $y \in T(v)$ for gadgets corresponding to vertices $u, v \in V$, then $\{u, v\} \notin E$ and thus u, v are independent in G. We let set S contain each vertex $v \in V$ whose corresponding gadget $T(v)$ contains a vertex of K. These vertices are all independent and also $|K| = |S|$. □

Proof: [Theorem 2] We suppose the existence of a polynomial-time approximation algorithm for d-Scattered Set with ratio $\Delta^{\lfloor d/2 \rfloor - \epsilon_1}$ for graphs of maximum degree Δ and some $0 < \epsilon_1 < d/2$ and assume Δ is sufficiently large for $\epsilon_1 > \frac{\lfloor \frac{d}{2} \rfloor^2}{\log_d(\Delta) + \lfloor d/2 \rfloor}$. We define $\epsilon_2 = \log_\Delta(d^{\lfloor d/2 \rfloor - \epsilon_1}) = \frac{\lfloor d/2 \rfloor - \epsilon_1}{\log_d(\Delta)}$ and $\epsilon' = \frac{\epsilon_1 - \epsilon_2 \lfloor d/2 \rfloor}{2\lfloor d/2 \rfloor}$. Note that $\epsilon' > 0$ because $\epsilon_1 > \frac{\lfloor \frac{d}{2} \rfloor^2}{\log_d(\Delta) + \lfloor d/2 \rfloor} \Leftrightarrow \frac{\log_d(\Delta)}{\lfloor d/2 \rfloor} + 1 > \frac{\lfloor d/2 \rfloor}{\epsilon_1} \Leftrightarrow \frac{\log_d(\Delta)}{\lfloor d/2 \rfloor} > \frac{\lfloor d/2 \rfloor - \epsilon_1}{\epsilon_1} \Leftrightarrow \epsilon_1 > \frac{\lfloor d/2 \rfloor - \epsilon_1}{\log_d(\Delta)} \lfloor d/2 \rfloor \Leftrightarrow \epsilon_1 > \epsilon_2 \lfloor d/2 \rfloor$.

Now, starting from a formula ϕ of SAT on N variables, where N is also sufficiently large, i.e $N > \Delta^{\frac{1}{5 + O(\epsilon')}}$, we use Theorem 1 to produce an instance $G = (V, E)$ of Independent Set on $|V| = N^{1+\epsilon'}\Delta^{1+\epsilon'}$ vertices and of maximum degree Δ, such that with high probability: if ϕ is satisfiable, then $\alpha(G) \geq N^{1+\epsilon'}\Delta$; if ϕ is not satisfiable, then $\alpha(G) \leq N^{1+\epsilon'}\Delta^{2\epsilon'}$. Thus approximating Independent Set in polynomial time on G within a factor of $\Delta^{1-2\epsilon'}$, for $\epsilon' > 0$, would permit us to decide if ϕ is satisfiable, with high probability.

We next use the above construction to create an instance G' of d-Scattered Set where the degree is bounded by $2\lfloor d/2 \rfloor(\delta - 1) + 2 = 2\lfloor d/2 \rfloor(\lceil \Delta^{\frac{1}{\lfloor d/2 \rfloor}} \rceil - 1) + 2$, by Lemma 1 where the maximum comes from the case where d is odd. To simplify the following, we observe $2\lfloor d/2 \rfloor(\lceil \Delta^{\frac{1}{\lfloor d/2 \rfloor}} \rceil - 1) + 2 < 2\lfloor d/2 \rfloor \cdot \Delta^{\frac{1}{\lfloor d/2 \rfloor}} + 2 = (d-1) \cdot \Delta^{\frac{1}{\lfloor d/2 \rfloor}} + 2 < d \cdot \Delta^{\frac{1}{\lfloor d/2 \rfloor}}$ and we claim this as our degree bound.

We then apply the supposed approximation for d-Scattered Set on G'. This returns a solution at most $(d \cdot \Delta^{\frac{1}{\lfloor d/2 \rfloor}})^{\lfloor d/2 \rfloor - \epsilon_1} = d^{\lfloor d/2 \rfloor - \epsilon_1} \cdot \Delta^{1 - \frac{\epsilon_1}{\lfloor d/2 \rfloor}} = \Delta^{1 - \frac{\epsilon_1}{\lfloor d/2 \rfloor} + \epsilon_2} = \Delta^{1-2\epsilon'}$ from the optimum. By Lemma 6 we can find a solution for Independent Set in G of the same size, meaning we can approximate $\alpha(G)$ within a factor of $\Delta^{1-2\epsilon'}$. This would allow us to decide if ϕ is satisfiable and thus solve SAT in polynomial time with two-sided bounded errors, implying NP⊆BPP. □

3.2 Approximation

We next show that any (degree-based) greedy polynomial-time approximation algorithm for d-Scattered Set achieves a ratio of $O(\Delta^{\lfloor d/2 \rfloor})$, thus improving upon the analysis of [13] and the $O(\Delta^{d-1})$- and $O(\Delta^{d-2}/d)$-approximations given therein.

Our strategy is to bound the size of the largest d-scattered set in any graph of maximum degree at most Δ and radius at most $d-1$, centered on some vertex v. The idea is that in one of its iterations our greedy algorithm would select v and thus exclude all other vertices within distance $d-1$ from v, yet an upper bound on the size of the largest possible d-scattered set can guarantee that the ratio of our algorithm will not be too large.

The following definition of our "merge" operation (see also Figure 3) will allow us to consider all possible graphs of a given radius and degree and provide upper bounds on the size of the optimal solution in such graphs. These bounds on the size of the optimum are then used to compare it to those solutions produced by our greedy scheme.

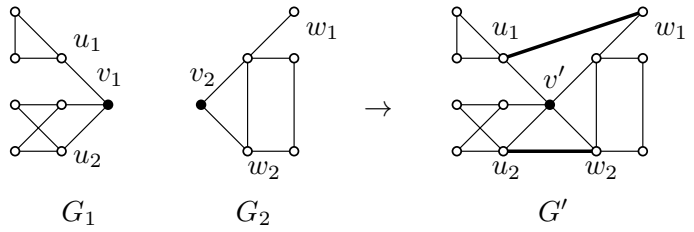

Figure 3: An example graph $G' = M_{G_1}^{G_2}(v_1, v_2, [u_1, u_2], [w_1, w_2])$ for G_1, G_2 shown on the left, with edges added by the third merge operation shown in bold.

Definition 3 (Merge operation) *For graphs $G_1 = (V_1, E_1), G_2 = (V_2, E_2)$, each of which is assumed to be connected, the* merged *graph $M_{G_1}^{G_2}(v_1, v_2, \mathbf{U}, \mathbf{W})$, where $\mathbf{U} = [u_1, \ldots, u_{k_1}]$, $\mathbf{W} = [w_1, \ldots, w_{k_2}]$ are ordered (possibly empty and with repetitions allowed) sequences of vertices from V_1 and V_2, respectively, is defined as the graph $G' = (V', E')$ obtained by:*

1. *Identification of vertex $v_1 \in G_1$ and vertex $v_2 \in G_2$, i.e. V' is composed of the union of V_1, V_2 after removal of vertices v_1, v_2 and addition of a new vertex v'.*
2. *Replacement of all edges of v_1, v_2 by new edges with v' as the new endpoint, i.e. E' contains all edges of E_1, E_2, where any edges incident on v_1 or v_2 are now incident on v'.*
3. *Addition of a number of edges between vertices of G_1, G_2, i.e. E' also contains one edge between every pair $\{u_i, w_j\}$ from $\mathbf{U}, \mathbf{W}$, for $i = j$.*

Lemma 7 *The maximum size of a d-scattered set in any graph of maximum degree at most Δ and radius at most $\lceil d/2 \rceil$ centered on some vertex v is at most Δ.*

Proof: Consider a graph of maximum degree Δ and radius $\lceil d/2 \rceil$ centered on some vertex v: the only pairs of vertices at distance d from each other must be at distance $\geq \lceil d/2 \rceil$ from v (or $\lfloor d/2 \rfloor$ in one side for odd d), as any vertex u at distance $< \lfloor d/2 \rfloor$ from v will be at distance $< d$ from any other vertex z in the graph, since z is at distance $\leq \lceil d/2 \rceil$ from v (due to the graph's radius). Furthermore, for every vertex in the d-scattered set, there must be an edge-disjoint path of length at least $\lceil d/2 \rceil$ to v that is not shared with any other such vertex, i.e. these paths can only share vertex v at distance $\lceil d/2 \rceil$ from their endpoints (the vertices that can be in a d-scattered set). As the degree of v is bounded by Δ, the number of such disjoint paths also cannot be more than Δ. □

Lemma 8 *Given two graphs G_1, G_2, for $G' = M_{G_1}^{G_2}(v_1, v_2, \mathbf{U}, \mathbf{W})$ and any $\mathbf{U}, \mathbf{W}$, it is $OPT_d(G') \leq OPT_d(G_1) + OPT_d(G_2)$.*

Proof: Assume for the sake of contradiction that $OPT_d(G') > OPT_d(G_1) + OPT_d(G_2)$ and let S denote an optimum d-scattered set in G' of this size, with $S_1 = S \cap \{V_1 \setminus \{v_1\} \cup \{v'\}\}$ and $S_2 = S \cap \{V_2 \setminus \{v_2\} \cup \{v'\}\}$ denoting the parts in G_1, G_2, respectively. Since $|S| > OPT_d(G_1) + OPT_d(G_2)$,

then for any pair of optimal $K_1 \subseteq V_1$, $K_2 \subseteq V_2$ in G_1, G_2, there must be at least one vertex s in S for which $s \notin K_1$ and $s \notin K_2$, but it must be $s \in S_1$ or $s \in S_2$ (or both, if $s = v'$).

Observe that the distance between any pair of vertices in G_1 (the same holds for G_2) cannot increase in G' after the merge operation, since identification of a pair of vertices between two graphs and addition of any number of edges between the two can only decrease their distance Thus if $s \in V_1$, then S_1 is also a d-scattered set in G_1 (potentially substituting v' for v_1) and so, if $|S_1| > |K_1|$ then K_1 was not optimal for G_1. If $|S_1| \leq |K_1|$, it must be $|S_2| > |K_2|$ contradicting the optimality of K_2 for G_2. Similarly, if $u \in V_2$ we have either $|S_1| > |K_1|$ or $|S_2| > |K_2|$.

If u is the merged vertex v' then there must be at least two other vertices added from V_1, V_2 for $|S| > |K_1| + |K_2|$, since S can only contain v' in the place of $v_1 \in K_1$ and $v_2 \in K_2$. In this case the same argument as above gives the contradiction. □

Lemma 9 *For any graph $G = (V, E)$ of maximum degree at most Δ and radius at most $d-1$ centered on some vertex v, it is $OPT_d(G) \leq O(\Delta^{\lfloor d/2 \rfloor})$.*

Proof: Any graph G of maximum degree at most Δ and radius at most $d-1$ centered on a vertex v can be obtained by the following process: we begin with a graph H of radius at most $\lfloor d/2 \rfloor - 1$ and maximum degree Δ. Let $\{v_1, \ldots, v_k\} \in H$ be the set of vertices at maximum distance from v, i.e. $dst_H(v, v_i) = \lfloor d/2 \rfloor - 1$. Since the degree of H is bounded by Δ, it must be $k \leq \Delta^{\lfloor d/2 \rfloor - 1}$. We now let H_i, for each $i \leq k$, denote a series of at most k graphs of radius at most $\lceil d/2 \rceil$ centered on a vertex v_i and maximum degree Δ.

Repeatedly applying the merge operation $M_H^{H_i}(v_1, v_i, \mathbf{U}, \mathbf{W})$ between graph H (or the result of the previous operation) and such a graph H_i we can obtain any graph G of radius at most $d-1$: identifying a vertex $v_j \in H$ (for $j \in [1, k]$) at maximum distance from v with the central vertex v_i of H_i and then adding any number of edges between the vertices of H and H_i (while respecting the maximum degree of Δ), we can produce any graph of radius $\leq d-1$, since the distance from v to each v_j is at most $\lfloor d/2 \rfloor - 1$ and from there to any vertex of H_i it is at most $\lceil d/2 \rceil$. The remaining structure of G can be constructed by the chosen structures of the graphs H, H_i and the added edges between them, i.e. the sequences $\mathbf{U}, \mathbf{W}$.

By Lemma 7 it is $OPT_d(H_i) \leq \Delta$ and by Lemma 8, it must be $OPT_d(G) \leq OPT_d(H) + \sum_{i=1}^{k} OPT_d(H_i) \leq 1 + \Delta \cdot \Delta^{\lfloor d/2 \rfloor - 1} \leq 1 + \Delta^{\lfloor d/2 \rfloor}$. □

Theorem 4 *Any degree-based greedy approximation algorithm for d-Scattered Set achieves a ratio of $O(\Delta^{\lfloor d/2 \rfloor})$ on graphs of degree bounded by Δ.*

Proof: Let $G = (V, E)$ be the input graph and consider the process of our supposed greedy algorithm: it picks a vertex v_i, removes it from consideration along with the set $V_i \subseteq V$ of vertices at distance at most $d-1$ from v_i and continues the process until there are no vertices left to consider. The sets $V_1, \ldots, V_{ALG}$ thus form a partition of G. By Lemma 9, the optimum size of a d-scattered set in any such V_i is at most $O(\Delta^{\lfloor d/2 \rfloor})$ and thus $OPT_d(G) \leq ALG \cdot O(\Delta^{\lfloor d/2 \rfloor})$, by Lemma 8, since G can be seen as the merged graph of $G[V_1], \ldots, G[V_{ALG}]$. □

3.3 Bipartite graphs

Here we consider bipartite graphs and show that d-Scattered Set is approximable to $2\sqrt{n}$ in polynomial time also for even values of d. Our algorithm will be applied on both sides of the bipartition and each time it will only consider vertices from one side as candidates for inclusion in the solution. Appropriate sub-instances of Set Packing are then defined and solved using the known $\sqrt{n}$-approximation for that problem.

Definition 5 *For a bipartite graph $G = (A \cup B, E)$, let $1OPT_d(G)$ denote the size of the largest* one-sided *d-scattered set of G, i.e. a set that only includes vertices from the same side of the bipartition A or B, but not both.*

Lemma 10 *For a bipartite graph $G = (A \cup B, E)$, it is $1OPT_d(G) \geq OPT_d(G)/2$.*

Proof: Consider an optimal solution $S \subseteq A \cup B$ with $|S| = OPT_d(G)$. Then at least half of the vertices of S are contained in one side of G, i.e. it is either $|S \cap A| \geq |S|/2$ or $|S \cap B| \geq |S|/2$ (or both if $|S \cap A| = |S \cap B| = |S|/2$). By definition, it is also $1OPT_d(G) \geq |S \cap A|$ and $1OPT_d(G) \geq |S \cap B|$. Thus in both cases it must be $1OPT_d(G) \geq OPT_d(G)/2$. □

Theorem 6 *For any bipartite graph $G = (A \cup B, E)$ of size n and d even, the d-*Scattered Set *problem can be approximated within a factor of $2\sqrt{n}$ in polynomial time.*

Proof: We will consider two cases based on the parity of $d/2$ and define appropriate Set Packing instances whose solutions are in a one-to-one correspondence with one-sided d-scattered sets in G. We will then be able to apply the $\sqrt{n}$-approximation for Set Packing of [17]. We will repeat this process for both sides A, B of the bipartition and retain the best solution found. Thus we will be able to approximate $1OPT_d(G)$ within a factor of $\sqrt{n}$ and then rely on Lemma 10 to obtain the claimed bound.

Our Set Packing instances are defined as follows: for $d/2$ even, we make a set c_i for every vertex a_i of A (i.e. from one side) and an element e_j for every vertex b_j of B (i.e. from the other side). For $d/2$ odd, we make a set c_i for every vertex a_i of A (again from one side) and an element e_j for every vertex b_j of B *and* an element r_i for every vertex $a_i \in A$ (i.e. from both sides). Note that $i, j \leq n$. In both cases we include an element corresponding to vertex $x \in G$ in a set corresponding to a vertex $y \in G$, if $dst_G(x, y) \leq d/2 - 1$. We then claim that for any given collection C of compatible (i.e. non-overlapping) sets in the above definitions, we can always find a one-sided d-scattered set $S \subseteq A$ in G with $|C| = |S|$ and vice-versa.

First consider the case where $d/2$ is even. Given a one-sided d-scattered set $S \subseteq A$, we let C include all the sets that correspond to some vertex in S and suppose for a contradiction that there exists a pair of sets $c_1, c_2 \in C$ that are incompatible, i.e. that there exists some element e with $e \in c_1$ and $e \in c_2$. Let $a_1, a_2 \in A$ be the vertices corresponding to sets c_1, c_2 and $b \in B$ be the vertex corresponding to element e. Then it must be $dst_G(a_1, b) \leq d/2 - 1$ since $e \in c_1$ and $dst_G(b, a_2) \leq d/2 - 1$ since $e \in c_2$, that gives $dst_G(v_1, v_2) \leq d - 2$, which contradicts S being a d-scattered set. On the other hand, given collection C of compatible sets we let $S \subseteq A$ include all the vertices corresponding to some set in C and suppose there exists a pair of vertices $a_1, a_2 \in S$ for which it is $dst_G(a_1, a_2) < d$. Since d is even and $a_1, a_2 \in A$, if $dst_G(a_1, a_2) < d$ it must be $dst_G(a_1, a_2) \leq d - 2$, as any shortest path between two vertices on the same side of a bipartite graph must be of even length. Thus there must exist at least one vertex $b \in B$ on a shortest path between a_1, a_2 in G for which it is $dst_G(a_1, b) \leq d/2 - 1$ and $dst_G(b, a_2) \leq d/2 - 1$. This means that the element e corresponding to vertex $b \in B$ must be included in both sets c_1, c_2 corresponding to vertices $a_1, a_2 \in A$, which contradicts the compatibility of sets in C.

We next consider the case where $d/2$ is odd. Given a one-sided d-scattered set $S \subseteq A$, we again let C include all sets that correspond to some vertex in S. If there exists a pair of sets $c_1, c_s \in C$ that contain the same element e corresponding to some vertex $b \in B$ or some element r that corresponds to a vertex $a \in A$, then by the same argument as in the even case we know that there must exist paths of length $\leq d/2 - 1$ from both vertices $a_1, a_2 \in A$ (corresponding to $c_1, c_2 \in C$) to vertex $b \in B$ or $a \in A$ and thus it must be $dst_G(a_1, a_2) < d$. On the other hand,

given a collection C of compatible sets we again let $S \subseteq A$ include all the vertices corresponding to sets in C. Supposing there exists a pair $a_1, a_2 \in S$ for which it is $dst_G(a_1, a_2) < d$, then again as d is even it must be $dst_G(a_1, a_2) \leq d-2$. This means there must be a vertex $a \in A$ on a shortest path between a_1 and a_2 for which $dst_G(a_1, a) \leq d/2 - 1$ and $dst_G(a, a_2) \leq d/2 - 1$, which means the corresponding sets $c_1, c_2 \in C$ must both contain element r that corresponds to this vertex $a \in A$, giving a contradiction.

Our algorithm then is as follows. For a given bipartite graph $G = (A \cup B, E)$, we define an instance of Set Packing as described above (depending on the parity of $d/2$) and apply the $\sqrt{n}$-approximation of [17]. Observe that $|A|, |B| \leq n$. We then exchange the sets A, B in the definitions of our instances and repeat the same process. This will return a solution S of size $|S| \geq \frac{1OPT_d(G)}{\sqrt{n}}$, which by Lemma 10 is $\geq \frac{OPT_d(G)}{2\sqrt{n}}$. □

4 Super-polynomial time

This section concerns itself with running-times that are not restricted to being functions polynomial in the size of the input. We begin with an upper bound on the size of the solution in any connected graph that is then employed in obtaining a simple exact exponential time algorithm.

Lemma 11 *The maximum size of any d-Scattered Set in a connected graph is* $\left\lfloor \frac{n}{\lfloor d/2 \rfloor} \right\rfloor$.

Proof: Given connected graph $G = (V, E)$, let $S \subseteq V$ be a d-scattered set in G. To each $u \in S$, we will assign all vertices at distance $< \lfloor d/2 \rfloor$: let $M(u) \coloneqq \{u\} \cup \{v \in V | dst(u, v) < \lfloor d/2 \rfloor\}$. Our aim is to show that for any $u \in S$, it must be $|M(u)| \geq \lfloor d/2 \rfloor$. In other words, for any vertex u in the solution there must be at least $\lfloor d/2 \rfloor - 1$ distinct vertices that are at distance $< \lfloor d/2 \rfloor$ from u and at distance $\geq \lfloor d/2 \rfloor$ from any vertex $w \in S$. Observe that if for some pair $u, w \in S$, we have $M(u) \cap M(w) \neq \emptyset$, then $dst(u, w) < d$, as there exists a vertex at distance $< \lfloor d/2 \rfloor$ from both u, w.

Consider some $u \in S$ and let k be the number of vertex-disjoint paths of length $< \lfloor d/2 \rfloor$ starting from u, i.e. such that no vertices are shared between them. Then it must be $|M(u) \setminus \{u\}| \geq k(\lfloor d/2 \rfloor - 1)$. If $V \subseteq M(u)$, then $OPT_d(G) = |S| = 1$ and the claim trivially holds, so we may assume that there is at least one vertex $z \notin M(u)$. This means $k \geq 1$, since G is connected and there must be (at least) one path from z to u of length $\geq \lfloor d/2 \rfloor$. It is then $|M(u) \setminus \{u\}| + 1 \geq \lfloor d/2 \rfloor - 1 + 1$ giving $|M(u)| \geq \lfloor d/2 \rfloor$.

This means that for each vertex u taken in any solution S there must be at least $\lfloor d/2 \rfloor$ distinct vertices in the graph, i.e. G must contain $|S|$ disjoint subsets of size at least $\lfloor d/2 \rfloor$ and the claim follows. □

Theorem 7 *The d-Scattered Set problem can be solved in* $O^*((ed)^{\frac{2n}{d}})$ *time.*

Proof: We simply try all sets of vertices of size at most $\left\lfloor \frac{n}{\lfloor d/2 \rfloor} \right\rfloor$ for feasibility and retain the best one found. By Lemma 11, the optimal solution must be contained therein. The number of sets we examine is $\leq \frac{2n}{d} \binom{n}{d/2} = O^*((ed)^{\frac{2n}{d}})$, giving also the running-time. □

4.1 Inapproximability

We now turn our attention to the problem's hardness of approximation in super-polynomial time. We use Theorem 1 in conjunction with straightforward reductions from Independent Set to d-Scattered Set for the two cases, that depend on the parity of d.

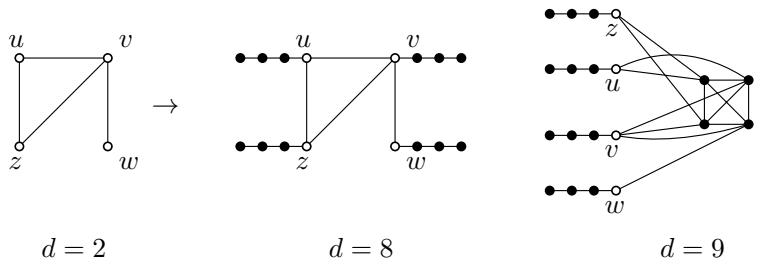

Figure 4: Examples of the constructions for even (center) and odd (right) values of d. Note the existence of an edge "gadget" for the odd case.

Theorem 8 *Under the randomized ETH, for any even $d \geq 4$, $\epsilon > 0$ and $\rho \leq (2n/d)^{5/6}$, no ρ-approximation for d-*Scattered Set *can take time $2^{\left(\frac{n^{1-\epsilon}}{\rho^{1+\epsilon}d^{1-\epsilon}}\right)} \cdot n^{O(1)}$.*

Proof: We suppose the existence of a ρ-approximation algorithm for d-Scattered Set of running-time $2^{\left(\frac{n^{1-\epsilon}}{d^{1-\epsilon}\rho^{1+\epsilon}}\right)} \cdot n^{O(1)}$ for some $\epsilon > 0$ and aim to show this would violate the (randomized) ETH. First let $\epsilon_1 > 0$ be such that $\epsilon > \epsilon_1$ and $\epsilon > \frac{2\epsilon_1}{1-3\epsilon_1}$. We next define $\epsilon_2 = \frac{1}{1-\epsilon_1} - 1$ and $r = \rho^{\frac{1-\epsilon_1}{1-3\epsilon_1}}$. Then, given a formula ϕ of 3SAT on N variables, we use the reduction of Theorem 1 with parameters r and ϵ_2 to build a graph G from ϕ, with $|V(G)| = N^{1+\epsilon_2}r^{1+\epsilon_2}$ and maximum degree r, such that with high probability: if ϕ is satisfiable then $\alpha(G) \geq N^{1+\epsilon_2}r$; if ϕ is not satisfiable then $\alpha(G) \leq N^{1+\epsilon_2}r^{2\epsilon_2}$. Thus an approximation algorithm with ratio $r^{1-2\epsilon_2}$ would permit us to decide if ϕ is satisfiable. We have that $r^{1-2\epsilon_2} = (\rho^{\frac{1-\epsilon_1}{1-3\epsilon_1}})^{1-2\epsilon_2} = (\rho^{\frac{1-\epsilon_1}{1-3\epsilon_1}})^{3-\frac{2}{1-\epsilon_1}} = \rho^{\frac{3+\frac{2\epsilon_1-2}{1-\epsilon_1}-3\epsilon_1}{1-3\epsilon_1}} = \rho$.

We will construct graph H from G as follows (similarly to Theorem 3.10 in [17], see also Figure 4): graph H contains a copy of G and a distinct path of $d/2-1$ edges attached to each vertex of G. Without loss of generality, we may assume that any d-scattered set will prefer selecting an endpoint of these attached paths than some vertex from G, as these selections would exclude strictly fewer vertices from the solution, i.e. any solution can only be improved by exchanging any vertex of G with selecting the (other) endpoint of the path attached to it. A pair of vertices that are endpoints of such paths (and not originally in G) will be at distance $\geq 2(d/2-1)+2 = d$, only if the vertices of G to which they are attached are non-adjacent, i.e. if the shortest path between them is of length at least 2. Thus, d-scattered sets in H are in one-to-one correspondence with independent sets in G and $\alpha(G) = OPT_d(H)$. The size of H is $n = |V(H)| = |V(G)|(d/2-1+1) = N^{1+\epsilon_2}r^{1+\epsilon_2}(d/2)$. Note also that $\rho \leq N^5$ and thus $\rho \leq (2n/d)^{5/6}$.

If ϕ is satisfiable then $OPT_d(H) = \alpha(G) \geq N^{1+\epsilon_2}r$, while if ϕ is not satisfiable then $OPT_d(H) = \alpha(G) \leq N^{1+\epsilon_2}r^{2\epsilon_2}$. Therefore, applying the supposed ρ-approximation for d-Scattered Set on H would permit us to solve 3SAT in time $2^{\left(\frac{n^{1-\epsilon}}{d^{1-\epsilon}\rho^{1+\epsilon}}\right)} \cdot n^{O(1)}$, with high probability. We next show that this would violate the ETH, i.e. $2^{\left(\frac{n^{1-\epsilon}}{d^{1-\epsilon}\rho^{1+\epsilon}}\right)} \cdot n^{O(1)} = 2^{o(N)}$. We have:

$$n = N^{1+\epsilon_2}r^{1+\epsilon_2}(d/2) \Rightarrow N = \left(\frac{2n}{d}\right)^{1-\epsilon_1} \cdot \frac{1}{\rho^{\left(\frac{1-\epsilon_1}{1-3\epsilon_1}\right)}} \tag{1}$$

And we then need to show:

$$2^{\left(\frac{n^{1-\epsilon}}{d^{1-\epsilon}\rho^{1+\epsilon}}\right)} \cdot n^{O(1)} = 2^{o(N)} = 2^{o\left(\left(\frac{2n}{d}\right)^{1-\epsilon_1} \cdot \frac{1}{\rho^{\left(\frac{1-\epsilon_1}{1-3\epsilon_1}\right)}}\right)} \tag{2}$$

Observe that, since $\epsilon > \epsilon_1$ and $\epsilon > \frac{2\epsilon_1}{1-3\epsilon_1}$, it is:

$$(n/d)^{1-\epsilon} < (2n/d)^{1-\epsilon_1} \tag{3}$$

$$(1/\rho^{1+\epsilon}) < \left(1/\rho^{\left(\frac{1-\epsilon_1}{1-3\epsilon_1}\right)}\right) \tag{4}$$

Which then gives:

$$\lim_{(n,\rho)\to\infty} \frac{2^{\left(\left(\frac{n}{d}\right)^{1-\epsilon} \cdot \frac{1}{\rho^{1+\epsilon}}\right)}}{2^{\left(\left(\frac{2n}{d}\right)^{1-\epsilon_1} \cdot \frac{1}{\rho^{\left(\frac{1-\epsilon_1}{1-3\epsilon_1}\right)}}\right)}} = 0 \tag{5}$$

□

The following reduction from Independent Set to d-Scattered Set for odd values of d uses a construction that includes a copy of every edge of the original graph (an *edge gadget*, see Figure 4). This necessity is responsible for the difference in running-times and is due to the parity idiosyncrasies of the problem as mentioned above.

Theorem 9 *Under the randomized ETH, for any odd $d \geq 5$, $\epsilon > 0$ and $\rho \leq (2n/d)^{5/6}$, no ρ-approximation for d-Scattered Set can take time $2^{\left(\frac{n^{1-\epsilon}}{\rho^{1+\epsilon}(d+\rho)^{1+\epsilon}}\right)} \cdot n^{O(1)}$.*

Proof: We suppose the existence of a ρ-approximation algorithm for d-Scattered Set of running-time $2^{\left(\frac{n^{1-\epsilon}}{\rho^{1+\epsilon}(d+\rho)^{1+\epsilon}}\right)} \cdot n^{O(1)}$ for some $\epsilon > 0$ and aim to show this would violate the (randomized) ETH. We let $\epsilon_1 > 0$ be such that $\epsilon > \frac{\epsilon_1}{1+\epsilon_1}$ and also $\epsilon > \frac{2\epsilon_1}{1-2\epsilon_1}$, as well as $\epsilon > \frac{2\epsilon_1^2+\epsilon_1}{1-2\epsilon_1^2-\epsilon_1}$. We then set $r = \rho^{\left(\frac{1}{1-2\epsilon_1}\right)}$. Given a formula ϕ of 3SAT on N variables, we build graph G from ϕ using the reduction of Theorem 1 with parameters r and ϵ_1. The size of G is $|V(G)| = N^{1+\epsilon_1} r^{1+\epsilon_1}$, its maximum degree is r and with high probability: if ϕ is satisfiable then $\alpha(G) \geq N^{1+\epsilon_1} r$; if ϕ is not satisfiable then $\alpha(G) \leq N^{1+\epsilon_1} r^{2\epsilon_1}$. Therefore an approximation algorithm with ratio $r^{1-2\epsilon_1} = \rho$ would permit us to decide if ϕ is satisfiable.

For odd $d \geq 5$, we will construct graph H from G as follows (again, a similar reduction is alluded to in the proof of Theorem 3.10 from [17] and partly also used for Corollary 1 from [12]): we make a vertex in H for each vertex of G and we also attach a distinct path of $(d-3)/2$ edges to each of them. We then make a vertex for every edge of G, turn all these vertices into a clique and also connect each one to the two vertices of H representing its endpoints. In this way, all pairs of vertices in H are at distance $\leq 2(d-1)/2+1 = d$ and the only vertices at exactly this distance are pairs of leaves on paths added to vertices that do not share a common neighbor representing some edge of G. Thus, d-scattered sets in H are again in one-to-one correspondence with independent sets in G and $\alpha(G) = OPT_d(H)$. The size of H is $n = |V(H)| = |V(G)|(d-1)/2 + |E(G)|$. The construction of [8] builds a graph where every vertex has degree at least one and at most

r, therefore $|E(G)| \geq |V(G)|$ and $|E(G)| \leq |V(G)|r/2$, that gives $n \leq N^{1+\epsilon_1} r^{1+\epsilon_1} (\frac{d+r-1}{2})$, while $\rho \leq N^5$, with $n \geq N\rho(d+1)/2$, that gives $\rho \leq (2n/d)^{5/6}$.

If ϕ is satisfiable then $OPT_d(H) = \alpha(G) \geq N^{1+\epsilon_1} r$, while if ϕ is not satisfiable then $OPT_d(H) = \alpha(G) \leq N^{1+\epsilon_1} r^{2\epsilon_1}$. Thus the supposed ρ-approximation for d-Scattered Set on H would permit us to solve 3SAT in time $2^{\left(\frac{n^{1-\epsilon}}{\rho^{1+\epsilon}(d+\rho)^{1+\epsilon}}\right)} \cdot n^{O(1)}$, with high probability. We next show that this would violate the ETH, i.e. $2^{\left(\frac{n^{1-\epsilon}}{\rho^{1+\epsilon}(d+\rho)^{1+\epsilon}}\right)} \cdot n^{O(1)} = 2^{o(N)}$. It is:

$$n \leq N^{1+\epsilon_1} r^{1+\epsilon_1} \left(\frac{d+r-1}{2}\right) \Rightarrow \tag{6}$$

$$\Rightarrow 2^N \geq 2^{\left(\left(\frac{2n}{d+r-1}\right)^{\frac{1}{1+\epsilon_1}} \cdot \frac{1}{r}\right)} = 2^{\left(\left(\frac{2n}{d+\rho\left(\frac{1}{1-2\epsilon_1}\right)-1}\right)^{\frac{1}{1+\epsilon_1}} \cdot \frac{1}{\rho\left(\frac{1}{1-2\epsilon_1}\right)}\right)} \tag{7}$$

Observe it is $(d+\rho)^{\frac{1}{1-2\epsilon_1}} > (d+\rho^{\frac{1}{1-2\epsilon_1}} - 1)$ and so $2^N > 2^{\left(\left(\frac{2n}{(d+\rho)^{\frac{1}{1-2\epsilon_1}}}\right)^{\frac{1}{1+\epsilon_1}} \cdot \frac{1}{\rho\left(\frac{1}{1-2\epsilon_1}\right)}\right)}$. We thus then require:

$$\lim_{(n,\rho)\to\infty} \frac{2^{\left(\frac{n^{1-\epsilon}}{\rho^{1+\epsilon}(d+\rho)^{1+\epsilon}}\right)}}{2^{\left(\left(\frac{2n}{(d+\rho)^{\frac{1}{1-2\epsilon_1}}}\right)^{\frac{1}{1+\epsilon_1}} \cdot \frac{1}{\rho\left(\frac{1}{1-2\epsilon_1}\right)}\right)}} = 0 \tag{8}$$

This is shown by the following inequalities:

$$\epsilon > \frac{\epsilon_1}{1+\epsilon_1} \Rightarrow n^{1-\epsilon} < 2n^{\frac{1}{1+\epsilon_1}} \tag{9}$$

$$\epsilon > \frac{2\epsilon_1^2+\epsilon_1}{1-2\epsilon_1^2-\epsilon_1} \Rightarrow \frac{1}{(d+\rho)^{1+\epsilon}} < \frac{1}{(d+\rho)^{\frac{1}{(1+\epsilon_1)(1-2\epsilon_1)}}} \tag{10}$$

$$\epsilon > \frac{2\epsilon_1}{1-2\epsilon_1} \Rightarrow \frac{1}{\rho^{1+\epsilon}} < \frac{1}{\rho^{(\frac{1}{1-2\epsilon_1})}} \tag{11}$$

□

4.2 Approximation

We complement the above hardness results with approximation algorithms of almost matching super-polynomial running-times. Similarly to the exact algorithm of Theorem 7, the upper bound from the beginning of this section is used for even values of d, while for the odd values this idea is combined with a greedy scheme based on minimum vertex degree.

Theorem 10 *For any even $d \geq 2$ and any $\rho \leq \frac{n}{\lfloor d/2 \rfloor}$, there is a ρ-approximation algorithm for d-Scattered Set of running-time $O^*((e\rho d)^{\frac{2n}{\rho d}})$.*

Proof: From Lemma 11 we know that the maximum size of a d-scattered set is $\lfloor \frac{n}{\lfloor d/2 \rfloor} \rfloor$. We thus simply try all sets of vertices of size at most $\frac{n}{\rho \lfloor d/2 \rfloor}$ for feasibility and retain the best one: these

are $\leq \frac{n}{\rho\lfloor d/2\rfloor}\binom{n}{\rho\lfloor d/2\rfloor} = O^*((e\rho d)^{\frac{2n}{\rho d}})$, that gives the running-time. If the graph is not connected, we can apply Lemma 11 to each connected component C of size n_C and then consider all subsets of size at most $\frac{n_C}{\rho\lfloor d/2\rfloor}$ in each C. □

Theorem 11 *For any odd $d \geq 3$ and any $\rho \leq \frac{n}{\lfloor d/2\rfloor}$, there is a ρ-approximation algorithm for d-*Scattered Set *of running-time $O^*((e\rho d)^{\frac{2n}{\rho(d+\rho)}})$.*

Proof: Let $q = (d-1)/2$ and G' be the q-th power of graph G. We then claim that any d-scattered set S in G is a 3-scattered set in G' and vice-versa: if S is a d-scattered set in G, then for any pair $u, v \in S$, it is $dst_G(u, v) \geq d$. For some pair of distinct vertices w, z on a shortest path between u, v in G, it must be $dst_G(u, w) = dst_G(z, v) = q$, while also $dst_G(u, z) = dst_G(w, v) > q$, i.e. w and z are two vertices on a shortest path from u to v, each at equal distance q from their closest endpoint (v or u). Then $dst_{G'}(u, w) = dst_{G'}(z, v) = 1$. Since w, z are distinct, it must be $dst_{G'}(w, z) \geq 1$ which gives also $dst_{G'}(u, v) \geq 3$, since $dst_{G'}(u, z) = dst_{G'}(w, v) > 1$.

If S is a 3-scattered set in G', then for any pair $u, v \in S$ it is $dst_{G'}(u, v) \geq 3$. Now, any shortest path in G between u, v must contain two distinct vertices w, z for which $dst_G(u, w) = q$ and $dst_G(z, v) = q$, while $dst_G(u, z) = dst_G(w, v) > q$. If no such pair of vertices exists in G, then $dst_{G'}(u, v) < 3$: any pair of vertices at distance $\leq q$ in G are adjacent in G' and so for the distance between u, v in G' to be at least 3, there must be two vertices each at distance $\geq q$ from u, v in G. From this we get that $dst_G(u, v) \geq 2q + 1 = d$, since $dst_G(w, z) \geq 1$, as $w \neq z$.

The algorithm then proceeds in two phases. For the first phase, so long as there exists an unmarked vertex v_i in G' of minimum degree $< \rho$, we mark v_i as 'selected' and add it to $S_1 \subseteq V$, marking all vertices at distance ≤ 2 from v_i in G' as 'excluded' and adding them to $X_i \subseteq V$. That is, $X_i = N^2_{G'}(v_i)$ and we let $X = N^2_{G'}(S_1)$, i.e. $X = X_1 \cup \cdots \cup X_{|S_1|}$. We also let the remaining (unmarked) vertices belong to $H \subseteq V$. Thus when this procedure terminates we have V partitioned into three sets S_1, X, H, while the degree of any vertex in H is $\geq \rho$. For the second phase, we try all subsets of vertices of H of size at most $\frac{2n}{\rho(\rho+\lfloor d/2\rfloor)}$ for feasibility and retain the best one. These are $\leq \frac{2n}{\rho(\rho+\lfloor d/2\rfloor)}\binom{n}{\rho/2(\rho+\lfloor d/2\rfloor)} = O^*((e\rho d)^{2n/\rho(d+\rho)})$, giving the upper bound on the running-time.

Now let S_1^* be a 3-scattered set of maximum size in the subgraph of G' induced by $S_1 \cup X$, i.e. $|S_1^*| = OPT_3(G'[S_1 \cup X])$ and S_2^* be a 3-scattered set of maximum size in the subgraph of G' induced by H, i.e. $|S_2^*| = OPT_3(G'[H])$. As the degree of any vertex $v_i \in S_1$ is $< \rho$, we have (1): $|S_1^*| < \rho|S_1|$, since for every vertex u in $N^1_{G'}(v_i)$, for $v_i \in S_1$, 3-scattered set S_1^* can contain at most one vertex w from $N^1_{G'}(u)$, as the distance between w and another neighbor of u is ≤ 2. For the second phase, it must be $|S_2^*| \leq n/\rho \Rightarrow 1/|S_2^*| \geq \rho/n$, since all vertices of H are of degree $\geq \rho$ and these neighborhoods are disjoint: if two vertices of S_2^* share a common neighbor then they cannot belong in a 3-scattered set. From Lemma 11, we also know that $|S_2^*| \leq n/\lfloor d/2\rfloor \Rightarrow 1/|S_2^*| \geq \lfloor d/2\rfloor/n$. Adding the two inequalities gives $2/|S_2^*| \geq \frac{\rho+\lfloor d/2\rfloor}{n} \Rightarrow |S_2^*| \leq \frac{2n}{\rho+\lfloor d/2\rfloor}$. Furthermore, it is $|S_2| \leq \frac{2n}{\rho(\rho+\lfloor d/2\rfloor)}$, by construction. Dividing the two inequalities gives (2): $\frac{|S_2^*|}{|S_2|} \leq \rho \Rightarrow |S_2^*| \leq \rho|S_2|$. From (1) and (2) we get that $|S_1^*| + |S_2^*| \leq \rho(|S_1| + |S_2|)$. It is $OPT_d(G) = OPT_3(G') \leq |S_1^*| + |S_2^*|$ since $S_1 \cup X$ and H form a partition of G'. Our algorithm returns a solution of size $|S_1| + |S_2|$ and thus our approximation ratio is $\frac{OPT_d(G)}{|S_1|+|S_2|} \leq \frac{|S_1^*|+|S_2^*|}{|S_1|+|S_2|} \leq \rho$. If the graph is not connected, we can apply Lemma 11 to each connected component C of size n_C and then try all subsets of size at most $\frac{n_C}{\rho\lfloor d/2\rfloor}$ in each C and obtain an additive version of (2) for each component. □

Treewidth of power graphs

We close this paper with a note on the treewidth of power graphs. Similar ideas as those used in the above results also point to the following upper bound on the increase in treewidth taking place when computing the power of a graph of bounded degree:

Theorem 12 *For any graph G of treewidth tw and maximum degree bounded by Δ, the treewidth tw' of the d-th power G^d is at most $tw' \leq tw \cdot \Delta \sum_{i=0}^{d/2-1}(\Delta-1)^i = O(tw \cdot \Delta^{d/2})$.*

Proof: Given a tree decomposition T of $G = (V, E)$ of width tw, we make a tree decomposition T' of $G^d = (V, E^d)$ by replacing the appearance of each vertex v in each bag of T with v and the set of vertices at distance at most $d/2$ from v in G, i.e. with $N_G^{d/2}(v) \cup \{v\}$. It is $|N_G^{d/2}(v)| \leq \sum_{i=0}^{d/2-1}(\Delta(\Delta-1)^i)$, from which we get the upper bound. This is a valid tree decomposition for G^d as: (a) all vertices appear in some bag of T' as they appeared in T, (b) for every edge $\{u, v\}$ in G^d, either $\{u, v\} \in E$ and there is a bag in T containing both u, v and thus there is one also in T', or $\{u, v\}$ was added to E^d due to the distance between u, v being $\leq d$ in G. In this case there must be at least one vertex w at distance $\leq d/2$ from both u and v in G, meaning there will be a bag in T' containing all three vertices u, v, w that was constructed from a bag of T that contains w.

Finally, (c) for every vertex v appearing in two bags X', Y' of T', vertex v also appears on every bag on the path from X' to Y' in T': consider (for a contradiction) the existence of a bag Z' on the path from X' to Y' in T' that does not contain v, and let X, Y, Z be the corresponding bags in T. Since X', Y' contain v, then both X and Y contain some vertex u at distance at most $d/2$ from v in G, or v itself, i.e. $u \in N_G^{d/2}(v) \cup \{v\}$. If both X and Y contain v, then as T is a valid tree decomposition, so does Z and therefore also Z'. Thus we may assume that at least one of X, Y do not contain v, as well as $v \notin Z$. As Z is a separator, then v must appear only on one side of Z in T. We assume (without loss of generality) that v only appears on the X-side of T (from Z) and v is not contained in Y. Thus Y must contain some vertex u at distance $\leq d/2$ from v in G. As Z is a separator, the path from v to u must contain at least one vertex $w \in Z$, at distance $< d/2$ from v. Thus Z' must also contain v, as it includes all vertices at distance $\leq d/2$ from w. □

As for graphs of maximum degree bounded by Δ, we have $\text{cw} \leq O(\Delta \cdot \text{tw})$ (see [9]) and $\text{tw} \leq O(\Delta \cdot \text{cw})$ (directly derived from the well-known result of Gurski and Wanke [16]), we also obtain the following corollary.

Corollary 13 *For any graph G of clique-width cw and maximum degree bounded by Δ, the clique-width cw' of the d-th power G^d is at most $cw' \leq O(cw \cdot \Delta^{d/2+2})$.*

5 Conclusion

In this paper we furthered our understanding of the d-Scattered Set problem by answering remaining questions on its (super-)polynomial (in-)approximability. In particular, we showed the following:

- A lower bound of $\Delta^{\lfloor d/2 \rfloor - \epsilon}$ on the approximation ratio of any polynomial-time algorithm for graphs of maximum degree Δ and an improved upper bound of $O(\Delta^{\lfloor d/2 \rfloor})$ on the approximation ratio of any greedy scheme for this problem.

- A polynomial-time approximation algorithm of ratio $2\sqrt{n}$ for bipartite graphs and even values of d, that complements known results by considering the only remaining open case.
- An exact exponential time algorithm of complexity $O^*((ed)^{\frac{2n}{d}})$, based on an upper bound on the size of any solution.
- A lower bound on the complexity of any ρ-approximation algorithm of (roughly) $2^{\frac{n^{1-\epsilon}}{\rho d}}$ for even d and $2^{\frac{n^{1-\epsilon}}{\rho(d+\rho)}}$ for odd d, under the randomized ETH.
- ρ-approximation algorithms of running-times $O^*((e\rho d)^{\frac{2n}{\rho d}})$ for even d and $O^*((e\rho d)^{\frac{2n}{\rho(d+\rho)}})$ for odd d that (almost) match the above lower bounds, thus giving a clear picture of the trade-off curve between approximation and running-time.

Apart from the possibility of "de-randomization" of the results above that use the randomized construction of [8] as a starting point, some remaining unanswered questions here would concern the complexity of the problem on chordal bipartite graphs (also mentioned as an open problem by [12]) as well as the functionality of the PTAS for planar graphs by the same authors, that only works for fixed values of d as it extends the well-known approach of [2] for obtaining such algorithms in planar graphs for e.g. Independent Set. Because this approach involves breaking down the graph into (roughly) d-outerplanar subgraphs and then exactly solving the problem in each of these using dynamic programming over their tree decompositions, for values of d that are not constant (say $d \geq \sqrt{n}$)) this is not achievable in polynomial time, due to the exponent of the treewidth algorithms depending on d. It would be interesting to see an extension of this (or some other) approach for the case of unbounded d, or, conversely, a hardness reduction proving it is unlikely. The hard part would have to involve a construction that is very efficient in terms of *crossing* gadgets in order to maintain planarity, or, from the other side, a way to optimally solve the problem in carefully constructed subgraphs without spending time exponential on d.

References

[1] N. Alon, U. Feige, A. Wigderson, and D. Zuckerman. Derandomized graph products. *Comput. Complex.*, 5(1):60–75, 1995. doi:10.1007/BF01277956.

[2] B. S. Baker. Approximation algorithms for np-complete problems on planar graphs. *J. ACM*, 41(1):153–180, 1994. doi:10.1145/174644.174650.

[3] P. Berman and M. Karpinski. On some tighter inapproximability results (extended abstract). In J. Wiedermann, P. van Emde Boas, and M. Nielsen, editors, *Automata, Languages and Programming, 26th International Colloquium, ICALP'99, Prague, Czech Republic, July 11-15, 1999, Proceedings*, volume 1644 of *Lecture Notes in Computer Science*, pages 200–209. Springer, 1999. doi:10.1007/3-540-48523-6_17.

[4] B. Bollobás and F. R. K. Chung. The diameter of a cycle plus a random matching. *SIAM J. Discret. Math.*, 1(3):328–333, 1988. doi:10.1137/0401033.

[5] É. Bonnet, M. Lampis, and V. T. Paschos. Time-approximation trade-offs for inapproximable problems. *J. Comput. Syst. Sci.*, 92:171–180, 2018. doi:10.1016/j.jcss.2017.09.009.

[6] N. Bourgeois, B. Escoffier, and V. T. Paschos. Approximation of max independent set, min vertex cover and related problems by moderately exponential algorithms. *Discret. Appl. Math.*, 159(17):1954–1970, 2011. doi:10.1016/j.dam.2011.07.009.

[7] A. Brouwer and W. Haemers. *Spectra of graphs*. Universitext. Springer, 2012.

[8] P. Chalermsook, B. Laekhanukit, and D. Nanongkai. Independent set, induced matching, and pricing: Connections and tight (subexponential time) approximation hardnesses. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 370–379. IEEE Computer Society, 2013. doi:10.1109/FOCS.2013.47.

[9] B. Courcelle. On the model-checking of monadic second-order formulas with edge set quantifications. *Discret. Appl. Math.*, 160(6):866–887, 2012. doi:10.1016/j.dam.2010.12.017.

[10] M. Cygan, L. Kowalik, M. Pilipczuk, and M. Wykurz. Exponential-time approximation of hard problems. *CoRR*, abs/0810.4934, 2008.

[11] M. Demange and V. T. Paschos. Improved approximations for maximum independent set via approximation chains. *Applied Mathematics Letters*, 10(3):105 – 110, 1997.

[12] H. Eto, F. Guo, and E. Miyano. Distance-d independent set problems for bipartite and chordal graphs. *J. Comb. Optim.*, 27(1):88–99, 2014. doi:10.1007/s10878-012-9594-4.

[13] H. Eto, T. Ito, Z. Liu, and E. Miyano. Approximability of the distance independent set problem on regular graphs and planar graphs. In T. H. Chan, M. Li, and L. Wang, editors, *Combinatorial Optimization and Applications - 10th International Conference, COCOA 2016, Hong Kong, China, December 16-18, 2016, Proceedings*, volume 10043 of *Lecture Notes in Computer Science*, pages 270–284. Springer, 2016. doi:10.1007/978-3-319-48749-6_20.

[14] H. Eto, T. Ito, Z. Liu, and E. Miyano. Approximation algorithm for the distance-3 independent set problem on cubic graphs. In S. Poon, M. S. Rahman, and H. Yen, editors, *WALCOM: Algorithms and Computation, 11th International Conference and Workshops, WALCOM 2017, Hsinchu, Taiwan, March 29-31, 2017, Proceedings*, volume 10167 of *Lecture Notes in Computer Science*, pages 228–240. Springer, 2017. doi:10.1007/978-3-319-53925-6_18.

[15] F. V. Fomin, D. Lokshtanov, V. Raman, and S. Saurabh. Bidimensionality and EPTAS. In D. Randall, editor, *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23-25, 2011*, pages 748–759. SIAM, 2011. doi:10.1137/1.9781611973082.59.

[16] F. Gurski and E. Wanke. The tree-width of clique-width bounded graphs without $K_{n,\,n}$. In U. Brandes and D. Wagner, editors, *Graph-Theoretic Concepts in Computer Science, 26th International Workshop, WG 2000, Konstanz, Germany, June 15-17, 2000, Proceedings*, volume 1928 of *Lecture Notes in Computer Science*, pages 196–205. Springer, 2000. doi:10.1007/3-540-40064-8_19.

[17] M. M. Halldórsson, J. Kratochvíl, and J. A. Telle. Independent sets with domination constraints. *Discret. Appl. Math.*, 99(1-3):39–54, 2000. doi:10.1016/S0166-218X(99)00124-9.

[18] M. M. Halldórsson and J. Radhakrishnan. Greed is good: Approximating independent sets in sparse and bounded-degree graphs. *Algorithmica*, 18(1):145–163, 1997. doi:10.1007/BF02523693.

[19] J. Håstad. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Mathematica*, 182:105–142, 1999.

[20] R. Impagliazzo and R. Paturi. On the complexity of k-sat. *J. Comput. Syst. Sci.*, 62(2):367–375, 2001. doi:10.1006/jcss.2000.1727.

[21] R. Impagliazzo, R. Paturi, and F. Zane. Which problems have strongly exponential complexity? *J. Comput. Syst. Sci.*, 63(4):512–530, 2001. doi:10.1006/jcss.2001.1774.

[22] I. Katsikarelis, M. Lampis, and V. T. Paschos. Improved (in-)approximability bounds for d-scattered set. In E. Bampis and N. Megow, editors, *Approximation and Online Algorithms - 17th International Workshop, WAOA 2019, Munich, Germany, September 12-13, 2019, Revised Selected Papers*, volume 11926 of *Lecture Notes in Computer Science*, pages 202–216. Springer, 2019. doi:10.1007/978-3-030-39479-0_14.

[23] I. Katsikarelis, M. Lampis, and V. T. Paschos. Structurally parameterized d-scattered set. *Discret. Appl. Math.*, 308:168–186, 2022. doi:10.1016/j.dam.2020.03.052.

[24] D. Marx and M. Pilipczuk. Optimal parameterized algorithms for planar facility location problems using voronoi diagrams. *ACM Trans. Algorithms*, 18(2):13:1–13:64, 2022. doi:10.1145/3483425.

[25] P. Montealegre and I. Todinca. On distance-d independent set and other problems in graphs with "few" minimal separators. In P. Heggernes, editor, *Graph-Theoretic Concepts in Computer Science - 42nd International Workshop, WG 2016, Istanbul, Turkey, June 22-24, 2016, Revised Selected Papers*, volume 9941 of *Lecture Notes in Computer Science*, pages 183–194, 2016. doi:10.1007/978-3-662-53536-3_16.

[26] M. Pilipczuk and S. Siebertz. Kernelization and approximation of distance-r independent sets on nowhere dense graphs. *Eur. J. Comb.*, 94:103309, 2021. doi:10.1016/j.ejc.2021.103309.