

Approximation Algorithms for Maximum Independent Set of Pseudo-Disks*

Timothy M. Chan[†] Sariel Har-Peled[‡]

January 26, 2012

Abstract

We present approximation algorithms for maximum independent set of pseudo-disks in the plane, both in the weighted and unweighted cases. For the unweighted case, we prove that a local search algorithm yields a PTAS. For the weighted case, we suggest a novel rounding scheme based on an LP relaxation of the problem, which leads to a constant-factor approximation.

Most previous algorithms for maximum independent set (in geometric settings) relied on packing arguments that are not applicable in this case. As such, the analysis of both algorithms requires some new combinatorial ideas, which we believe to be of independent interest.

1 Introduction

Let $F = \{f_1, \dots, f_n\}$ be a set of n objects in the plane, with weights $w_1, w_2, \dots, w_n > 0$, respectively. In this paper, we are interested in the problem of finding an independent set of maximum weight. Here a set of objects is *independent*, if no pair of objects intersect.

A natural approach to this problem is to build an *intersection graph* $G = (V, E)$, where the objects form the vertices, and two objects are connected by an edge if they intersect, and weights are associated with the vertices. We want the maximum independent set in G . This is of course an **NP-COMplete** problem, and it is known that no approximation factor is possible within $|V|^{1-\varepsilon}$ for any $\varepsilon > 0$ if $\text{NP} \neq \text{ZPP}$ [Has96]. In fact, even if the maximum degree of the graph is bounded by 3, no PTAS is possible in this case [BF99].

In geometric settings, better results are possible. If the objects are fat (e.g., disks and squares), PTASs are known. One approach [Cha03, EJS05] relies on a hierarchical spatial

*A preliminary version of the paper appeared in SoCG 2009 [CH09].

[†]School of Computer Science; University of Waterloo; 200 University Ave West; Waterloo, Ontario N2L 3G1; Canada; tmchan@uwaterloo.ca; <http://www.cs.uwaterloo.ca/~tmchan/>.

[‡]Department of Computer Science; University of Illinois; 201 N. Goodwin Avenue; Urbana, IL, 61801, USA; sariel@uiuc.edu; <http://www.uiuc.edu/~sariel/>.

subdivision, such as a quadtree, combined with dynamic programming techniques [Aro98]; it works even in the weighted case. Another approach [Cha03] relies on a recursive application of a nontrivial generalization of the planar separator theorem [LT79, SW98]; this approach is limited to the unweighted case. If the objects are not fat, only weaker results are known. For the problem of finding a maximum independent set of unweighted axis-parallel rectangles, an $O(\log \log n)$ -approximation algorithm was recently given by Chalermsook and Chuzhoy [CC09]. For line segments, a roughly $O(\sqrt{\text{Opt}})$ -approximation is known [AM06]; recently, Fox and Pach [FP11] have improved the approximation factor to n^ϵ for not only for line segments but curve segments that intersect a constant number of times.

In this paper we are interested in the problem of finding a large independent set in a set of weighted or unweighted pseudo-disks. A set of objects is a collection of *pseudo-disks*, if the boundary of every pair of them intersects at most twice. This case is especially intriguing because previous techniques seem powerless: it is unclear how one can adapt the quadtree approach [Cha03, EJS05] or the generalized separator approach [Cha03] for pseudo-disks.

Even a constant-factor approximation in the unweighted case is not easy. Consider the most obvious greedy strategy for disks (or fat objects): select the object $f_i \in F$ of the smallest radius, remove all objects that intersect f_i from F , and repeat. This is already sufficient to yield a constant-factor approximation by a simple packing argument [EKNS00, MBH⁺95]. However, even this simplest algorithm breaks down for pseudo-disks—as pseudo-disks are defined “topologically”, how would one define the “smallest” pseudo-disk in a collection?

Independent set via local search. Nevertheless, we are able to prove that a different strategy can yield a constant-factor approximation for unweighted pseudo-disks: local search. In the general settings, local search was used to get (roughly) a $\Delta/4$ approximation to independent set, where Δ is the maximum degree in the graph, see [Hal98] for a survey. In the geometric settings, Agarwal and Mustafa [AM06, Lemma 4.2] had a proof that a local search algorithm gives a constant-factor approximation for the special case of pseudo-disks that are rectangles; their proof does not immediately work for arbitrary pseudo-disks. Our proof provides a generalization of their lemma.

In fact, we are able to do more: we show that local search can actually yield a PTAS for unweighted pseudo-disks! This gives us by accident a new PTAS for the special case of disks and squares. Though the local-search algorithm is slower than the quadtree-based PTAS in these special cases [Cha03], it has the advantage that it only requires the intersection graph as input, not its geometric realization; previously, an algorithm with this property was only known in further special cases, such as unit disks [NHK05]. Our result uses the planar separator theorem, but in contrast to the separator-based method in [Cha03], a standard version of the theorem suffices and is needed only in the analysis, not in the algorithm itself.

Planar graphs are special cases of disk intersection graphs, and so our result applies. Of course, PTASs for planar graphs have been around for quite some time [LT79, Bak94], but the fact that a simple local search algorithm already yields a PTAS for planar graphs is apparently not well known, if it was known at all.

We can further show that the same local search algorithm gives a PTAS for independent

set for fat objects in any fixed dimension, reproving known results in [Cha03, EJS05].

This strategy, unfortunately, works only in the unweighted case.

Independent set via LP. It is easy to extract a large independent set from a sparse unweighted graph. For example, greedily, we can order the vertices from lowest to highest degree, and pick them one by one into the independent set, if none of its neighbors was already picked into the independent set. Let d_G be the average degree in G . Then a constant fraction of the vertices have degree $O(d_G)$, and the selection of such a vertex can eliminate $O(d_G)$ candidates. Thus, this yields an independent set of size $\Omega(n/d_G)$. Alternatively, for better constants, we can order the vertices by a random permutation and do the same. Clearly, the probability of a vertex v to be included in the independent set is $1/(d(v) + 1)$. An easy calculation leads to Turán’s theorem, which states that any graph G has an independent set of size $\geq n/(d_G + 1)$ [AS00].

Now, our intersection graph G may not be sparse. We would like to “sparsify” it, so that the new intersection graph is sparse and the number of vertices is close to the size of the optimal solution. Interestingly, we show that this can be done by solving the LP relaxation of the independent set problem. The relaxation provides us with a fractional solution, where every object f_i has value $x_i \in [0, 1]$ associated with it. Rounding this fractional solution into a feasible solution is not a trivial task, as no such scheme exists in the general case. Our basic approach is somewhat similar to the local-ratio technique [BYBFR04], but, more precisely, it is a variant of the contention resolution scheme of Chekuri *et al.* [CVZ11]. To this end, we prove a technical lemma (see Lemma 4.1) that shows that the total sum of terms of the form $x_i x_j$, over pairs $f_i f_j$ that intersect is bounded by the boundary complexity of the union of \mathcal{E} objects of F , where \mathcal{E} is the size of the fractional solution. The proof contains a nice application of the standard Clarkson technique [CS89].

This lemma implies that on average, if we pick f_i into our random set of objects, with probability x_i , then the resulting intersection graph would be sparse. This is by itself sufficient to get a constant-factor approximation for the unweighted case. For the weighted case, we follow a greedy approach: we examine the objects in a certain order (based on a quantity we call “resistance”), and choose an object with probability around x_i , *on condition that* it does not intersect any previously chosen object. We argue, for our particular order, that each object is indeed chosen with probability $\Omega(x_i)$. This leads to a constant-factor approximation for weighted pseudo-disks.

Interestingly, the rounding scheme we use can be used in more general settings, when one tries to find an independent set that maximizes a submodular target function. See Section 4.6 for details.

Linear union complexity. Our LP analysis works more generally for any class of objects with linear union complexity. We assume that the boundary of the union of any k of these objects has at most ϱk vertices, for some fixed ϱ . For pseudo-disks, the boundary of the union is made out of at most $6n - 12$ arcs, implying $\varrho = 6$ in this case [KLPS86].

A family F of simply connected regions bounded by simple closed curves in general

position in the plane is *k-admissible* (with k even) if for any pair $f_i, f_j \in F$, we have: (i) $f_i \setminus f_j$ and $f_j \setminus f_i$ are connected, and (ii) their boundary intersect at most k times. Whitesides and Zhao [WZ90] showed that the union of such n objects has at most $3kn - 6$ arcs; that is, $\varrho = 3k$. So, our LP analysis applies to this class of objects as well. For more results on union complexity, see the sermon by Agarwal *et al.* [APS08].

Our local-search PTAS works more generally for unweighted admissible regions in the plane. For an arbitrary class of unweighted objects with linear union complexity in the plane, local search still yields a constant-factor approximation.

Rectangles. LP relaxation has been used before, notably, in Chalermsook and Chuzhoy’s recent breakthrough in the case of axis-parallel rectangles [CC09], but their analysis is quite complicated. Although rectangles do not have linear union complexity in general, we observe in Section 5 that a variant of our approach can yield a readily accessible proof of a sublogarithmic $O(\log n / \log \log n)$ approximation factor for rectangles, even in the weighted case, where previously only logarithmic approximation is known [AvKS98, BDMR01, Cha04] (Chalermsook and Chuzhoy’s result is better but currently is applicable only to unweighted rectangles).

Discussion. Local search and LP relaxation are of course staples in the design of approximation algorithms, but are not seen as often in computational geometry. Our main contribution lies in the fusion of these approaches with combinatorial geometric techniques.

In a sense, one can view our results as complementary to the known results on approximate geometric set cover by Brönnimann and Goodrich [BG95] and Clarkson and Varadarajan [CV07]. They consider the problem of finding the minimum number of objects in F to cover a given point set. Their results imply a constant-factor approximation for families of objects with linear union complexity, for instance. One version of their approaches is indeed based on LP relaxation [Lon01, ERS05]. The “dual” hitting set problem is to find the minimum number of points to pierce a given set of objects. Brönnimann and Goodrich’s result combined with a recent result of Pyrga and Ray [PR08] also implies a constant-factor approximation for pseudo-disks for this piercing problem. The piercing problem is actually the dual of the independent set problem (this time, we are referring to linear programming duality). We remark, however, that the rounding schemes for set cover and piercing are based on different combinatorial techniques, namely, ε -nets, which are not sufficient to deal with independent set (one obvious difference is that independent set is a maximization problem).

In Theorem 4.6, we point out a combinatorial consequence of our LP analysis: for any collection of unweighted pseudo-disks, the ratio of the size of the minimum piercing set to the size of maximum independent set is at most a constant. (It is easy to see that the ratio is always at least 1; for disks or fat objects, it is not difficult to obtain a constant upper bound by packing arguments.) This result is of independent interest; for example, getting tight bounds on the ratio for axis-parallel rectangles is a long-standing open problem.

In an interesting independent development, Mustafa and Ray [MR09] have recently applied the local search paradigm to obtain a PTAS for the geometric set cover problem for

(unweighted) pseudo-disks and admissible regions.

2 Preliminaries

In the following, we have a set F of n objects in the plane, such that the union complexity of any subset $X \subseteq F$ is bounded by $\varrho |X|$, where ϱ is a constant. Here, the *union complexity* of X is the number of arcs on the boundary of the union of the objects of X . Let $\mathcal{A}(F)$ denote the arrangement of F , and $\mathcal{V}(F)$ denote the set of vertices of $\mathcal{A}(F)$.

In the following, we assume that deciding if two objects intersect takes constant time.

3 Approximation by Local Search: Unweighted Case

3.1 The algorithm

In the unweighted case, we may assume that no object is fully contained in another.

We say that a subset L of F is *b -locally optimal* if L is an independent set and one cannot obtain a larger independent set from L by deleting at most b objects and inserting at most $b + 1$ objects of F .

Our algorithm for the unweighted case simply returns a b -locally optimal solution for a suitable constant b , by performing a local search. We start with $L \leftarrow \emptyset$. For every subset $X \subseteq F \setminus L$ of size at most $b + 1$, we verify that X by itself is independent, and furthermore, that the set $Y \subseteq L$ of objects intersecting the objects of X , is of size at most $|X| - 1$. If so, we do $L \leftarrow (L \setminus Y) \cup X$. Every such exchange increases the size of L by at least one, and as such it can happen at most n times. Naively, there are $\binom{n}{b+1}$ subsets X to consider, and for each such subset X it takes $O(nb)$ time to compute Y . Therefore, the running time is bounded by $O(n^{b+3})$. (The running time can be probably improved by being a bit more careful about the implementation.)

3.2 Analysis

We present two alternative ways to analyze this algorithm. The first approach uses only the fact that the union complexity is low. The second approach is more direct, and uses the property that the regions are admissible.

3.2.1 Analysis using union complexity

The following lemma by Afshani and Chan [AC06], which was originally intended for different purposes, will turn out to be useful here (the proof exploits linearity of planar graphs and the Clarkson technique [CS89]):

Lemma 3.1 *Suppose we have n disjoint simply connected regions in the plane and a collection of disjoint curves, where each curve intersects at most k regions. Call two curves*

equivalent if they intersect precisely the same subset of regions. Then the number of equivalent classes is at most c_0nk^2 for some constant c_0 .

Let \mathcal{S} be an optimal solution, and let \mathbf{L} be a b -locally optimal solution. We will upper-bound $|\mathcal{S}|$ in terms of $|\mathbf{L}|$.

Let $\mathcal{S}_{>b}$ denote the set of objects in \mathcal{S} that intersect at least $b + 1$ objects of \mathbf{L} . Let $\mathcal{S}_{\leq b}$ be the set of remaining objects in \mathcal{S} .

If $f_i \in \mathcal{S}$ intersects $f_j \in \mathbf{L}$, then the pair of objects contributes at least two vertices to the boundary of the union of $\mathcal{S} \cup \mathbf{L}$. Indeed, the objects of \mathcal{S} (resp. \mathbf{L}) are disjoint from each other since this is an independent set, and no object is contained inside another (by assumption). We remind the reader that for any subset $X \subseteq \mathbf{F}$, the union complexity of the regions of X is $\leq \varrho|X|$. As such, the union complexity of $\mathcal{S}_{>b} \cup \mathbf{L}$ is $\leq \varrho(|\mathcal{S}_{>b}| + |\mathbf{L}|)$. Thus,

$$\begin{aligned} 2(b+1)|\mathcal{S}_{>b}| &\leq \varrho(|\mathcal{S}_{>b}| + |\mathbf{L}|) \\ \implies |\mathcal{S}_{>b}| &\leq \frac{\varrho}{2(b+1) - \varrho} |\mathbf{L}|. \end{aligned}$$

On the other hand, by applying Lemma 3.1 with \mathbf{L} as the regions and the boundaries of $\mathcal{S}_{\leq b}$ as the curves, the objects in $\mathcal{S}_{\leq b}$ form at most $c_0b^2|\mathbf{L}|$ equivalent classes. Each equivalent class contains at most b objects: Otherwise we would be able to remove b objects from \mathbf{L} and intersect $b + 1$ objects in this equivalence class to get an independent set larger than \mathbf{L} . This would contradict the b -local optimality of \mathbf{L} . Thus, $|\mathcal{S}_{\leq b}| \leq c_0b^3|\mathbf{L}|$.

Combining the two inequalities, we get

$$|\mathcal{S}| \leq |\mathcal{S}_{\leq b}| + |\mathcal{S}_{>b}| \leq \left(c_0b^3 + \frac{\varrho}{2(b+1) - \varrho} \right) |\mathbf{L}|.$$

For example, we can set $b = \lceil \varrho/2 \rceil$ and the approximation factor is $O(\varrho^3)$.

Theorem 3.2 *Given a set of n unweighted objects in the plane with linear union complexity, for a sufficiently large constant b , any b -locally optimal independent set has size $\Omega(\text{opt})$, where opt is the size of the maximum independent set of the objects.*

3.2.2 Better analysis for admissible regions

A set of regions \mathbf{F} is **admissible**, if for any two regions $f, f' \in \mathbf{F}$, we have that $f \setminus f'$ and $f' \setminus f$ are both simply connected (i.e., connected and contains no holes). Note that we do not care how many times the boundaries of the two regions intersect, and furthermore, by definition, no region is contained inside another.

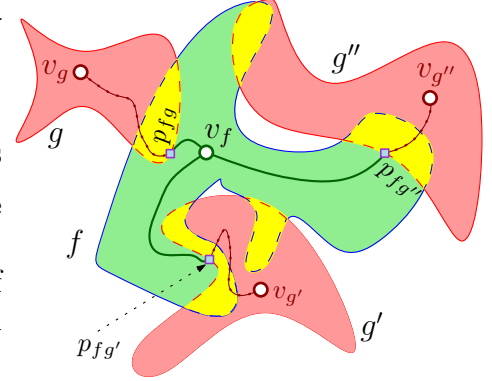
Lemma 3.3 *Let \mathbf{F} be a set of admissible regions, and consider a independent set of regions $I \subseteq \mathbf{F}$, and a region $f \in \mathbf{F} \setminus I$. Then, the **core** region $f \setminus I = f \setminus \bigcup_{g \in I} g$ is non-empty and simply connected.*

Proof: It is easy to verify that for the regions of I to split f into two connected components, they must intersect, which contradicts their disjointness. ■

Lemma 3.4 *Let $X, Y \subseteq \mathcal{F}$ be two independent sets of regions. Then the intersection graph G of $X \cup Y$ is planar.*

Proof: Lemma 3.3 implies the planarity of this graph.

Indeed, for a region $f \in X$, the core $f' = f \setminus \bigcup_{g \in Y} g$ is non-empty and simply connected. Place a vertex v_f inside this region, and for every object $g \in Y$ that intersects f , create a curve from v_f to a point $p_{f,g}$ on the boundary of g that lies inside f . Clearly, we can create these curves in such a way that they do not intersect each other.



Similarly, for every region $g \in Y$, we place a vertex v_g inside g , and connect it to all the points $p_{f,g}$ placed on its boundary, by curves that are contained in g , and they are interior disjoint. Clearly, together, these vertices and curves form a planar drawing of G . ■

We need the following version of the planar separator theorem. Below, for a set of vertices U in a graph G , let $\Gamma(U)$ denote the set of neighbors of U , and let $\bar{\Gamma}(U) = \Gamma(U) \cup U$.

Lemma 3.5 ([Fre87]) *There are constants c_1, c_2 and c_3 , such that for any planar graph $G = (V, E)$ with n vertices, and a parameter r , one can find a set of $X \subseteq V$ of size at most $c_1 n / \sqrt{r}$, and a partition of $V \setminus X$ into n/r sets $V_1, \dots, V_{n/r}$, satisfying: (i) $|V_i| \leq c_2 r$, (ii) $\Gamma(V_i) \cap V_j = \emptyset$, for $i \neq j$, and (iii) $|\Gamma(V_i) \cap X| \leq c_3 \sqrt{r}$.*

Let \mathcal{S} be the optimal solution and \mathbf{L} be a b -locally optimal solution. Consider the bipartite intersection graph G of $\mathcal{S} \cup \mathbf{L}$. By Lemma 3.4, we can apply Lemma 3.5 to G , for $r = b / (c_2 + c_3)$. Note that $|\bar{\Gamma}(V_i)| \leq c_2 r + c_3 \sqrt{r} < b$ for each i . Let

$$\begin{aligned} s_i &= |V_i \cap \mathcal{S}|, \\ \ell_i &= |V_i \cap \mathbf{L}|, \\ \text{and } b_i &= |\Gamma(V_i) \cap X|, \text{ for each } i. \end{aligned}$$

Observe that $\ell_i + b_i \geq s_i$, for all i . Indeed, otherwise, we can throw away the vertices of $\mathbf{L} \cap \bar{\Gamma}(V_i)$ from \mathbf{L} , and replace them by $V_i \cap \mathcal{S}$, resulting in a better solution. This would contradict the local optimality of \mathbf{L} . Thus,

$$\begin{aligned} |\mathcal{S}| &\leq \sum_i s_i + |X| \leq \sum_i \ell_i + \sum_i b_i + |X| \\ &\leq |\mathbf{L}| + c_3 \sqrt{r} \cdot \frac{|\mathcal{S}| + |\mathbf{L}|}{r} + c_1 \frac{|\mathcal{S}| + |\mathbf{L}|}{\sqrt{r}} \\ &\leq |\mathbf{L}| + (c_1 + c_3) \frac{|\mathcal{S}| + |\mathbf{L}|}{\sqrt{r}}. \end{aligned}$$

It follows that $|\mathcal{S}| \leq (1 + O(1/\sqrt{b}))|\mathbf{L}|$. We can set b to the order of $1/\varepsilon^2$, and we get the following.

Theorem 3.6 *Given a set of n unweighted admissible regions in the plane, any b -locally optimal independent set has size $\geq (1 - O(1/\sqrt{b}))\text{opt}$, where opt is the size of the maximum independent set of the objects. In particular, one can compute an independent set of size $\geq (1 - \varepsilon)\text{opt}$, in time $n^{O(1/\varepsilon^2)}$.*

3.2.3 Analysis for fat objects in any fixed dimension

We show that the same algorithm gives a PTAS for the case when the objects in F are fat. This result in fact holds in any fixed dimension d . For our purposes, we use the following definition of fatness: the objects in F are *fat* if for every axis-aligned hypercube B of side length r , we can find a constant number c of points such that every object that intersects B and has diameter at least r contains one of the chosen points.

Smith and Wormald [SW98] proved a family of geometric separator theorems, one version of which will be useful for us and is stated below (see also [Cha03]):

Lemma 3.7 ([SW98]) *Given a collection of n fat objects in a fixed dimension d with constant maximum depth, there exists an axis-aligned hypercube B such that at most $2n/3$ objects are inside B , at most $2n/3$ objects are outside B , and at most $O(n^{1-1/d})$ objects intersect the boundary of B .*

We need the following extension of Smith and Wormald’s separator theorem to multiple clusters (whose proof is similar to the extension of the standard planar separator theorem in [Fre87]):

Lemma 3.8 *There are constants c_1, c_2, c_3 and c_4 , such that for any intersection graph $G = (V, E)$ of n fat objects in a fixed dimension d with constant maximum depth, and a parameter r , one can find a set of $X \subseteq V$ of size at most $c_1 n/r^{1/d}$, and a partition of $V \setminus X$ into n/r sets $V_1, \dots, V_{n/r}$, satisfying: (i) $|V_i| \leq c_2 r$, (ii) $\Gamma(V_i) \cap V_j = \emptyset$, for $i \neq j$, and (iii) $\sum_i |\Gamma(V_i) \cap X| \leq c_3 n/r^{1/d}$, and (iv) $|\Gamma(V_i) \cap X| \leq c_4 r$.*

Proof: Assume that all objects are unmarked initially. We describe a recursive procedure for a given set S of objects. If $|S| \leq c_2 r$, then S is a “leaf” subset and we stop the recursion. Otherwise, we apply Lemma 3.7. Let S' and S'' be the subset of all objects inside and outside the separator hypercube B respectively. Let \widehat{S} be the subset of all objects intersecting the boundary of B . We mark the objects in \widehat{S} and recursively run the procedure for the subset $S' \cup \widehat{S}$ and for the subset $S'' \cup \widehat{S}$.

Note that some objects may be marked more than once. Let X be the set of all objects that have been marked at least once. For each leaf subset S_i , generate a subset V_i of all unmarked objects in S_i . Property (i) is obvious. Properties (ii) and (iv) hold, because the unmarked objects in each leaf subset S_i can only intersect objects within S_i and cannot intersect unmarked objects in other S_j ’s.

The total number of marks satisfies the recurrence $T(n) = 0$ if $n \leq c_2 r$, and

$$T(n) \leq \max_{\substack{n_1, n_2 \leq 2n/3 \\ n_1 + n_2 \leq n}} T(n_1 + O(n^{1-1/d})) + T(n_2 + O(n^{1-1/d})) + O(n^{1-1/d})$$

otherwise. The solution is $T(n) = O(n/r^{1/d})$. Thus, we have $|X| = O(n/r^{1/d})$. Furthermore, for each object $f \in X$, the number of leaf subsets that f is in is equal to 1 plus the number of marks that f receives. Thus, (iii) follows. ■

Let \mathcal{S} be the optimal solution and \mathcal{L} be a b -locally optimal solution. Consider the bipartite intersection graph G of $\mathcal{S} \cup \mathcal{L}$, which has maximum depth 2. We proceed as in the proof from Section 3.2.2, using Lemma 3.8 instead of Lemma 3.5. Note that (iii)–(iv) are weaker properties but are sufficient for the same proof to go through. The only differences are that square roots are replaced by d th roots, and we now set $r = b/(c_2 + c_4)$, so that $|\overline{\Gamma}(V_i)| \leq c_2r + c_4r < b$. We conclude:

Theorem 3.9 *Given a set of n fat objects in a fixed dimension d , any b -locally optimal independent set has size $\geq (1 - O(1/b^{1/d}))\text{opt}$, where opt is the size of the maximum independent set of the objects. In particular, one can compute an independent set of size $\geq (1 - \varepsilon)\text{opt}$, in time $n^{O(1/\varepsilon^d)}$.*

4 Approximation by LP Relaxation: Weighted Case

4.1 The algorithm

We are interested in computing a maximum-weight independent set of the objects in $\mathbf{F} = \{f_1, \dots, f_n\}$, with weights w_1, \dots, w_n , respectively. To this end, let us solve the following LP relaxation:

$$\begin{aligned} \max \quad & \sum_{i=1}^n w_i x_i & (1) \\ & \sum_{f_i \ni \mathbf{p}} x_i \leq 1 & \forall \mathbf{p} \in \mathcal{V}(\mathbf{F}) \\ & 0 \leq x_i \leq 1, \end{aligned}$$

where $\mathcal{V}(\mathbf{F})$ denotes the set of vertices of the arrangement $\mathcal{A}(\mathbf{F})$.

In the following, x_i will refer to the value assigned to this variable by the solution of the LP. Similarly, $\text{Opt} = \sum_i w_i x_i$ will denote the weight of the relaxed optimal solution, which is at least the weight opt of the optimal integral solution.

We will assume, for the time being, that no two objects of \mathbf{F} fully contain each other.

For every object f_i , let its **resistance** be the total sum of the values of the objects that intersect it. Formally, we have that

$$\eta(f_i, \mathbf{F}) = \sum_{\substack{f_j \in \mathbf{F} \setminus \{f_i\} \\ f_i \cap f_j \neq \emptyset}} x_j.$$

We pick the object in \mathbf{F} with minimal resistance, and set it as the first element in the permutation Π of the objects. We compute the permutation by performing this “extract-min”, with the variant that objects that are already in the permutation are ignored. Formally,

if we computed the first i objects in the permutation to be $\Pi_i = \langle \pi_1, \dots, \pi_i \rangle$, then the $(i+1)$ th object π_{i+1} is the one realizing

$$\eta_{i+1} = \min_{f \in \mathcal{F} \setminus \Pi_i} \eta(f, \mathcal{F} \setminus \Pi_i). \quad (2)$$

The algorithm starts with an empty candidate set C and an empty independent set I , and scans the objects according to the permutation in reverse order. At the i th stage, the algorithm first decides whether to put the object π_{n-i} in C , by flipping a coin that with probability $x(\pi_{n-i})/\tau$ comes up heads, where $x(\pi_{n-i})$ is the value assigned by the LP to the object π_{n-i} and τ is some parameter to be determined shortly. If π_{n-i} is put into C then we further check whether π_{n-i} intersects any of the objects already added to the independent set I . If it does not intersect any objects in I , then it adds π_{n-i} to I and continues to the next iteration.

In the end of the execution, the set I is returned as the desired independent set.

4.2 Analysis

Let \mathcal{F} be a set of n objects in the plane, and let $u(m)$ be the maximum union complexity of $m \leq n$ objects of \mathcal{F} . Furthermore, we assume that the function $u(\cdot)$ is a monotone increasing function which is well behaved; namely, $u(n)/n$ is a non-decreasing function, and there exists a constant c , such that $u(xr) \leq c u(r)$, for any r and $1 \leq x \leq 2$. In the following, a vertex \mathbf{p} of $\mathcal{V}(\mathcal{H})$ is denoted by (\mathbf{p}, i, j) , to indicate that it is the result of the intersection of the i th and j th object.

The key to our analysis lies in the following inequality, which we prove by adapting the Clarkson technique [CS89].

Lemma 4.1 *Let \mathcal{H} be any subset of \mathcal{F} . Then $\sum_{(\mathbf{p}, i, j) \in \mathcal{V}(\mathcal{H})} x_i x_j = O(u(\mathcal{E}(\mathcal{H})))$, where $\mathcal{E}(\mathcal{H}) = \sum_{f \in \mathcal{H}} x(f)$.*

Proof: Consider a random sample \mathcal{R}' of \mathcal{H} , where an object f_i is being picked up with probability $x_i/2$. Clearly, we have that $(\mathbf{p}, i, j) \in \mathcal{V}(\mathcal{F})$ appears on the boundary of the union of the objects of \mathcal{R}' , if and only if f_i and f_j are being picked, and none of the objects that cover \mathbf{p} are being chosen into \mathcal{R}' . In particular, let $\mathcal{U}(\mathcal{R}')$ denote the vertices on the boundary of the union of the objects of \mathcal{R}' . We have

$$\Pr \left[(\mathbf{p}, i, j) \in \mathcal{U}(\mathcal{R}') \right] = \frac{x_i}{2} \cdot \frac{x_j}{2} \prod_{\substack{f_k \ni \mathbf{p}, \\ k \neq i, k \neq j}} \left(1 - \frac{x_k}{2} \right) \geq \frac{x_i x_j}{4} \left[1 - \sum_{\substack{f_k \ni \mathbf{p}, \\ k \neq i, k \neq j}} \frac{x_k}{2} \right] \geq \frac{x_i x_j}{8},$$

by the inequality $\prod_k (1 - a_k) \geq 1 - \sum_k a_k$ for $a_k \in [0, 1]$, since $\sum_{f_k \ni \mathbf{p}} x_k \leq 1$ (as the LP solution is valid). On the other hand, the number of vertices on the union is $|\mathcal{U}(\mathcal{R}')| \leq u(|\mathcal{R}'|)$.

Thus,

$$\sum_{(p,i,j) \in \mathcal{V}(\mathbf{F})} \frac{x_i x_j}{8} \leq \sum_{(p,i,j) \in \mathcal{V}(\mathbf{F})} \Pr[(p, i, j) \in \mathcal{U}(\mathbf{R}')] = \mathbf{E}[|\mathcal{U}(\mathbf{R}')|] \leq \mathbf{E}[\mathbf{u}(|\mathbf{R}'|)].$$

To bound last expression, observe that $\mu = \mathbf{E}[|\mathbf{R}'|] = \mathcal{E}(\mathbf{H})/2$. Furthermore, by Chernoff inequality, $\Pr[|\mathbf{R}'| > (t+1)\mu] \leq 2^{-t}$. Thus, $\mathbf{E}[\mathbf{u}(|\mathbf{R}'|)] \leq \sum_{t=1}^{\infty} 2^{-t+1} \mathbf{u}(t\mu) = \sum_{t=1}^{\infty} 2^{-t+1} t^{O(1)} \mathbf{u}(\mu) = O(\mathbf{u}(\mu)) = O(\mathbf{u}(\mathcal{E}(\mathbf{H})))$, since $\mathbf{u}(\cdot)$ is well behaved. ■

Lemma 4.2 *For any i , the resistance of the i th object π_i (as defined by Eq. (2)) is $\eta_i = O\left(\frac{\mathbf{u}(\mathcal{E}(\mathbf{F}))}{\mathcal{E}(\mathbf{F})}\right)$.*

Proof: Fix an i , and let $\mathbf{K} = \mathbf{F} \setminus \{\pi_1, \dots, \pi_{i-1}\}$. By Lemma 4.1,

$$\begin{aligned} \sum_{f_j \in \mathbf{K}} x_j \eta(f_j, \mathbf{K}) &= 2 \sum_{(p,i,j) \in \mathcal{V}(\mathbf{K})} x_i x_j = O(\mathbf{u}(\mathcal{E}(\mathbf{K}))) \\ \implies \sum_{f_j \in \mathbf{K}} \frac{x_j}{\mathcal{E}(\mathbf{K})} \eta(f_j, \mathbf{K}) &= O\left(\frac{\mathbf{u}(\mathcal{E}(\mathbf{K}))}{\mathcal{E}(\mathbf{K})}\right) = O\left(\frac{\mathbf{u}(\mathcal{E}(\mathbf{F}))}{\mathcal{E}(\mathbf{F})}\right) \end{aligned}$$

by the monotonicity of $\mathbf{u}(n)/n$, where $\mathcal{E}(\mathbf{K}) = \sum_{f_j \in \mathbf{K}} x_j$. It follows that $\min_{f_j \in \mathbf{K}} \eta(f_j, \mathbf{K}) = O\left(\frac{\mathbf{u}(\mathcal{E}(\mathbf{F}))}{\mathcal{E}(\mathbf{F})}\right)$. ■

Lemma 4.3 *For a sufficiently large constant c , setting $\tau = c \frac{\mathbf{u}(\mathcal{E}(\mathbf{F}))}{\mathcal{E}(\mathbf{F})}$, the algorithm in Section 4.1 outputs in expectation an independent set of weight $\Omega((n/\mathbf{u}(n))\text{Opt})$.*

Proof: Indeed, the j th object in the permutation is added to C with probability x_j/τ . Let \mathbf{K} be the set of all objects of \mathbf{F} that were already considered and intersect f_j . Clearly, $\mathcal{E}(\mathbf{K})$ is exactly the resistance η_j of f_j . Furthermore by picking c large enough, we have that $\mathcal{E}(\mathbf{K}) = \eta_j \leq \tau/2$ by Lemma 4.2. This implies that

$$\begin{aligned} \Pr[f_j \in I \mid f_j \in C] &= \Pr[\mathbf{K} \cap C = \emptyset \mid f_j \in C] = \Pr[\mathbf{K} \cap C = \emptyset] = \prod_{f_j \in \mathbf{K}} \left(1 - \frac{x_j}{\tau}\right) \\ &\geq 1 - \sum_{f_j \in \mathbf{K}} \frac{x_j}{\tau} = 1 - \frac{\mathcal{E}(\mathbf{K})}{\tau} \geq \frac{1}{2}, \end{aligned}$$

by the inequality $\prod_k (1 - a_k) \geq 1 - \sum_k a_k$ for $a_k \in [0, 1]$. Now, we have that

$$y_j = \Pr[f_j \in I] = \Pr[f_j \in I \mid f_j \in C] \cdot \Pr[f_j \in C] \geq \frac{x_j}{2\tau}.$$

As such, the expected value of the independent set output is

$$\sum_j y_j w_j = \sum_j \frac{x_j}{2\tau} w_j = \Omega\left(\frac{\text{Opt}}{\tau}\right) = \Omega\left(\frac{n}{\mathbf{u}(n)} \text{Opt}\right),$$

as $\mathcal{E}(\mathbf{F}) \leq n$. ■

4.3 Remarks

Variation. In the conference version of this paper [CH09], we proposed a different variant of the algorithm, where instead of ordering the objects by increasing resistance, we order the objects by decreasing weights. An advantage of the resistance-based algorithm is that it is oblivious to (i.e., does not look at) the input weights. This feature is shared, for example, by Varadarajan’s recent algorithm for weighted geometric set cover via “quasi-random sampling” [Var10]. Another advantage of the resistance-based algorithm is its extendibility to other settings; see Section 4.6.

Derandomization. The variance of the expected weight of the returned independent set I could be high, but fortunately the algorithm can be derandomized by the standard method of conditional probabilities/expectations [MR95]. To this end, observe that the above analysis provide us with a constructive way to estimate the weight of the generated solution. It is now straightforward to decide for each region whether to include it or not inside the generated solution, using conditional probabilities. Indeed, for each object we compute the expected weight of the solution if it is include in the solution, and if it is not included in the solution, and pick the one that has higher value.

Coping with object containment. We have assumed that no object is fully contained in another, but this assumption can be removed by adding the constraint $\sum_{f_i \subset f_j} x_j \leq 1$ for each i to the LP. Then, for any subset H of F , we have

$$\sum_{\substack{f_i, f_j \in H \\ f_i \subset f_j}} x_i x_j \leq \mathcal{E}(H),$$

and so Lemma 4.1 still holds. The rest of the analysis then holds verbatim.

Time to solve the LP. This LP is a packing LP with $O(n^2)$ inequalities, and n variables. As such, it can be $(1 + \varepsilon)$ -approximated in $O(n^3 + \varepsilon^{-2} n^2 \log n) = O(n^3)$ by a randomized algorithm that succeeds with high probability [KY07]. For our purposes, it is sufficient to set ε to be a sufficient small constant, say $\varepsilon = 10^{-4}$.

We have thus proved:

Theorem 4.4 *Given a set of n weighted objects in the plane with union complexity $O(u(n))$, one can compute an independent set of total weight $\Omega((n/u(n))\text{opt})$, where opt is the maximum weight over all independent sets of the objects. The running time of the randomized algorithm is $O(n^3)$, and polynomial for the deterministic version.*

The running time of the deterministic algorithm of Theorem 4.4 is dominated by the time it takes to deterministically solve (approximately) the LP. One can use the ellipsoid algorithm to this end, but faster algorithm are known, see [KY07] and references therein.

Corollary 4.5 *Given a set of n weighted pseudo-disks in the plane, one can compute, in $O(n^3)$ time, a constant factor approximation to the maximum weight independent set of pseudo-disks.*

Theorem 4.4 can be applied to cases where the union complexity is low. Even in the case of fat objects, where PTASs are known [Cha03, EJS05], the above approach is still interesting as it can be extended to more general settings, as noted in Section 4.6.

4.4 A combinatorial result: piercing number

In the unweighted case, we obtain the following result as a byproduct:

Theorem 4.6 *Given a set of n pseudo-disks in the plane, let opt be the size of the maximum independent set and let opt' be the size of the minimum set of points that pierce all the pseudo-disks. Then $\text{opt} = \Omega(\text{opt}')$.*

Proof: By the preceding analysis, we have $\text{opt} = \Omega(\text{Opt})$, i.e., the integrality gap of our LP is a constant. (Here, all the weights are equal to 1.)

For piercing, the LP relaxation is

$$\begin{aligned} \min \quad & \sum_{\mathbf{p} \in \mathcal{V}(S)} y_{\mathbf{p}} \\ & \sum_{\mathbf{p} \in f_i} y_{\mathbf{p}} \geq 1 && \forall i = 1, \dots, n \\ & 0 \leq y_{\mathbf{p}} \leq 1. \end{aligned}$$

Let Opt' be the value of this LP. Known analysis [Lon01, ERS05] implies that the integrality gap of this LP is constant if there exist ε -nets of linear size for a corresponding class of hypergraphs formed by objects in F and points in $\mathcal{V}(S)$. Pyrga and Ray [PR08, Theorem 12] obtained such an existence proof for this (“primal”) hypergraph for pseudo-disks. Thus, $\text{opt}' = O(\text{Opt}')$.

To conclude, observe that the two LPs are precisely the dual of each other, and so $\text{Opt} = \text{Opt}'$. ■

4.5 A discrete version of the independent set problem

We now show that our algorithm can be extended to solve a variant of the independent set problem where we are given not only a set F of n weighted objects but also a set P of m points. The goal is to select a maximum-weight subset $\mathcal{S} \subseteq F$ such that each point $\mathbf{p} \in P$ is contained in at most one object of \mathcal{S} . (The original problem corresponds to the case where P is the entire plane.) Unlike in the original independent set problem, it is not clear if local search yields good approximation here, even in the unweighted case.

We can use the same LP as in Section 4.1 to solve this problem, except that we now have a constraint for each $p \in P$ instead of each $p \in \mathcal{V}(F)$. In the rest of the algorithm and analysis, we just reinterpret “ $f_i \cap f_j \neq \emptyset$ ” to mean “ $f_i \cap f_j \cap P \neq \emptyset$ ”.

Lemma 4.1 is now replaced by the following.

Lemma 4.7 *Let H be any subset of F . Then
$$\sum_{\substack{f_i \cap f_j \cap P \neq \emptyset \\ f_i, f_j \in H}} x_i x_j = O(u(\mathcal{E}(H))).$$*

Proof: Consider a random sample R of H , where an object f_i is being picked up with probability $x_i/2$. Let $\mathcal{VD}(R)$ be the cells in the vertical decomposition of the complement of the union $\mathcal{U}(R)$. For a cell $\Delta \in \mathcal{VD}(R)$, let $x_\Delta = \sum_{f_i \in H, f_i \cap \text{int}(\Delta) \neq \emptyset} x_i$ be the total energy of the objects of H that intersects the interior of Δ . A minor modification of the analysis of Clarkson [CS89] implies that, for any constant c , it holds $\mathbf{E} \left[\sum_{\Delta \in \mathcal{VD}(R)} (x_\Delta)^c \right] = O(\mathbf{E}[u(\mathcal{E}(H))])$.

A point of $p \in P$ is *active* for R , if it is outside the union of objects of R . Let P' be the set of active points in P . We have that

$$\sum_{f_i \cap f_j \cap P' \neq \emptyset} x_i x_j \leq \sum_{\Delta \in \mathcal{VD}(R)} x_\Delta^2.$$

Furthermore, by arguing as in Lemma 4.1, every point $p \in P$ has probability at least $\prod_{f_i \in H, p \in f_i} (1 - x_i/2) \geq 1/2$ to be active. Thus,

$$\frac{1}{2} \sum_{f_i \cap f_j \cap P \neq \emptyset} x_i x_j \leq \mathbf{E} \left[\sum_{f_i \cap f_j \cap P' \neq \emptyset} x_i x_j \right] \leq \mathbf{E} \left[\sum_{\Delta \in \mathcal{VD}(R)} x_\Delta^2 \right] = O(\mathbf{E}[u(\mathcal{E}(H))]) = O(u(\mathcal{E}(H))),$$

again, by arguing as in Lemma 4.1. ■

The above proof is inspired by a proof from [AACS98]. There is an alternative argument based on *shallow cuttings*, but the known proof for the existence of such cuttings requires a more complicated sampling analysis [Mat92].

The rest of the analysis then goes through unchanged. We therefore obtain an $O(1)$ -approximation algorithm for the discrete independent set problem for unweighted or weighted pseudo-disks in the plane.

4.6 Contention resolution and submodular functions

The algorithm of Theorem 4.4 can be interpreted as a contention resolution scheme; see Chekuri *et al.* [CVZ11] for details. The basic idea is that given a feasible fractional solution $x \in [0, 1]^n$, a **contention resolution scheme** scales down every coordinate of x (by some constant b) such that given a random sample C of the objects according to x (i.e., the i th object f_i is picked with probability bx_i), the contention resolution scheme computes (in our

case) an independent set I such that $\Pr[f_i \in I \mid f_i \in C] \geq c$, for some positive constant c . The proof of Lemma 4.3 implies exactly this property in our case.

As such, we can apply the results of Chekuri *et al.* [CVZ11] to our settings. In particular, they show that one can obtain constant approximation to the optimal solution, when considering independence constraints and submodular target function. Intuitively, submodularity captures the diminishing-returns nature of many optimization problems. Formally, a function $g : 2^{\mathbf{F}} \rightarrow \mathbb{R}$ is **submodular** if $g(X \cup Y) + g(X \cap Y) \leq g(X) + g(Y)$, for any $X, Y \subseteq \mathbf{F}$.

As a concrete example, consider a situation where each object in \mathbf{F} represents a coverage area by a single antenna. If a point is contained inside such an object, it is fully serviced. However, even if it is not contained in a object, it might get some reduced coverage from the closest object in the chosen set. In particular, let $\nu(r)$ be some coverage function which is the amount of coverage a point gets if it is at distance r from the closest object in the current set I . We assume here $\nu(\cdot)$ is a monotone decreasing function. Because of interference between antennas we require that the regions these antennas represent do not intersect (i.e., the set of antennas chosen needs to be an independent set).

Lemma 4.8 *Let \mathbf{P} be a set of points, and let \mathbf{F} be a set of objects in the plane. Let $\nu(\cdot)$ be a monotone decreasing function. For a subset $\mathbf{H} \subseteq \mathbf{F}$, consider the target function*

$$\alpha_{\mathbf{P}}(\mathbf{H}) = \sum_{\mathbf{p} \in \mathbf{P}} \nu(\mathbf{d}_{\mathbf{p}}(\mathbf{H})),$$

where $\mathbf{d}_{\mathbf{p}}(\mathbf{H})$ is the distance of \mathbf{p} to its nearest neighbor in \mathbf{H} . Then the function $\alpha_{\mathbf{P}}(\mathbf{H})$ is submodular.

Proof: The proof is not hard and is included for the sake of completeness. For a point $\mathbf{p} \in \mathbf{P}$, it is sufficient to prove that the function $\nu(\mathbf{d}_{\mathbf{p}}(\mathbf{H}))$ is submodular, as $\alpha_{\mathbf{P}}(\mathbf{H})$ is just the sum of these functions, and a sum of submodular functions is submodular.

To prove the latter, it is sufficient to prove that for any sets $X \subseteq Y \subseteq \mathbf{F}$, and an object $f \in \mathbf{F} \setminus Y$,

$$\nu(\mathbf{d}_{\mathbf{p}}(X \cup \{f\})) - \nu(\mathbf{d}_{\mathbf{p}}(X)) \geq \nu(\mathbf{d}_{\mathbf{p}}(Y \cup \{f\})) - \nu(\mathbf{d}_{\mathbf{p}}(Y)).$$

To this end, let x and y be the closest objects to \mathbf{p} in X and Y , respectively. Similarly, let ℓ_x, ℓ_y, ℓ_f be the distance of \mathbf{p} to x, y and f , respectively. The above then becomes

$$\nu(\min(\ell_x, \ell_f)) - \nu(\ell_x) \geq \nu(\min(\ell_y, \ell_f)) - \nu(\ell_y). \quad (3)$$

Observe that as $X \subseteq Y$, it holds that $\ell_x \geq \ell_y$, so $\nu(\ell_y) \geq \nu(\ell_x)$ as ν is monotone decreasing. Now, one of the following holds:

- If $\ell_f \leq \ell_y \leq \ell_x$ then Eq. (3) becomes $\nu(\ell_f) - \nu(\ell_x) \geq \nu(\ell_f) - \nu(\ell_y)$, which holds.
- If $\ell_y \leq \ell_f \leq \ell_x$ then Eq. (3) becomes $\nu(\ell_f) - \nu(\ell_x) \geq \nu(\ell_y) - \nu(\ell_y)$, which is equivalent to $\nu(\ell_f) \geq \nu(\ell_x)$. This in turn holds by the decreasing monotonicity of ν .
- If $\ell_y \leq \ell_x \leq \ell_f$ then Eq. (3) becomes $0 = \nu(\ell_x) - \nu(\ell_x) \geq \nu(\ell_y) - \nu(\ell_y) = 0$.

We conclude that $\alpha_{\mathbf{P}}(\cdot)$ is submodular. ■

To solve our problem, using the framework of Chekuri *et al.* [CVZ11], we need the following:

- (A) The target function is indeed submodular and can be computed efficiently. This is Lemma 4.8.
- (B) State an LP that solves the fractional problem (and its polytope contains the optimal integral solution). This is just the original LP, see Eq. (1).
- (C) Observe that our rounding (i.e., contention resolution) scheme is still applicable in this case. This follows by Lemma 4.3.

One can now plug this into the algorithm of Chekuri *et al.* [CVZ11] and get an $\Omega(\alpha)$ -approximation algorithm, where α is the rounding scheme gap. The algorithm of Chekuri *et al.* [CVZ11] uses a continuous optimization to find the maximum of a multi-linear extension of the target function inside the feasible polytope, and then uses this fractional value with the rounding scheme to get the desired approximation.

We thus get the following.

Problem 4.9 *Let \mathbf{P} be a set of n points in the plane, and let \mathbf{F} be a set of m objects in the plane. Let $\nu(r)$ be a monotone decreasing function, which returns the amount of coverage a point gets if it is at distance r from one of the regions of \mathbf{F} . Consider the scoring function that for an independent set $\mathbf{H} \subseteq \mathbf{F}$ returns the total coverage it provides; that is,*

$$\alpha_{\mathbf{P}}(\mathbf{H}) = \sum_{\mathbf{p} \in \mathbf{P}} \nu(d_{\mathbf{p}}(\mathbf{H})).$$

*We refer to the problem of computing the independent set maximizing this function as the **partial coverage problem**.*

Theorem 4.10 *Given a set of n points in the plane, and set m of unweighted objects in the plane with union complexity $O(u(n))$, one can compute, in polynomial time, an independent set. Furthermore, this independent set provides an $\Omega(n/u(n))$ -approximation to the optimal solution of the partial coverage problem.*

Observe that the above algorithm applies for any pricing function that is submodular. In particular, one can easily encode into this function weights for the ranges, or other similar considerations.

5 Weighted Rectangles

5.1 The algorithm

For the (original) independent set problem in the case of weighted axis-aligned rectangles, we can solve the same LP, where the set $\mathcal{V}(\mathbf{F})$ contains both intersection points and corners of the given rectangles.

Define two subgraphs \mathcal{G}_1 and \mathcal{G}_2 of the intersection graph: if the boundaries of f_i and f_j intersect zero or two times, put $f_i f_j$ in \mathcal{G}_1 ; if the boundaries intersect four times instead, put $f_i f_j$ in \mathcal{G}_2 .

We first extract an independent set I of \mathcal{G}_1 using the algorithm of Theorem 4.4.

It is well known (e.g., see [AG60]) that \mathcal{G}_2 forms a perfect graph (specifically, a comparability graph), so find a Δ -coloring of the rectangles of I in \mathcal{G}_2 , where Δ denotes the maximum clique size, i.e., the maximum depth in $\mathcal{A}(\mathbf{R})$. Let I' be the color subclass of I of the largest total weight. Clearly, the objects in I' are independent, and we output this set.

5.2 Analysis

As in Lemma 4.1, let H be any subset of F . Observe that if $f_i f_j \in \mathcal{G}_1$, then f_i contains a corner of f_j or vice versa. Letting V_j denote the corners of f_j , we have

$$\begin{aligned} \sum_{f_i, f_j \in H, f_i f_j \in \mathcal{G}_1} x_i x_j &\leq \sum_{f_j \in H} \sum_{p \in V_j} \sum_{f_i \in H, p \in f_i} x_i x_j \leq \sum_{f_j \in H} \sum_{p \in V_j} x_j \sum_{f_i \in H, p \in f_i} x_i \leq \sum_{f_j \in H} \sum_{p \in V_j} x_j = \sum_{f_j \in H} 4x_j \\ &= 4\mathcal{E}(H). \end{aligned}$$

Applying the same analysis as before, we conclude that the expected total weight of I (i.e., the independent set of \mathcal{G}_1) computed by the algorithm is of size $\Omega(\text{Opt})$.

To analyze I' , we need a new lemma which bounds the maximum depth of \mathbf{R} :

Lemma 5.1 $\Delta = O(\log n / \log \log n)$ with probability at least $1 - 1/n$.

Proof: Fix a parameter $t > 1$. Fix a point $\mathbf{p} \in \mathcal{V}(\mathbf{F})$. The depth of \mathbf{p} in $\mathcal{A}(\mathbf{R})$, denoted by $\text{depth}(\mathbf{p}, \mathbf{R})$, is a sum of independent 0-1 random variables with overall mean $\mu = \sum_{p \in f_i} x_i \leq 1$. By the Chernoff bound [MR95, page 68],

$$\Pr[\text{depth}(\mathbf{p}, \mathbf{R}) > (1 + \delta)\mu] < \left[\frac{e^\delta}{(1 + \delta)^{1 + \delta}} \right]^\mu$$

for any $\delta > 0$ (possibly large). By setting δ so that $t = (1 + \delta)\mu$, this probability becomes less than $(e/t)^t$. Since $|\mathcal{V}(\mathbf{F})| = O(n^2)$, the probability that $\Delta > t$ is at most $O((e/t)^t n^2)$, which is at most $1/n$ by setting the value of t to be $\Theta(\log n / \log \log n)$. ■

By construction of I' , we know that

$$\sum_{f_i \in I'} w_i \geq \frac{1}{\Delta} \sum_{f_i \in I} w_i \geq \frac{1}{t} \sum_{f_i \in I} w_i - \text{Opt} \cdot 1_{\Delta > t}$$

where 1_A denotes the indicator variable for event A . With $t = \Theta(\log n / \log \log n)$, we conclude that

$$\begin{aligned} \mathbf{E} \left[\sum_{f_i \in I'} w_i \right] &\geq \Omega(\log \log n / \log n) \mathbf{E} \left[\sum_{f_i \in I} w_i \right] - \text{Opt}/n \\ &\geq \Omega(\log \log n / \log n) \cdot \text{Opt}. \end{aligned}$$

5.3 Remarks

Derandomization. This algorithm can also be derandomized by the method of conditional expectations. The trick is to consider the following random variable

$$Z := \frac{1}{t} \sum_{f_i \in I} w_i - \text{Opt} \cdot \sum_{\mathbf{p} \in \mathcal{V}(\mathbf{F})} (1 + \delta_{\mathbf{p}})^{\text{depth}(\mathbf{p}, \mathbf{R}) - t},$$

where $\delta_{\mathbf{p}}$ is the δ from the proof of Lemma 5.1 and t is the same as before. This variable Z lower-bounds $\sum_{f_i \in I} w_i$ (the bound is trivially true if $\Delta > t$, since Z would be negative). Our analysis still implies that $\mathbf{E}[Z] \geq \Omega(\log \log n / \log n) \cdot \text{Opt}$ (since the standard proof of the Chernoff bound [MR95, pages 68–69] actually shows that

$$\Pr[\text{depth}(\mathbf{p}, \mathbf{R}) > (1 + \delta)\mu] < \frac{\mathbf{E}[(1 + \delta)^{\text{depth}(\mathbf{p}, \mathbf{R})}]}{(1 + \delta)^{(1 + \delta)\mu}} \leq \left[\frac{e^\delta}{(1 + \delta)^{1 + \delta}} \right]^\mu,$$

which for $\delta = \delta_{\mathbf{p}}$ and $\mu \leq 1$ implies $\mathbf{E}[(1 + \delta_{\mathbf{p}})^{\text{depth}(\mathbf{p}, \mathbf{R}) - t}] < (e/t)^t$. The advantage of working with Z is that we can calculate $\mathbf{E}[Z]$ exactly in polynomial time, even when conditioned to the events that some objects are known to be in or not in \mathbf{R} (since $\text{depth}(\mathbf{p}, \mathbf{R})$ is a sum of independent 0-1 random variables, making $(1 + \delta_{\mathbf{p}})^{\text{depth}(\mathbf{p}, \mathbf{R})}$ a product of independent random variables).

We have thus proved:

Theorem 5.2 *Given a set of n weighted axis-aligned boxes in the plane, one can compute in polynomial time an independent set of total weight $\Omega(\log \log n / \log n) \cdot \text{opt}$, where opt is the maximum weight over all independent sets of the objects.*

Higher dimensions. By a standard divide-and-conquer method [AvKS98], we get an approximation factor of $O(\log^{d-1} n / \log \log n)$ for weighted axis-aligned boxes in any constant dimension d .

Acknowledgments

We thank Esther Ezra for discussions on the discrete version of the independent set problem considered in Section 4.3. The somewhat cleaner presentation in the paper, compared to the preliminary version [CH09], was suggested by Chandra Chekuri. The results of Section 4.6 were inspired by discussions with Chandra Chekuri. We also thank the anonymous referee for their comments.

References

- [AACS98] P. K. Agarwal, B. Aronov, T. M. Chan, and M. Sharir. On levels in arrangements of lines, segments, planes, and triangles. *Discrete Comput. Geom.*, 19:315–331, 1998.

- [AC06] P. Afshani and T. M. Chan. Dynamic connectivity for axis-parallel rectangles. In *Proc. 14th European Sympos. Algorithms*, volume 4168 of *Lecture Notes Comput. Sci.*, pages 16–27, 2006.
- [AG60] E. Asplund and B. Gröbaum. On a coloring problem. *Math. Scand.*, 8:181–188, 1960.
- [AM06] P. K. Agarwal and N. H. Mustafa. Independent set of intersection graphs of convex objects in 2D. *Comput. Geom. Theory Appl.*, 34(2):83–95, 2006.
- [APS08] P. K. Agarwal, J. Pach, and M. Sharir. State of the union—of geometric objects. In J. E. Goodman, J. Pach, and R. Pollack, editors, *Surveys in Discrete and Computational Geometry Twenty Years Later*, volume 453 of *Contemporary Mathematics*, pages 9–48. AMS, 2008.
- [Aro98] S. Arora. Polynomial time approximation schemes for Euclidean TSP and other geometric problems. *J. Assoc. Comput. Mach.*, 45(5):753–782, Sep 1998.
- [AS00] N. Alon and J. H. Spencer. *The Probabilistic Method*. Wiley Inter-Science, 2nd edition, 2000.
- [AvKS98] P. K. Agarwal, M. van Kreveld, and S. Suri. Label placement by maximum independent set in rectangles. *Comput. Geom. Theory Appl.*, 11:209–218, 1998.
- [Bak94] B. S. Baker. Approximation algorithms for NP-complete problems on planar graphs. *J. Assoc. Comput. Mach.*, 41:153–180, 1994.
- [BDMR01] P. Berman, B. DasGupta, S. Muthukrishnan, and S. Ramaswami. Efficient approximation algorithms for tiling and packing problems with rectangles. *J. Algorithms*, 41:443–470, 2001.
- [BF99] P. Berman and T. Fujito. On approximation properties of the independent set problem for low degree graphs. *Theo. Comp. Sci.*, 32(2):115–132, 1999.
- [BG95] H. Brönnimann and M. T. Goodrich. Almost optimal set covers in finite VC-dimension. *Discrete Comput. Geom.*, 14:263–279, 1995.
- [BYBFR04] R. Bar-Yehuda, K. Bendel, A. Freund, and D. Rawitz. Local ratio: A unified framework for approximation algorithms. in memoriam: Shimon even 1935-2004. *ACM Comput. Surv.*, 36:422–463, December 2004.
- [CC09] P. Chalermsook and J. Chuzhoy. Maximum independent set of rectangles. In *Proc. 20th ACM-SIAM Sympos. Discrete Algorithms*, pages 892–901, 2009.
- [CH09] T. M. Chan and S. Har-Peled. Approximation algorithms for maximum independent set of pseudo-disks. In *Proc. 25th Annu. ACM Sympos. Comput. Geom.*, pages 333–340, 2009. cs.uiuc.edu/~sariel/papers/08/w_indep.

- [Cha03] T. M. Chan. Polynomial-time approximation schemes for packing and piercing fat objects. *J. Algorithms*, 46(2):178–189, 2003.
- [Cha04] T. M. Chan. A note on maximum independent sets in rectangle intersection graphs. *Inform. Process. Lett.*, 89:19–23, 2004.
- [CS89] K. L. Clarkson and P. W. Shor. Applications of random sampling in computational geometry, II. *Discrete Comput. Geom.*, 4:387–421, 1989.
- [CV07] K. L. Clarkson and K. R. Varadarajan. Improved approximation algorithms for geometric set cover. *Discrete Comput. Geom.*, 37(1):43–58, 2007.
- [CVZ11] C. Chekuri, J. Vondrák, and R. Zenklusen. Submodular function maximization via the multilinear relaxation and contention resolution schemes. In *Proc. 43rd Annu. ACM Sympos. Theory Comput.*, 2011. To appear.
- [EJS05] T. Erlebach, K. Jansen, and E. Seidel. Polynomial-time approximation schemes for geometric intersection graphs. *SIAM J. Comput.*, 34(6):1302–1323, 2005.
- [EKNS00] A. Efrat, M. J. Katz, F. Nielsen, and M. Sharir. Dynamic data structures for fat objects and their applications. *Comput. Geom. Theory Appl.*, 15:215–227, 2000.
- [ERS05] G. Even, D. Rawitz, and S. Shahar. Hitting sets when the VC-dimension is small. *Inform. Process. Lett.*, 95(2):358–362, 2005.
- [FP11] J. Fox and J. Pach. Computing the independence number of intersection graphs. In *Proc. 22nd ACM-SIAM Sympos. Discrete Algorithms*, pages 1161–1165, 2011.
- [Fre87] G. N. Frederickson. Fast algorithms for shortest paths in planar graphs, with applications. *SIAM J. Comput.*, 16(6):1004–1022, 1987.
- [Hal98] M. M. Halldórsson. Approximations of independent sets in graphs. In *The 2nd Intl. Work. Approx. Algs. Combin. Opt. Problems*, pages 1–13, 1998.
- [Has96] J. Hastad. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Mathematica*, pages 627–636, 1996.
- [KLPS86] K. Kedem, R. Livne, J. Pach, and M. Sharir. On the union of Jordan regions and collision-free translational motion amidst polygonal obstacles. *Discrete Comput. Geom.*, 1:59–71, 1986.
- [KY07] C. Koufogiannakis and N. E. Young. Beating simplex for fractional packing and covering linear programs. In *Proc. 48th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 494–506, 2007.

- [Lon01] P. M. Long. Using the pseudo-dimension to analyze approximation algorithms for integer programming. In *Proc. 7th Workshop Algorithms Data Struct.*, volume 2125 of *Lecture Notes Comput. Sci.*, pages 26–37, 2001.
- [LT79] R. J. Lipton and R. E. Tarjan. A separator theorem for planar graphs. *SIAM J. Appl. Math.*, 36:177–189, 1979.
- [Mat92] J. Matoušek. Reporting points in halfspaces. *Comput. Geom. Theory Appl.*, 2(3):169–186, 1992.
- [MBH⁺95] M. V. Marathe, H. Breu, H. B. Hunt III, S. S. Ravi, and D. J. Rosenkrantz. Simple heuristics for unit disk graphs. *Networks*, 25:59–68, 1995.
- [MR95] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, New York, NY, 1995.
- [MR09] N. H. Mustafa and S. Ray. PTAS for geometric hitting set problems via local search. In *Proc. 25th Annu. ACM Sympos. Comput. Geom.*, pages 17–22, 2009.
- [NHK05] T. Nieberg, J. Hurink, and W. Kern. A robust PTAS for maximum weight independent set in unit disk graphs. In *Proc. 30th Int. Workshop Graph-Theoretic Concepts in Comput. Sci.*, volume 3353 of *Lecture Notes Comput. Sci.*, pages 214–221, 2005.
- [PR08] E. Pyrga and S. Ray. New existence proofs for ε -nets. In *Proc. 24th ACM Sympos. Comput. Geom.*, pages 199–207, 2008.
- [SW98] W. D. Smith and N. C. Wormald. Geometric separator theorems and applications. In *Proc. 39th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 232–243, 1998.
- [Var10] K. Varadarajan. Weighted geometric set cover via quasi-uniform sampling. In *Proc. 42nd Annu. ACM Sympos. Theory Comput.*, pages 641–648, 2010.
- [WZ90] S. Whitesides and R. Zhao. k -admissible collections of Jordan curves and offsets of circular arc figures. Technical Report SOCS 90.08, McGill Univ., Montreal, Quebec, 1990.