

# Independent and Hitting Sets of Rectangles Intersecting a Diagonal Line: Algorithms and Complexity

José Correa · Laurent Feuilloley ·  
Pablo Pérez-Lantero · José A. Soto

Received: 27 December 2013 / Revised: 26 October 2014 / Accepted: 27 December 2014 /  
Published online: 21 January 2015  
© Springer Science+Business Media New York 2015

**Abstract** Given a set of  $n$  axis-parallel rectangles in the plane, finding a maximum independent set (MIS), a maximum weighted independent set (WMIS), and a minimum hitting set (MHS), are basic problems in computational geometry and combinatorics. They have attracted significant attention since the sixties, when Wegner conjectured that the duality gap, equal to the ratio between the size of MIS and the size of MHS, is always bounded by a universal constant. An interesting case is when there exists a diagonal line that intersects each of the given rectangles. Indeed, Chepoi and Felsner recently gave a 6-approximation algorithm for MHS in this setting, and showed that the duality gap is between  $3/2$  and 6. We consider the same setting and improve upon these

---

An extended abstract of a preliminary version of this work appears in the Proceedings of the 11th Latin American Theoretical Informatics Symposium, LATIN 2014.

---

J. Correa  
Department of Industrial Engineering, Universidad de Chile,  
Santiago, Chile  
e-mail: correa@uchile.cl

L. Feuilloley  
Department of Computer Science, ENS Cachan,  
Cachan, France  
e-mail: laurent.feuilloy@ens-cachan.fr

P. Pérez-Lantero  
Escuela de Ingeniería Civil en Informática, Universidad de Valparaíso,  
Valparaíso, Chile  
e-mail: pablo.perez@uv.cl

J. A. Soto  
Department of Mathematical Engineering and CMM, Universidad de Chile,  
Santiago, Chile  
e-mail: jsoto@dim.uchile.cl

results. First, we derive an  $O(n^2)$ -time algorithm for the WMIS when, in addition, every pair of intersecting rectangles have a common point below the diagonal. This improves and extends a classic result of Lubiw, and gives a 2-approximation algorithm for WMIS. Second, we show that MIS is NP-hard. Finally, we prove that the duality gap is between 2 and 4. The upper bound, which implies a 4-approximation algorithm for MHS, follows from simple combinatorial arguments, whereas the lower bound represents the best known lower bound on the duality gap, even in the general setting of the rectangles.

**Keywords** Independent sets · Rectangle selection · Duality gap

## 1 Introduction

Given a set of  $n$  weighted axis-parallel rectangles in the plane, a subset of rectangles is called *independent* if the rectangles in the subset are pairwise disjoint. A finite point set is called a *hitting set* if the points stab all the rectangles. In this paper, we are concerned with three natural problems: finding a *maximum independent set of rectangles* (MIS); finding a *maximum weight independent set of rectangles* (WMIS); and finding a *minimum hitting set* (MHS). As these problems appear to be basic questions with important applications, they have attracted significant attention from combinatorial and computational viewpoints.

All MIS, WMIS, and MHS have long been known to be strongly NP-hard [12, 18]. Consequently, efforts have been put in designing approximation algorithms and polynomial-time exact algorithms for special cases. The current best known approximation factors are  $O(\log \log n)$  for MIS [7] and  $O(\log n / \log \log n)$  for WMIS [8]. Recently, Adamaszek and Wiese [1] designed a pseudo-polynomial time algorithm that finds a  $(1 + \varepsilon)$ -approximate solution for WMIS, however, the existence of polynomial-time approximation algorithms with constant performance guarantee remains largely open. A similar situation occurs for MHS where the current champion is an algorithm with an approximation factor of  $O(\log \log n)$  [3], while polynomial-time constant-factor approximation algorithms are unknown. In more restricted settings, polynomial-time exact algorithms do exist. For instance, if all rectangles are intervals, the underlying intersection graph<sup>1</sup> is an interval graph and even linear-time algorithms, assuming that the input is sorted beforehand, are known for all three problems [13]. Beyond interval graphs, Lubiw [22] devised a cubic-time algorithm for computing a maximum weight independent set of point-intervals, which can be seen as a set of rectangles having their upper-right corners along the same diagonal. More recently, Soto and Telha [24] considered the case where, given two prescribed points sets  $A$  and  $B$  on the plane with total size  $m$ , the set of rectangles consists of all the rectangles whose upper-right corner and lower-left corner belong to  $A$  and  $B$ , respectively. They designed an algorithm that computes both MIS and MHS in the time required to do matrix multiplication on  $m \times m$  matrices, and showed that WMIS is NP-hard. On the

<sup>1</sup> Given a finite set of rectangles in the plane, the *intersection graph* is the graph with the rectangles as vertices, and two rectangles are adjacent if and only if they intersect.

other hand, Caraballo et al. [5] considered the same setting with prescribed points sets  $A$  and  $B$ , although the set of rectangles consists of all the rectangles having a diagonal connecting a point in  $A$  with a point in  $B$ . They proved that MIS is NP-hard and admits a polynomial-time 4-approximation. Very recently, Ahmadinejad and Zarrabi-Zadeh [2] considered the case where all rectangles are attached to the boundary of a rectangular region and give an  $O(n^4)$  time exact algorithm for WMIS, improving on a result of Kong et al. [15]. Finally, there are also known polynomial-time approximation schemes for special cases, including the results of Chan and Har-Peled [8] for pseudo-disks, and Mustafa and Ray [23] for unit-height rectangles.

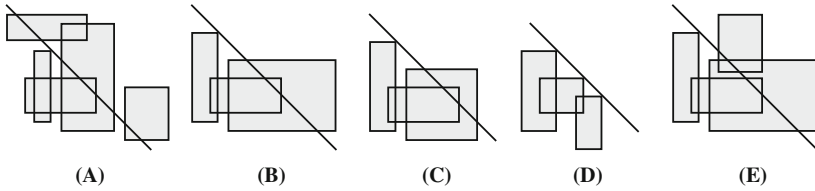
From a more combinatorial perspective, research has mostly focused on understanding the relation between the size of a maximum independent set and that of a minimum hitting set. Obviously, the former one is at most the latter one since a different point will be needed to hit each rectangle in an independent set. Furthermore, if all the rectangles have height zero (i.e. the intersection graph is an interval graphs), this inequality is actually an equality and this still holds in the case studied by Soto and Telha [24]. In general, the *duality gap*, denoted  $\delta_{\text{GAP}}$ , is the maximum ratio between these quantities, and the term arises as MHS is the integral version of the dual of the natural linear programming relaxation of MIS. Then,  $\delta_{\text{GAP}} = 1$  in interval graphs and in the framework of Soto and Telha. A natural question is whether  $\delta_{\text{GAP}}$  is bounded for general sets of rectangles. Indeed, already in the sixties, Wegner [26] conjectured  $\delta_{\text{GAP}} \leq 2$  for arbitrary rectangle sets, whereas Gyárfás and Lehel [14] proposed the weaker conjecture that  $\delta_{\text{GAP}}$  is bounded by a universal constant. The largest known lower bound on the duality gap is  $5/3$ , and was obtained by Fon-Der-Flass and Kostochka [11]. To date, the best known upper bound is only super-constant and was obtained by Károlyi [19] who proved that  $\delta_{\text{GAP}} = O(\log \alpha)$ , where  $\alpha$  denotes the size of the maximum independent set. For some special sets of rectangles,  $\delta_{\text{GAP}}$  is indeed a constant. In particular, when all rectangles intersect a given diagonal line, Chepoi and Felsner [9] proved  $\delta_{\text{GAP}} \leq 6$ , and obtained further improvements for more restricted classes [9, 16].

In this paper, we study the problems MIS, WMIS, and MHS for restricted sets of rectangles in which a given diagonal pierces every rectangle. Informally speaking, our main results are: a quadratic-time algorithm for WMIS when every pair of intersecting rectangles have a common point below the diagonal, improving and extending the work of Lubiw [22]; an NP-hardness proof for MIS in this class of rectangles; and new lower and upper bounds on the duality gap of 2 and 4, respectively. For a more precise description, we first need some definitions.

### 1.1 Notation and Classes of Rectangle Sets

Let  $D$  be the diagonal line defined by the equation  $y = -x$ . We call the closed halfplanes  $y \geq -x$  and  $y \leq -x$ , the *halfplanes* of  $D$ , whose intersection is  $D$ . The points in the bottom (resp. top) halfplane are said to be *below* (resp. *above*)  $D$ . Let  $\mathcal{R}$  be a set of  $n$  closed axis-parallel rectangles in  $\mathbb{R}^2$ , so that the diagonal  $D$  pierces each rectangles of  $\mathcal{R}$ .<sup>2</sup> Given  $r \in \mathcal{R}$ , let  $\ell^r$  and  $u^r$  denote the lower-left and upper-right

<sup>2</sup> We observe that the line  $D$  can be replaced by any decreasing bijective real function, so that the boundary of each  $r \in \mathcal{R}$  intersects the curve of that function in at most two points. By applying a suitable piecewise



**Fig. 1** Examples of rectangle sets: **a** diagonal-pierced,  $\mathcal{G}_{\text{pierce}}$ , **b** diagonal-side-pierced,  $\mathcal{G}_{\text{side}}$ , **c** diagonal-corner-separated,  $\mathcal{G}_{\text{c-sep}}$ , **d** diagonal-touched,  $\mathcal{G}_{\text{touch}}$ , **e** sub-diagonal-intersecting,  $\mathcal{G}_{\text{sub-int}}$

corners of  $r$ , respectively, and let  $w_r$  denote the weight of  $r$ . For a point  $v \in \mathbb{R}^2$ , let  $v_x$  and  $v_y$  denote the  $x$ - and  $y$ -coordinates of  $v$ , respectively. We assume without loss of generality that no two corners in the set  $\{\ell^r, u^r \mid r \in \mathcal{R}\}$  have the same  $x$ - or  $y$ -coordinate. Here, we consider five classes of rectangle sets intersected by  $D$ , depicted in Fig. 1.

**Definition 1** (Classes of rectangle sets and intersection graphs)

- (1)  $\mathcal{R}$  is *diagonal-pierced* if  $r \cap D \neq \emptyset$  for all  $r \in \mathcal{R}$ .
- (2)  $\mathcal{R}$  is *diagonal-side-pierced* if there is a side (upper, lower, left, or right) such that  $D$  intersects all  $r \in \mathcal{R}$  on that particular side.
- (3)  $\mathcal{R}$  is *diagonal-corner-separated* if there is a halfplane of  $D$  containing the same three corners of all rectangles of  $\mathcal{R}$ .
- (4)  $\mathcal{R}$  is *diagonal-touched* if either all the upper-right corners, or all the lower-left corners, of the rectangles in  $\mathcal{R}$  are in  $D$ .
- (5) A diagonal-pierced set  $\mathcal{R}$  is *sub-diagonal-intersecting* if every pair of intersecting rectangles have a common point in the bottom halfplane of  $D$ .

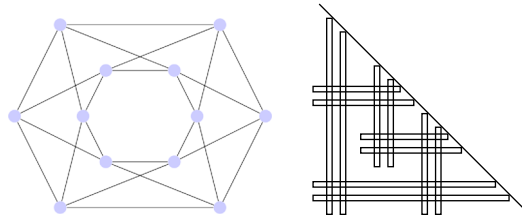
Let  $\mathcal{G}_{\text{pierce}}$  be the class (i.e. set) of the intersection graphs of all diagonal-pierced rectangle sets. Analogously, let  $\mathcal{G}_{\text{side}}$ ,  $\mathcal{G}_{\text{c-sep}}$ ,  $\mathcal{G}_{\text{touch}}$ , and  $\mathcal{G}_{\text{sub-int}}$  denote the intersection graph classes for the diagonal-side-pierced, diagonal-corner-separated, diagonal-touched, and sub-diagonal-intersecting sets of rectangles, respectively.

By rotating or reflecting the plane, we can make the following assumptions: In the diagonal-side-pierced rectangle sets, we assume that the common side of intersection is the upper one; in the diagonal-corner-separated rectangle sets, that the upper-right corner is on the top halfplane of  $D$  and the other three corners are in the bottom one; and in the diagonal-touched rectangle sets, that the corner contained in  $D$  is the upper-right one. Under these assumptions, the diagonal-pierced sets generalize the diagonal-side-pierced sets; the diagonal-side-pierced sets generalize the diagonal-corner-separated sets; and the diagonal-corner-separated sets generalize the diagonal-touched sets. We observe that these classes have appeared in the literature under different names. Hixon [16] calls the graphs in  $\mathcal{G}_{\text{touch}}$  *hook graphs*, Catanzaro et al. [6] call them *max point-tolerance graphs*, Soto and Thraves [25] call

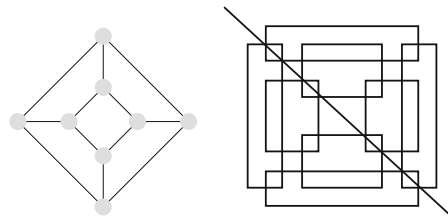
Footnote 2 continued

linear transformations on both coordinates, we can transform the rectangle set into one with the same intersection graph such that every rectangle is pierced by  $D$ .

**Fig. 2**  $G_{2C_6}$ , the doubled 6-cycle and its diagonal touching representation



**Fig. 3** The cube graph and diagonal intersecting representation



them AND(1) *graphs*, whereas Chepoi and Felsner [9] call those in  $\mathcal{G}_{\text{pierce}}$  *separable rectangle graphs*.

Although we work directly with the rectangle representation of these classes, it is worth noting that the underlying intersection graphs satisfy the following inclusions. Furthermore, given a graph, we do not know of polynomial-time algorithms to recognize the classes defined here.

$$\mathcal{G}_{\text{touch}} \subsetneq \mathcal{G}_{\text{c-sep}} = \mathcal{G}_{\text{side}} = \mathcal{G}_{\text{sub-int}} \subsetneq \mathcal{G}_{\text{pierce}}.$$

Indeed, note first that  $\mathcal{G}_{\text{c-sep}} \subseteq \mathcal{G}_{\text{side}} \subseteq \mathcal{G}_{\text{sub-int}}$ , where the first inclusion is straightforward, and for the second one we note that two rectangles of a diagonal-side-pierced set intersect if and only if they have a point in common in the bottom halfplane of  $D$ . Note also that by replacing each rectangle  $r$  of a sub-diagonal-intersecting set with the minimum rectangle containing the region of  $r$  that is below the diagonal, we get a diagonal-corner-separated set with the same intersection graph. By the previous discussion we conclude that  $\mathcal{G}_{\text{c-sep}} = \mathcal{G}_{\text{side}} = \mathcal{G}_{\text{sub-int}}$ . As mentioned earlier,  $\mathcal{G}_{\text{touch}} \subseteq \mathcal{G}_{\text{sub-int}}$  and that  $\mathcal{G}_{\text{c-sep}} \subseteq \mathcal{G}_{\text{pierce}}$  are direct consequences of the definitions of the classes.

To prove that  $\mathcal{G}_{\text{touch}} \neq \mathcal{G}_{\text{sub-int}}$ , consider the graph  $G_{2C_6}$  of Fig. 2 together with a universal vertex. This new graph still belongs to  $\mathcal{G}_{\text{sub-int}}$  since in  $\mathcal{G}_{\text{sub-int}}$  one can always create a rectangle intersecting all other rectangles and being a universal vertex of the underlying graph. However, although  $G_{2C_6}$  is in  $\mathcal{G}_{\text{touch}}$  (see the rectangle representation on the right of the figure), when a universal vertex is added the graph is not in  $\mathcal{G}_{\text{touch}}$  anymore. This can be proved by case analysis and using the following characterization of  $\mathcal{G}_{\text{touch}}$ , independently found by Hixon [16] and Soto and Thraves [25]: A graph  $G = (V, E)$  belongs to  $\mathcal{G}_{\text{touch}}$  if and only if there exists a total order  $<$  on  $V$  such that: for all  $a < b < c < d$  in  $V$ , if both  $(a, c) \in E$  and  $(b, d) \in E$  then  $(b, c) \in E$ .

Finally, to see that  $\mathcal{G}_{\text{c-sep}} \subsetneq \mathcal{G}_{\text{pierce}}$ , consider the cube depicted in Fig. 3, which is in  $\mathcal{G}_{\text{pierce}}$  but not in  $\mathcal{G}_{\text{sub-int}}$ . The first assertion follows directly from the figure on

the right, while the second one follows again by distinguishing the possible cases describing how the rectangles intersect the diagonal.

## 1.2 Our Results

In this paper, we study the problems MIS, WMIS, and MHS with input  $\mathcal{R}$  drawn for Definition 1. We thus denote the corresponding solutions as  $\text{MIS}(\mathcal{R})$ ,  $\text{WMIS}(\mathcal{R})$ , and  $\text{MHS}(\mathcal{R})$ , respectively, and let  $\text{mis}(\mathcal{R}) = |\text{MIS}(\mathcal{R})|$ ,  $\text{wmis}(\mathcal{R}) = \sum_{r \in \text{WMIS}(\mathcal{R})} w_r$ , and  $\text{mhs}(\mathcal{R}) = |\text{MHS}(\mathcal{R})|$ .

In Sect. 2, we give a quadratic-time algorithm to solve  $\text{WMIS}(\mathcal{R})$  when  $\mathcal{R}$  is sub-diagonal-intersecting, and a polynomial-time 2-approximation for  $\text{WMIS}(\mathcal{R})$  when  $\mathcal{R}$  is diagonal-pierced. The former one is the first polynomial-time algorithm for solving WMIS on a natural class of rectangle sets which includes diagonal-corner-separated rectangle sets, and improves upon previous work in the area. Specifically, for diagonal-touched rectangle sets, the best known algorithm to solve WMIS is due to Lubiw [22], who designed a cubic-time algorithm for the problem in the context of *interval systems*. More precisely, a collection of *point-intervals* is a set  $Q = \{(p_i, I_i)\}_{i=1}^n$  of  $n$  point-interval pairs, where  $I_i = [\text{left}(I_i), \text{right}(I_i)] \subseteq \mathbb{R}$  is an interval and  $p_i$  is a point in  $I_i$ , for each  $i \in [n] = \{1, \dots, n\}$ . The set  $Q$  is called *independent* if for  $k \neq j$ , we have  $p_k \notin I_j$  or  $p_j \notin I_k$ . Given a finite set  $Q$  of weighted point-intervals, Lubiw designed a dynamic programming based algorithm to find a maximum weight independent set of  $Q$ . It is easy to see [24] that this problem is equivalent to that of solving  $\text{WMIS}(\mathcal{R})$  for the diagonal-touched set  $\mathcal{R} = \{r_i\}_{i=1}^n$ , where  $r_i$  is the rectangle with upper-right corner  $(p_i, -p_i)$ , lower-left corner  $(\text{left}(I_i), -\text{right}(I_i))$ , and weight equal to that of  $(p_i, I_i)$ . Lubiw's algorithm was recently rediscovered by Hixon [16] and Catanzaro et al. [6]. Our algorithm, like Lubiw's, is based on dynamic programming. However, rather than decomposing the instance into small triangles and computing the optimal solution for the rectangles contained in every possible triangle, our approach involves computing the optimal solutions restricted to what we call a *harpoon*, which is defined for every pair of rectangles. We show that the amortized time cost of computing the optimal solution for all harpoons is constant, leading to an overall quadratic time. Interestingly, it is possible to show that our algorithm is an extension of the linear-time algorithm for maximum weighted independent set of intervals [17]. Our polynomial-time 2-approximation algorithm for  $\text{WMIS}(\mathcal{R})$ , when  $\mathcal{R}$  is diagonal-pierced, improves upon the 6-approximation of Chepoi and Felsner [9] which solves only  $\text{MIS}(\mathcal{R})$ .

In Sect. 3, we prove  $\delta_{\text{GAP}} \leq 2$  for diagonal-touched sets;  $\delta_{\text{GAP}} \leq 3$  for sub-diagonal-intersecting sets, and  $\delta_{\text{GAP}} \leq 4$  for diagonal-pierced sets. These bounds yield straightforward polynomial-time 2, 3, and 4-approximation algorithms for MHS on each class. They can also be used as approximation algorithms for MIS with the same approximation guarantee, however, as discussed in the previous paragraph, we have an exact algorithm for WMIS on the first two classes, and a 2-approximation for the last one. The 4-approximation for MHS in diagonal-pierced sets is the best approximation known, and improves upon the 6-approximation of Chepoi and Felsner [9], who also give a 3-approximation for diagonal-side-pierced sets based on a different method. For diagonal-touched sets, Hixon [16] independently showed  $\delta_{\text{GAP}} \leq 2$ . To complement

the previous results, we prove  $\delta_{\text{GAP}} \geq 2$  for diagonal-pierced sets. We do this by exhibiting an infinite set of instances whose  $\delta_{\text{GAP}}$  is arbitrarily close to 2. Similar instances were obtained, and communicated to us, by Cibulka et al. [10]. Note that this lower bound of 2 improves upon the  $5/3$  of Fon-Der-Flaass and Kostochka [11], which was the best known lower bound for the duality gap of general rectangle sets.

In Sect. 4, we prove that solving MIS (and then WMIS) on diagonal-pierced sets is NP-hard. In light of our polynomial-time algorithm for sub-diagonal-intersecting sets, the latter hardness result exhibits what is, in a way, a class at the boundary between polynomial-time solvability and NP-hardness. Three decades ago, Fowler et al. [12] (see also Asano [4]) established that solving MIS is NP-hard, by actually showing that this is the case even for squares. It is worth mentioning that diagonal-pierced sets constitute the first natural subclass for which NP-hardness of MIS has been shown since then. Our proof actually uses only rectangles that touch the diagonal line, but that may intersect above or below it, and uses a reduction from PLANAR 3-SAT.

Finally, combining the results of Chalermsook and Chuzhoy [7], and Aronov et al. [3], we show in Sect. 5 that for a general set  $\mathcal{R}$  of rectangles the duality gap is  $O(\log^2 \log \text{mis}(\mathcal{R}))$ , improving upon the  $O(\log \text{mis}(\mathcal{R}))$  bound of Károlyi [19].

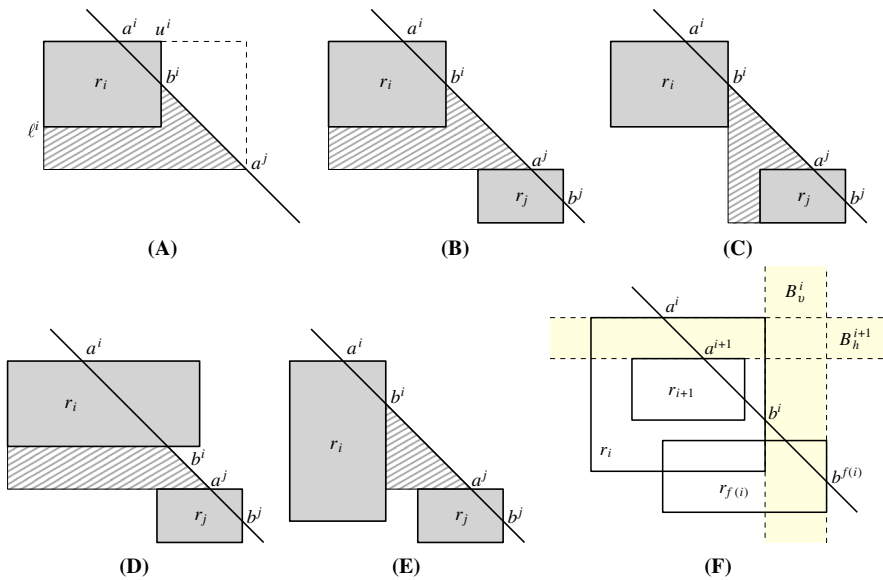
## 2 Algorithms for WMIS

Let  $r_1, r_2, \dots, r_n$  denote the  $n$  rectangles of  $\mathcal{R}$ , respectively. For each  $i \in [n]$ , let  $a^i$  and  $b^i$  denote, respectively, the higher and lower intersection points between  $D$  and the boundary of  $r_i$ . W.l.o.g., we assume  $a_x^1 < a_x^2 < \dots < a_x^n$ . For every  $i \in [n]$ , we define  $f(i)$  as follows: If there exists an index  $j \in [n]$  such that  $b_x^i < b_x^j$ , then  $f(i)$  denotes such index  $j$  with minimum  $b_x^j$ . Otherwise, we let  $f(i) = n + 1$ . For ease of notation, for every  $i \in [n]$ , let  $u^i = u^{r_i}$ ,  $\ell^i = \ell^{r_i}$ ,  $w_i = w_{r_i}$ . Let also  $\text{sub}(r_i)$  denote the intersection between  $r_i$  and the lower halfplane of  $D$ .

The idea behind Lubiw's algorithm [22] for solving WMIS on diagonal-touched rectangle sets is to compute the optimal independent set  $\text{OPT}_{ij}$  restricted to the set of rectangles  $r \in \mathcal{R}$  such that  $r$  is contained in the triangle with vertex set  $\{u^i, u^j, (u_x^i, u_y^j)\}$ , for every pair  $i < j$ . The principle exploited is that in  $\text{OPT}_{ij}$  there exists one rectangle, say  $r_k$ , with  $i < k < j$  and  $\text{OPT}_{ij} = \text{OPT}_{ik} \cup \{r_k\} \cup \text{OPT}_{kj}$ . Using this idea, Lubiw devise a dynamic programming algorithm with overall time complexity  $O(n^3)$ . Below, we present a new algorithm, which works for the more general sub-diagonal-intersecting sets and is based on more elaborate ideas, involving regions that we call *harpoons* and partitions of the set  $\mathcal{R}$  according to the incidences of the lower-left corners in horizontal and vertical *strips*.

### 2.1 Algorithm for Sub-diagonal-Intersecting Rectangle Sets

For any pair of rectangles  $r_i$  and  $r_j$  ( $i < j$ ) we define  $H_{i,j}$  and  $H_{j,i}$ , two shapes that we call *harpoons* (see Fig. 4). Formally, the *horizontal harpoon*  $H_{i,j}$  consists of the points below the diagonal  $D$  in the set obtained by subtracting the rectangle  $r_i$  from the closed box having the points  $(\ell_x^i, a_y^i)$  and  $a^j$  as vertices. Similarly, the *vertical*



**Fig. 4** **A** The construction of the harpoon  $H_{i,j}$ . **B** Harpoon  $H_{i,j}$ . **C** Harpoon  $H_{j,i}$ . **D, E** Particular degenerated cases of the harpoon  $H_{i,j}$  with  $i < j$  (the symmetric cases occur for  $H_{j,i}$ ). **F** The construction of the strips  $B_h^i$  and  $B_v^i$

harpoon  $H_{j,i}$  consists of the points below  $D$  in the set obtained by subtracting  $r_j$  from the box having the points  $(b_x^j, \ell_y^j)$  and  $b^i$  as vertices. Also, for every rectangle  $r_i$  (resp. such that  $r_{f(i)}$  is defined) we define  $B_h^i$  (resp.  $B_v^i$ ) as the open horizontal strip that goes through  $a^{i-1}$  and  $a^i$  (resp. as the open vertical strip that goes through  $b^i$  and  $b^{f(i)}$ ).

In our algorithm, called WMIS-SUBDIAGONALINTERSECT and explained below, we will compute  $S(i, j) = \text{wmis}(\mathcal{R}_{i,j})$ , where  $\mathcal{R}_{i,j} = \{r \in \mathcal{R} \mid \text{sub}(r) \subseteq H_{i,j}\}$ . We define two dummy rectangles  $r_0$  and  $r_{n+1}$ , at the two “ends” of the diagonal  $D$ , such that the harpoons  $H_{0,n+1}$  and  $H_{n+1,0}$  defined by these rectangles contain all other rectangles. As previously defined, two rectangles of  $\mathcal{R}$  intersect if and only if they intersect below the diagonal. Therefore, we have  $\text{wmis}(\mathcal{R}) = S(0, n + 1)$ .

**Algorithm WMIS-SUBDIAGONALINTERSECTING:**

1. *Initialization:* We partition the set of rectangles according to the horizontal strips containing their lower-left corner, and we do the same for the vertical strips. Formally, we compute in a preprocessing step, the sets  $\hat{B}_h^i = \{r \in \mathcal{R} \mid \text{sub}(r) \in B_h^i\}$  for  $i = 1, \dots, n + 1$ , and the sets  $\hat{B}_v^i = \{r \in \mathcal{R} \mid \ell^r \in B_v^i\}$  for  $i = 0, \dots, n$ .
2. *Main loop:* We compute the values  $S(i, j)$  by using dynamic programming, starting with  $S(i, i) = 0$  for  $i = 0, \dots, n + 1$ . Assume that for some  $t \in \{0, \dots, n\}$  all the values  $S(i, j)$  for every pair  $i, j$  such that  $|i - j| \leq t$  are computed. Then,  $S(i, j)$  with  $|i - j| = t + 1$  is computed as follows:



- If  $i < j$ , then initialize  $S(i, j) = S(i, j - 1)$  and let  $\hat{B}_{i,j} := \hat{B}_h^j$ . Otherwise, if  $i > j$ , then initialize  $S(i, j) = S(i, f(j))$  and let  $\hat{B}_{i,j} := \hat{B}_v^j$ .
  - For each rectangle  $r_k \in \hat{B}_{i,j}$  such that  $r_k \in \mathcal{R}_{i,j}$ , we set  $m = w_k + \max\{S(i, k), S(k, i)\} + S(k, j)$  and update  $S(i, j) := \max\{m, S(i, j)\}$ .
3. *Output:* Return  $S(0, n + 1)$ .

It is trivial to modify Algorithm WMIS-SUBDIAGONALINTERSECT to return not only  $wmis(\mathcal{R})$  but also the solution rectangle set  $WMIS(\mathcal{R})$ . We now establish the running time of this algorithm.

**Theorem 1** *Algorithm WMIS-SUBDIAGONALINTERSECT runs in  $O(n^2)$  time.*

*Proof* The initialization stage requires  $O(n)$  time if the rectangles are already sorted, otherwise it requires  $O(n \log n)$  time. The time to compute  $S(i, j)$  for given values of  $i$  and  $j$  is within  $O(1 + |\hat{B}_{i,j}|)$ , since checking whether a given rectangle belongs to  $\mathcal{R}_{i,j}$  or  $\mathcal{R}_{j,i}$  takes constant time. Since each rectangle is at most once in some set  $\hat{B}_h^i$ , and at most once in some set  $\hat{B}_v^i$ , the time to fill the whole table  $S(\cdot, \cdot)$  is within:

$$\sum_{(i,j) \in [n]^2} O(1 + |\hat{B}_{i,j}|) = O(n^2). \quad \square$$

In order to analyze the correctness of our algorithm, we define a partial order over the elements of  $\mathcal{R}$ .

**Definition 2** The (strict) *onion ordering*  $<$  in  $\mathcal{R}$  is defined as

$$r_i < r_j \iff r_i \cap r_j = \emptyset, \ell_x^i < \ell_x^j, \text{ and } \ell_y^i < \ell_y^j.$$

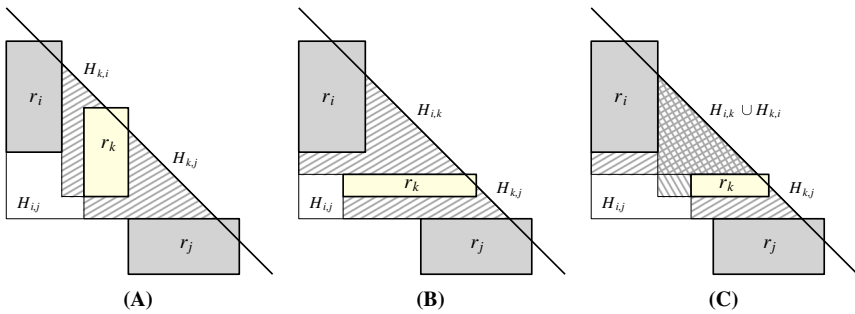
Clearly,  $<$  is a strict partial ordering in  $\mathcal{R}$ . We say that  $r_i$  is *dominated* by  $r_j$  if  $r_i < r_j$ ; in other words,  $r_i$  is dominated by  $r_j$  if  $r_i \cap r_j = \emptyset$  and  $\ell^i$  is dominated by  $\ell^j$  under the standard vector dominance in  $\mathbb{R}^2$ .

For any  $k$  such that  $r_k \in \mathcal{R}_{i,j}$ , let  $W_k(i, j)$  denote a maximum weight independent set that contains  $r_k$  and rectangles in  $\mathcal{R}_{i,j}$  not dominated by  $r_k$  in the onion ordering, and  $S_k(i, j)$  denote the total weight of  $W_k(i, j)$ .

**Lemma 1** *For any rectangle  $r_k \in \mathcal{R}_{i,j}$ , the following relation holds:*

$$S_k(i, j) = w_k + \max\{S(i, k), S(k, i)\} + S(k, j).$$

*Proof* Since  $r_k \in \mathcal{R}_{i,j}$ , we have that  $r_i, r_k$  and  $r_j$  are pairwise disjoint and  $\min\{i, j\} < k < \max\{i, j\}$ . Assume that the harpoon  $H_{i,j}$  is horizontal, i.e.,  $i < j$  (the proof for  $i > j$  is analogous). In particular, we know that  $a^i, b^i, a^k, b^k, a^j$ , and  $b^j$  appear in this order on the diagonal. There are three cases for the relative positions of the two rectangles  $r_i$  and  $r_k$  (see Fig. 5).



**Fig. 5** The three cases for a rectangle  $r_k$  contained in the set  $\mathcal{R}_{i,j}$  induced by the horizontal harpoon  $H_{i,j}$

CASE 1:  $r_i$  and  $r_k$  are separated by a vertical line, but not separated by a horizontal one (see Fig. 5a). Since  $H_{i,k} \subseteq H_{k,i}$ , we conclude that  $W_k(i, j) \setminus \{r_k\} \subseteq \mathcal{R}_{k,i} \cup \mathcal{R}_{k,j}$ . Since  $\mathcal{R}_{k,i} \cap \mathcal{R}_{k,j} = \emptyset$ , the correctness follows.

CASE 2:  $r_i$  and  $r_k$  are separated by a horizontal line, but not by a vertical one (see Fig. 5b). The proof follows almost exactly as in CASE 1.

CASE 3:  $r_i$  and  $r_k$  are separated by both a horizontal line and a vertical line (see Fig. 5c). By geometric and minimality arguments, we have  $W_k(i, j) \setminus \{r_k\} \subseteq \mathcal{R}_{i,k} \cup \mathcal{R}_{k,i} \cup \mathcal{R}_{k,j}$ . Finally, note that every independent set contained in the union of  $\mathcal{R}_{i,k}$  and  $\mathcal{R}_{k,i}$ , must be fully contained in one of them. Since these sets are disjoint from  $\mathcal{R}_{k,j}$ , the result holds. □

**Theorem 2** Algorithm WMIS-SUBDIAGONALINTERSECT returns WMIS( $\mathcal{R}$ ).

*Proof* We use induction. For the trivial sets  $\mathcal{R}_{i,i}$ , the maximum weight independent set has weight 0, since it is the empty set. The correctness follows directly from Lemma 1 and the next implications: For  $i \neq j$ ,

$$\begin{aligned}
 i < j &\implies S(i, j) = \max\{S(i, j - 1), \max\{S_k(i, j) \mid r_k \in \hat{B}_h^j \cap \mathcal{R}_{i,j}\}\}, \\
 i > j &\implies S(i, j) = \max\{S(i, f(j)), \max\{S_k(i, j) \mid r_k \in \hat{B}_v^j \cap \mathcal{R}_{i,j}\}\}.
 \end{aligned}$$

Indeed, assume that  $i < j$  (the case where  $i > j$  is analogous). Let  $\mathcal{S}$  be the WMIS corresponding to  $S(i, j)$ . Suppose there exists  $r_m \in \hat{B}_h^j \cap \mathcal{S}$  and assume w.l.o.g. that  $r_m$  is minimal in  $\mathcal{S}$  with respect to domination. Since  $\mathcal{S} \setminus \{r_m\}$  does not contain rectangles dominated by  $r_m$ ,  $S(i, j) = S_m(i, j)$ . Otherwise,  $\hat{B}_h^j \cap \mathcal{S} = \emptyset$ , in which case it follows that  $S(i, j) = S(i, j - 1)$ . □

### 2.2 A 2-Approximation for WMIS on Diagonal-Pierced Sets

We use the previous algorithm to obtain a 2-approximation for diagonal-pierced rectangle sets.

**Theorem 3** There exists a polynomial-time 2-approximation algorithm for WMIS on diagonal-pierced rectangle sets.

*Proof* Let  $\mathcal{R}_{\text{upper}} \subseteq \mathcal{R}$  denote the set of rectangles whose upper side is pierced by  $D$ , and let  $\mathcal{R}_{\text{left}} = \mathcal{R} \setminus \mathcal{R}_{\text{upper}}$ . Observe that  $D$  pierces the left side of each rectangle in  $\mathcal{R}_{\text{left}}$ . More importantly,  $\mathcal{R}_{\text{upper}}$  is a sub-diagonal-intersecting set of rectangles, and  $\mathcal{R}_{\text{left}}$  is in fact another sub-diagonal-intersecting set of rectangles in the instance obtained by reflecting  $\mathcal{R}$  with respect to  $D$ . Using Theorems 1 and 2, each of  $\text{WMIS}(\mathcal{R}_{\text{upper}})$  and  $\text{WMIS}(\mathcal{R}_{\text{left}})$  can be computed in  $O(n^2)$  time. We return the set of maximum total weight between  $\text{WMIS}(\mathcal{R}_{\text{upper}})$  and  $\text{WMIS}(\mathcal{R}_{\text{left}})$ , which is a 2-approximation of  $\text{WMIS}(\mathcal{R})$ .  $\square$

### 3 Duality Gap and Other Approximation Algorithms

In this section, we study the duality gap, that is, the largest possible ratio between the size of the minimum hitting set and the size of the maximum independent set, on some of the classes of rectangle sets defined before.

**Theorem 4** *The duality gap for diagonal-touched rectangle sets is between 3/2 and 2. For sub-diagonal-intersecting rectangle sets, it is between 3/2 and 3, and for diagonal-pierced sets it is between 2 and 4.*

We will prove the upper bounds and the lower bounds separately.

*Proof of the Upper Bounds in Theorem 4* Let  $\mathcal{R}$  be a rectangle set in the plane, that can be in one of the three classes described in the theorem. In the case in which  $\mathcal{R}$  is sub-diagonal-intersecting we first replace each rectangle  $r \in \mathcal{R}$  by the minimal one containing the region of  $r$  that is below the diagonal. The modified set has the same intersection graph as before, but it is diagonal-corner-separated. In particular, the region of each rectangle that is above the diagonal is a triangle or a single point.

Let  $\mathcal{R}_x$  and  $\mathcal{R}_y$  denote the projections of the rectangles in  $\mathcal{R}$  on the  $x$ - and  $y$ -axis, respectively, which can be regarded as sets of intervals. Let  $\mathcal{I}_x$  and  $\mathcal{I}_y$  denote maximum independent sets of  $\mathcal{R}_x$  and  $\mathcal{R}_y$ , respectively, and  $P_x$  and  $P_y$  denote minimum hitting sets of  $\mathcal{R}_x$  and  $\mathcal{R}_y$ , respectively. Since interval graphs are perfect, we have  $|P_x| = |\mathcal{I}_x|$  and  $|P_y| = |\mathcal{I}_y|$ . Further, since rectangles with disjoint projections on the  $x$ -axis (resp. on the  $y$ -axis) are disjoint, we have

$$\text{mis}(\mathcal{R}) \geq \max\{|\mathcal{I}_x|, |\mathcal{I}_y|\} = \max\{|P_x|, |P_y|\}.$$

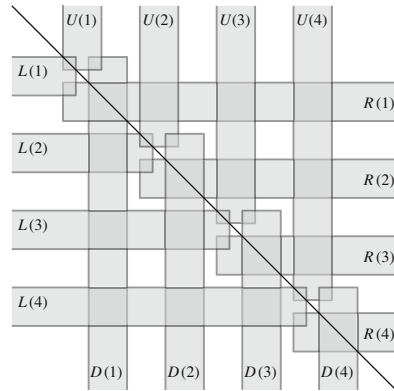
Let  $\mathcal{P} = P_x \times P_y \subset \mathbb{R}^2$ , which is a hitting set of  $\mathcal{R}$ , and let the set  $\mathcal{P}^-$  (resp.  $\mathcal{P}^+$ ) denote the points in  $\mathcal{P}$  that are below (resp. above) the diagonal  $D$ . Consider the following subsets of  $\mathcal{P}$ :

$$\begin{aligned} \mathcal{F}^- &= \{p \in \mathcal{P}^- : \nexists q \in \mathcal{P}^-, p_x < q_x \text{ and } p_y < q_y\}, \\ \mathcal{F}^+ &= \{p \in \mathcal{P}^+ : \nexists q \in \mathcal{P}^+, q_x < p_x \text{ and } q_y < p_y\}, \\ \mathcal{F}^* &= \{p \in \mathcal{P}^+ : \nexists q \in \mathcal{P}^+ \setminus \{p\}, q_x \leq p_x \text{ and } q_y \leq p_y\}. \end{aligned}$$

The set  $\mathcal{F}^-$  (resp.  $\mathcal{F}^+$ ) forms the closest “staircase” to the diagonal  $D$  that is below (resp. above)  $D$ . The set  $\mathcal{F}^*$  corresponds to the lower-left bending points of the



**Fig. 7** The set  $\mathcal{R}_4$ . The diagonal line shows that this set is diagonal-pierced



as a diagonal-touched set, that  $\text{mis}(\mathcal{R}) = 2$  and  $\text{mhs}(\mathcal{R}) = 3$ , and so the claim holds. Of course this lower bound applies also to sub-diagonal-intersecting sets.

The lower bound of 2 for diagonal-pierced sets is asymptotically attained by a sequence of rectangle sets  $\{\mathcal{R}_k\}_{k \in \mathbb{Z}^+}$ . We will describe the sequence in terms of unbounded 3-sided rectangles which intersect the diagonal, but it is easy to transform each  $\mathcal{R}_k$  into a set of bounded rectangles by confining them to a big enough bounding box.

For  $i \in \mathbb{Z}^+$ , define the  $i$ th layer as  $\mathcal{L}_i = \{U(i), D(i), L(i), R(i)\}$ , and for  $k \in \mathbb{Z}^+$ , define the  $k$ th instance as  $\mathcal{R}_k = \bigcup_{i=1}^k \mathcal{L}_i$ , where:

$$\begin{aligned}
 U(i) &= [2i, 2i + 1] \times [-(2i + \frac{1}{3}), +\infty), \\
 D(i) &= [2i + \frac{2}{3}, 2i + \frac{5}{3}] \times (-\infty, -2i], \\
 L(i) &= (-\infty, 2i + \frac{1}{3}] \times [-2i - 1, -2i], \\
 R(i) &= [2i, \infty) \times [-(2i + \frac{5}{3}), -(2i + \frac{2}{3})].
 \end{aligned}$$

Consider the instance  $\mathcal{R}_k$  (see Fig. 7 for  $\mathcal{R}_4$ ) with  $k$  layers of rectangles. It is immediately clear that  $\text{mhs}(\mathcal{R}) \geq 2k$ , since every point of the plane hits at most two rectangles.

Let us prove that  $I = \text{MIS}(\mathcal{R}_k)$  satisfies  $|I| \leq k + 2$ , which implies that  $\delta_{\text{GAP}} = \text{mhs}(\mathcal{R}_k)/\text{mis}(\mathcal{R}_k) \geq 2 - 4/(k + 2)$  is arbitrarily close to 2. To this end, let  $i_D = \min\{i : D(i) \in I \text{ or } i = k + 1\}$  and  $i_R = \min\{i : R(i) \in I \text{ or } i = k + 1\}$ . When  $i_D = i_R = k + 1$ , it is immediate that  $|I| \leq k$ . Then, assume w.l.o.g. that  $i_D < i_R$ . Since for  $i = 1, \dots, i_D - 1$  the set  $I$  contains neither  $D(i)$  nor  $R(i)$ , we have that  $I$  contains at most one rectangle on each of the layers  $\mathcal{L}_1, \dots, \mathcal{L}_{i_D-1}$ . It then follows that

$$\sum_{i=1}^{i_D-1} |I \cap \mathcal{L}_i| \leq i_D - 1.$$

Similarly, for  $i = i_D + 1, \dots, i_R - 1$ ,  $I$  contains neither  $L(i)$  nor  $R(i)$ , thus

$$\sum_{i=i_D+1}^{i_R-1} |I \cap \mathcal{L}_i| \leq i_R - i_D - 1.$$

Finally, we have for  $i = i_R + 1, \dots, k$  that  $I$  contains neither  $L(i)$  nor  $U(i)$ , and of the layer  $\mathcal{L}_{i_R}$ ,  $I$  contains at most two rectangles; thus

$$\sum_{i=i_R}^k |I \cap \mathcal{L}_i| \leq k - i_R + 2.$$

To conclude, note that  $I$  may contain at most two rectangles of the layer  $\mathcal{L}_{i_D}$ . Hence, we have

$$|I| = \sum_{i=1}^k |I \cap \mathcal{L}_i| \leq (i_D - 1) + 2 + (i_R - i_D - 1) + (k - i_R + 2) = k + 2. \quad \square$$

**Corollary 1** *There is a simple polynomial-time 2-approximation algorithm for MHS on diagonal-touched sets, a 3-approximation for MHS on sub-diagonal-intersecting sets, and a 4-approximation polynomial time algorithm for MHS on diagonal-pierced sets.*

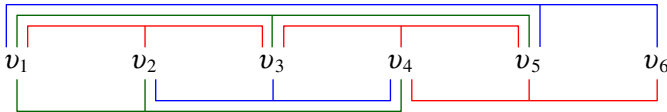
*Proof* The algorithm consists in computing and returning  $\mathcal{F}^-$  for the first case,  $\mathcal{F}^- \cup \mathcal{F}^*$  for the second case, and  $\mathcal{F}^- \cup \mathcal{F}^+$  for the third case.  $\square$

#### 4 NP-Hardness of MIS on Diagonal-Pierced Sets

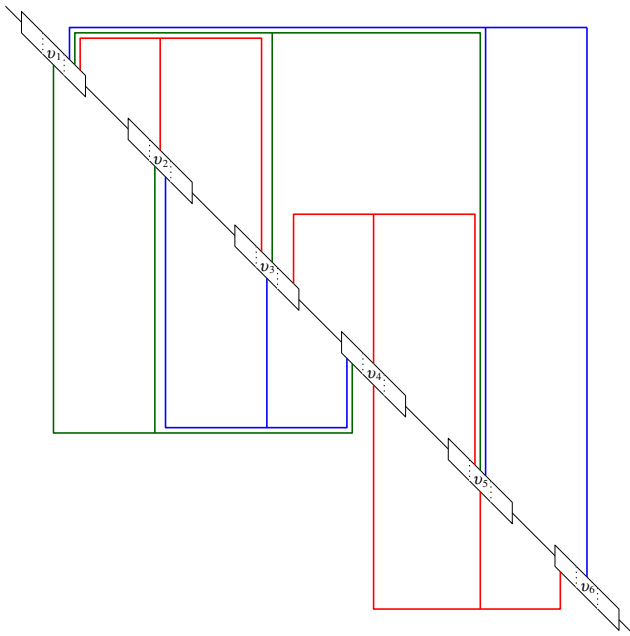
In this section, we prove the following theorem. It is worth noting that the class of rectangles referred to in the theorem is not the class of diagonal-touched rectangles: the diagonal may touch some rectangles on their lower-left corners and others on their upper-right corners.

**Theorem 5** *Solving MIS is NP-hard on diagonal-pierced sets of rectangles, even if the diagonal intersects each rectangle on a corner.*

*Proof* We use a reduction from the PLANAR 3-SAT which is NP-complete [21]. Given a Boolean formula  $\varphi$ , its associated graph  $G(\varphi)$  is the bipartite graph having one vertex per variable and one vertex per clause, and in which a variable vertex is connected to a clause vertex if either the variable or its negation appears as a literal in the clause. The input of the PLANAR 3-SAT consists of a Boolean formula  $\varphi$  in 3-CNF whose associated graph  $G(\varphi)$  is planar, and the formula is accepted if and only if there exists an assignment to its variables such that in each clause at least one literal is satisfied. Let  $\varphi$  be a PLANAR 3-SAT formula. Note that  $G(\varphi)$  can be represented in the plane as in Fig. 8, where all variables lie on an horizontal line, and all clauses



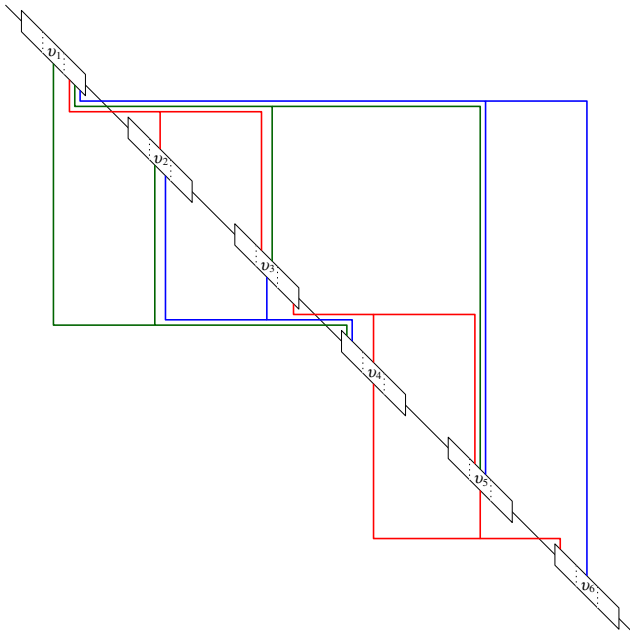
**Fig. 8** Planar embedding of  $\varphi = (v_1 \vee \bar{v}_2 \vee v_3) \wedge (v_3 \vee \bar{v}_4 \vee \bar{v}_5) \wedge (\bar{v}_1 \vee \bar{v}_3 \vee v_5) \wedge (v_1 \vee \bar{v}_2 \vee v_4) \wedge (\bar{v}_2 \vee \bar{v}_3 \vee \bar{v}_4) \wedge (\bar{v}_4 \vee v_5 \vee \bar{v}_6) \wedge (\bar{v}_1 \vee v_5 \vee v_6)$



**Fig. 9** The embedding  $E_1$  obtained (essentially) from  $E_0$  by arranging the variables as segments in the diagonal  $D$  and extending the combs. For clarity, each variable segment is represented by a *parallelogram*

are represented by *non-intersecting* three-legged combs [20]. We identify each clause with its corresponding comb, and vice versa. Using this embedding as base, which can be constructed in a grid of polynomial size [20], we construct a set  $\mathcal{R}$  of rectangles intersecting the diagonal  $D$ , such that there exists in  $\mathcal{R}$  an independent set of some given number of rectangles if and only if  $\varphi$  is accepted.

Let  $\varphi$  be an instance of the PLANAR 3-SAT, with  $n$  variables and  $m$  clauses, and let  $E_0$  denote the above embedding of  $\varphi$ . For any variable  $v$ , let  $d(v)$  denote the number of clauses in which  $v$  appears. We assume that every variable appears in each clause at most once. Given any clause  $C$  with variables  $u, v$ , and  $w$ , such that  $u, v$ , and  $w$  appear in this order from left to right in the embedding  $E_0$ , we say that  $u$  is the *left* variable of  $C$ , and that  $w$  is the *right* variable. Given  $E_0$ , using simple geometric transformations we can obtain the next slightly different planar embedding  $E_1$  of  $\varphi$ . Essentially,  $E_1$  can be obtained from  $E_0$  by arranging the variables in the diagonal  $D$  and extending the combs. Such an embedding  $E_1$  has the next further properties (see Fig. 9):



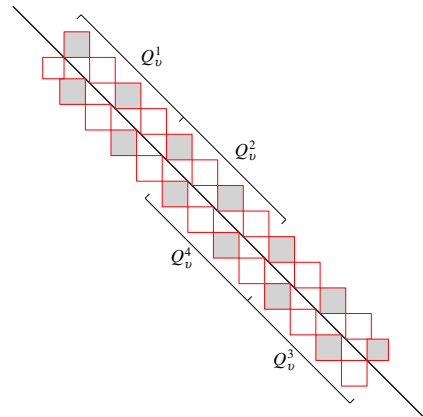
**Fig. 10** The embedding  $E_2$  obtained from  $E_1$  by modifying the combs and inverting one leg in each comb

- (1) Each variable  $v$  is represented by a segment  $S_v \subset D$ , divided into three equal parts:  $S_v^l$  is the left part,  $S_v^m$  the middle one, and  $S_v^r$  the right one. The segments  $S = \{S_v \mid v \text{ is a variable}\}$ , are pairwise disjoint and equally spaced in  $D$ , and appear in  $D$  from left to right in the same order as the variables appear in  $E_0$ . Let  $\delta$  denote the vertical gap between successive segments in  $S$ .
- (2) If  $v$  is the left variable of a clause  $C$  above  $D$ , then the left leg of  $C$  (i.e. the one corresponding to  $v$ ) contacts the interior of  $S_v^r$ . Otherwise, it contacts the interior of  $S_v^m$ .
- (3) If  $v$  is the right variable of a clause  $C$  below  $D$ , then the right leg of  $C$  (i.e. the one corresponding to  $v$ ) contacts the interior of  $S_v^l$ . Otherwise, it contacts the interior of  $S_v^m$ .

The above properties (1)–(3) of  $E_1$  allow us to obtain the next embedding  $E_2$  of  $\varphi$  (refer to Fig. 10). Let  $v$  be any variable and let  $C_1, C_2, \dots, C_k$  be the clauses above the diagonal having  $v$  as left variable, sorted according to the left-to-right order of the contact points of their left legs with  $S_v^r$ . Let  $s_1, s_2, \dots, s_k$  denote the horizontal segments of  $C_1, C_2, \dots, C_k$ , respectively. Assume w.l.o.g. that  $s_1, s_2, \dots, s_k$  are equally spaced at a distance less than  $\delta/2k$ . Push downwards simultaneously  $s_1, s_2, \dots, s_k$  to modify  $C_1, C_2, \dots, C_k$  so that  $s_1$  is now below  $S_v$ , the vertical gap between  $s_k$  and  $S_v$  is less than  $\delta/2$ , and the left legs of  $C_1, C_2, \dots, C_k$  are inverted and make contact with  $S_v^r$  from below. Further, by inverting the left-to-right order of the contact points of  $C_1, C_2, \dots, C_k$  with  $S_v^r$ , the clauses  $C_1, C_2, \dots, C_k$  become pairwise disjoint. Proceed similarly (and symmetrically) with the clauses below the diagonal having  $v$  as



**Fig. 11** The variable gadget for the variable  $v$ . The even numbered squares are shaded



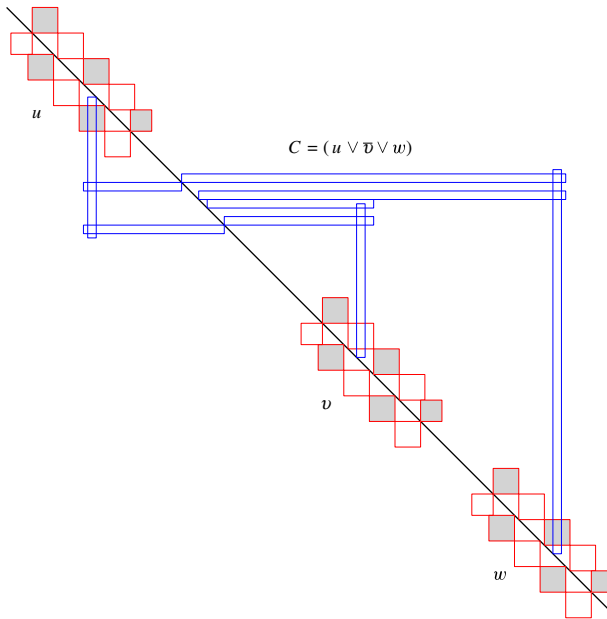
right variable. It can be verified that this new embedding  $E_2$  has no crossings among the combs and variable segments.

Using the embedding  $E_2$ , we construct a set  $\mathcal{R}$  of rectangles via variable gadgets and clause gadgets.

*Variable gadgets* For each variable  $v$ , the segment  $S_v$  is replaced by a necklace  $Q_v$  of  $12 \cdot d(v) + 2$  squares (their intersection graph is a cycle), so that each square intersects the diagonal  $D$  on a corner and only consecutive squares pairwise intersect (see Fig. 11). We number these squares consecutively in clockwise order, starting from the topmost one which is numbered 0. Let  $Q_v^1 \subset Q_v$  be the first top-down  $2 \cdot d(v)$  squares above  $D$ ,  $Q_v^2 \subset Q_v$  the second top-down  $2 \cdot d(v)$  squares above  $D$ ,  $Q_v^3 \subset Q_v$  the first bottom-up  $2 \cdot d(v)$  squares below  $D$ , and  $Q_v^4 \subset Q_v$  the second bottom-up  $2 \cdot d(v)$  squares below  $D$ . Since any clause can contact:  $S_v^\ell$  from above,  $S_v^r$  from below, and  $S_v^m$  from either above or below, we identify  $S_v^\ell$  with  $Q_v^1$ ,  $S_v^m$  with  $Q_v^2 \cup Q_v^4$ , and  $S_v^r$  with  $Q_v^3$ .

*Clause gadgets* Let  $C$  be a clause with variables  $u, v$ , and  $w$ , appearing in this order from left to right in  $E_0$ . We represent  $C$  by the set  $Q_C$  of nine thin rectangles, three vertical and six horizontal, as in Fig. 12. The vertical rectangles of  $Q_C$  represent the three legs of  $C$ , and for  $z = u, v, w$  the vertical rectangle corresponding to  $z$  intersects a unique rectangle  $R_z$  of  $Q_v$ , and  $D$  as well, so that  $R_z$  is even numbered if and only if  $z$  appears as positive in  $C$ . Furthermore, if  $C$  is above  $D$  in  $E_1$  then  $R_u \in Q_u^3$ ,  $R_v \in Q_v^2$ , and  $R_w \in Q_w^2$ . Otherwise, if  $C$  is below  $D$  in  $E_1$  then  $R_u \in Q_u^4$ ,  $R_v \in Q_v^4$ , and  $R_w \in Q_w^1$ . Observe that since for every variable  $z$  each of the sets  $Q_z^1, Q_z^2, Q_z^3, Q_z^4$  contains  $2 \cdot d(z)$  squares, we can guarantee that each square of the variable gadget is intersected by at most one vertical rectangle of the clause gadgets.

*Reduction* Observe that in each variable  $v$ , the set  $Q_v$  has exactly two maximum independent sets of rectangles of size  $6 \cdot d(v) + 1$ : the set  $Q_{v,0} \subset Q_v$  of the even-numbered squares and the set  $Q_{v,1} \subset Q_v$  of the odd-numbered squares. We consider that  $v = 1$  (i.e.  $v$  is true) if we select  $Q_{v,1}$  as a maximum independent set of rectangles of  $Q_v$ , and consider  $v = 0$  (i.e.  $v$  is false) if  $Q_{v,0}$  is selected.



**Fig. 12** Clause gadget for  $C = (u \vee \bar{v} \vee w)$ . The variable  $u$  is positive in  $C$  and the vertical rectangle of  $C$  corresponding to  $u$  intersects  $D$  and an even numbered square of  $Q_u^3$ . The variable  $v$  appears negative and the vertical rectangle of  $C$  corresponding to  $v$  intersects  $D$  and an odd numbered square of  $Q_v^4$ . The variable  $w$  appears positive and the vertical rectangle of  $C$  corresponding to  $w$  intersects  $D$  and an even numbered square of  $Q_w^3$

Observe that the next statements are satisfied:

- (1) if  $v = 1$  then the vertical rectangles of the clause gadgets in which  $v$  appears as positive, together with  $Q_{v,1}$ , form an independent set of rectangles.
- (2) if  $v = 0$  then the vertical rectangles of the clause gadgets in which  $v$  appears as negative, together with  $Q_{v,0}$ , form an independent set of rectangles.
- (3) For each clause  $C$ , any maximum independent set of rectangles of  $Q_C$  has size 4, and among its elements there must be a vertical rectangle.
- (4) For each clause  $C$  and each variable  $v$  in  $C$ : there exists an independent set of size  $(6 \cdot d(v) + 1) + 4$  in  $Q_v \cup Q_C$ , such that the vertical rectangle of  $C$  corresponding to  $v$  is selected, if and only if either  $v$  appears as positive in  $C$  and  $Q_{v,1}$  is selected or  $v$  appears as negative in  $C$  and  $Q_{v,0}$  is selected.

Let  $\mathcal{R}$  be the set of the rectangles of all variable gadgets and clause gadgets. From the above observations, we claim that  $\varphi$  can be accepted if and only if  $\mathcal{R}$  has an independent set of exactly  $\sum_v (6 \cdot d(v) + 1) + 4m$  rectangles.

Therefore, MIS is NP-hard on diagonal-pierced sets of rectangles, even if the diagonal intersects each rectangle on a corner. □

## 5 Duality Gap of General Rectangle Sets

In this section, we prove that the duality gap of general rectangle sets  $\mathcal{R}$  is  $O(\log^2 \log \text{mis}(\mathcal{R}))$ , improving on the best known bound of  $O(\log \text{mis}(\mathcal{R}))$  [19]. Surprisingly, this new bound follows by combining known ideas [3, 7].

**Theorem 6** *For every rectangle set  $\mathcal{R}$  we have  $\frac{\text{mhs}(\mathcal{R})}{\text{mis}(\mathcal{R})} \leq O(\log^2 \log \text{mis}(\mathcal{R}))$ .*

In order to prove this bound, we need to briefly recall the natural linear programming formulations for MIS and MHS. A point  $p$  in the plane is called a *witness* of an inclusion maximal clique of the intersection graph of  $\mathcal{R}$  if it hits every rectangle of this clique. Given any (possibly infinite) hitting set  $H \subseteq \mathbb{R}^2$  containing the witness points of all maximal cliques in  $\mathcal{R}$ , define the following polytopes:

$$\text{Pol}_H(\mathcal{R}) = \left\{ x \in \mathbb{R}^{\mathcal{R}} : \sum_{r \in \mathcal{R}: p \in r} x_r \leq 1 \text{ for all } p \in H, x \geq 0 \right\},$$

$$\text{Dual}_H(\mathcal{R}) = \left\{ y \in \mathbb{R}^H : \sum_{p \in H \cap r} y_p \geq 1 \text{ for all } r \in \mathcal{R}, y \geq 0 \right\}.$$

and the following dual linear programs:

$$\text{LP}_H(\mathcal{R}) = \max \left\{ \sum_{r \in \mathcal{R}} x_r : x \in \text{Pol}_H(\mathcal{R}) \right\},$$

$$\text{LP}'_H(\mathcal{R}) = \min \left\{ \sum_{p \in H} y_p : y \in \text{Dual}_H(\mathcal{R}) \right\}.$$

It is easy to see that the previous linear programs are relaxations of MIS and MHS respectively. Therefore,

$$\text{mis}(\mathcal{R}) \leq \text{LP}_H(\mathcal{R}) = \text{LP}'_H(\mathcal{R}) \leq \text{mhs}(\mathcal{R}). \quad (1)$$

Furthermore, the value  $\text{LP}_H(\mathcal{R})$  does not depend on the hitting set  $H$  chosen, and so, we can drop the subindex  $H$  in (1).

Chalermsook and Chuzhoy [7] showed that for any set of rectangles  $\mathcal{R}$  having their corners in the grid  $[t]^2 = [t] \times [t]$ , it is possible to find an independent set  $\mathcal{Q}$  with

$$\text{mis}(\mathcal{R}) \leq \text{LP}_{[t]^2}(\mathcal{R}) \leq |\mathcal{Q}| \cdot O(\log \log t). \quad (2)$$

On the other hand, Aronov et al. [3] proved that for every set  $\mathcal{R}$  of rectangles, there exist a polynomial-time computable hitting set  $P$  with

$$|P| \leq O\left(\text{LP}(\mathcal{R}) \cdot \log \log \text{LP}(\mathcal{R})\right). \quad (3)$$

To prove Theorem 6, we require the following lemma.

**Lemma 2** *For every set of rectangles  $\mathcal{R}$ , with  $\alpha = \text{mis}(\mathcal{R})$ , there is another set  $\mathcal{R}'$  of rectangles with corners in the grid  $[\alpha]^2$  such that*

$$\text{mis}(\mathcal{R}') \leq \text{mis}(\mathcal{R}) \leq \text{LP}(\mathcal{R}) \leq 9 \cdot \text{LP}(\mathcal{R}'). \tag{4}$$

*Proof* Let  $\mathcal{R}_x$  (resp.  $\mathcal{R}_y$ ) be the set of intervals obtained by projecting  $\mathcal{R}$  on the  $x$ -axis (resp.  $y$ -axis), and let  $P_x$  (resp.  $P_y$ ) be minimum hitting sets for  $\mathcal{R}_x$  (resp.  $\mathcal{R}_y$ ). Similar to the proof of the upper bounds of Theorem 4, we have

$$\max\{|P_x|, |P_y|\} = \max\{\text{mis}(\mathcal{R}_x), \text{mis}(\mathcal{R}_y)\} \leq \text{mis}(\mathcal{R}) = \alpha.$$

Consider the grid  $P_x \times P_y$  of size at most  $\alpha \times \alpha$ . By translating and piece-wise scaling the plane, we can identify  $P_x$  with the set  $\{(i, 0) : 1 \leq i \leq |P_x|\}$  and  $P_y$  with the set  $\{(0, j) : 1 \leq j \leq |P_y|\}$  without changing the intersection graph associated with  $\mathcal{R}$ . Thus, we can identify the grid  $P_x \times P_y$  with a subgrid of  $[\alpha] \times [\alpha]$ . Note that this grid is a hitting set of  $\mathcal{R}$ .

Further, consider the set  $\tilde{\mathcal{R}} = \{R \cap [1, \alpha] \times [1, \alpha] : R \in \mathcal{R}\}$ . That is,  $\tilde{\mathcal{R}}$  is obtained by trimming the rectangles to the rectangular region  $[1, \alpha] \times [1, \alpha]$ . It is easy to see that this operation does not change the intersection graph of the set either. So, for our purposes, we will assume w.l.o.g. that  $\mathcal{R} = \tilde{\mathcal{R}}$ .

Let  $\mathcal{R}'$  be the set of rectangles obtained by replacing each rectangle  $r$  of  $\mathcal{R}$  by the minimum possible rectangle in the plane containing  $r$  and having all its corners in the grid  $[\alpha] \times [\alpha]$ . That is, we replace the rectangle  $r$  defined by  $\ell^r$  and  $u^r$  by the rectangle  $r_+$  defined by  $\tilde{\ell}^r = (\lfloor \ell_x^r \rfloor, \lfloor \ell_y^r \rfloor)$  and  $\tilde{u}^r = (\lceil u_x^r \rceil, \lceil u_y^r \rceil)$ , where  $\lfloor \cdot \rfloor$  and  $\lceil \cdot \rceil$  are the floor and ceiling functions, respectively.

The first inequality of (4) follows since any independent set of  $\mathcal{R}'$  induces an independent set of  $\mathcal{R}$  of the same size. The second inequality follows from (1). The only non-trivial inequality is the last one.

Since  $[\alpha]^2$  is a set hitting all maximal cliques of  $\mathcal{R}$  and  $\mathcal{R}'$ , we have  $\text{LP}_{[\alpha]^2}(\mathcal{R}) = \text{LP}(\mathcal{R})$  and  $\text{LP}_{[\alpha]^2}(\mathcal{R}') = \text{LP}(\mathcal{R}')$ . Consider a fractional optimal solution  $y'$  for  $\text{LP}'_{[\alpha]^2}(\mathcal{R}')$  and recall that the support of  $y'$  is contained in  $[\alpha]^2$ . Observe that if  $p$  is a point in the support of  $y'$  that fractionally hits some grown rectangle  $r_+$ , then either  $p$ , one of its four immediate neighbors in the grid, or one of its four diagonal neighbors in the grid will hit the original rectangle  $r$ . Define  $y$  as

$$y_q = y'_q + \sum_{p \in N_q} y'_p, \text{ for all } q \in [\alpha]^2, \tag{5}$$

where  $N_q \subseteq [\alpha]^2$  denotes the set of points  $p$  such that  $p$  is an immediate or diagonal neighbor of  $q$ . By the previous observation,  $y$  is a fractional feasible solution for the dual of  $\text{LP}_{[\alpha]^2}(\mathcal{R})$ , and by definition, its value is at most nine times the value of  $y'$ .  $\square$

*Proof of Theorem 6.* Let  $\mathcal{R}$  be a set of rectangles with  $\text{mis}(\mathcal{R}) = \alpha$  and  $\mathcal{R}'$  the set guaranteed by Lemma 2. Then, by combining (2) with (3), we have:

$$\begin{aligned}
\text{mhs}(\mathcal{R}) &\leq O(\text{LP}(\mathcal{R}) \cdot \log \log \text{LP}(\mathcal{R})) \\
&\leq O(\text{LP}(\mathcal{R}') \cdot \log \log \text{LP}(\mathcal{R}')) \\
&= O(\text{LP}_{[\alpha]^2}(\mathcal{R}') \cdot \log \log \text{LP}_{[\alpha]^2}(\mathcal{R}')) \\
&\leq O(\alpha(\log \log \alpha) \cdot \log \log(\alpha \log \log \alpha)) \\
&= O(\alpha(\log \log \alpha)^2). \quad \square
\end{aligned}$$

## 6 Discussion

To conclude the paper, we mention open problems that are worth further investigation. First, note that the computational complexity of MHS is open for all classes of rectangle sets considered in this paper. The complexity of recognizing the intersection graphs of different rectangle sets is also open. It is known that the most general version of this problem, that is, to recognize whether a given graph is the intersection graph of a set of rectangles, is NP-complete [27]. However, little is known for restricted classes. Finally, it would be interesting to determine the duality gap for the classes of rectangle sets studied here.

**Acknowledgments** We thank Vít Jelínek for allowing us to include the lower bound example in Fig. 7, and Flavio Guíñez and Mauricio Soto for stimulating discussions. We also thank two anonymous reviewers whose many suggestions greatly improved the presentation of the paper. This work was supported by Núcleo Milenio Información y Coordinación en Redes ICM/FIC P10-024F and done while the second author was visiting Universidad de Chile. The third author was also supported by FONDECYT Grant 11110069, and the fourth author by FONDECYT Grant 11130266.

## References

1. Adamaszek, A., Wiese, A.: Approximation schemes for maximum weight independent set of rectangles. In: 2013 IEEE 54th Annual Symposium on Foundations of Computer Science (FOCS), pp. 400–409. IEEE, New York (2013)
2. Ahmadinejad, A., Zarrabi-Zadeh, H.: The maximum disjoint set of boundary rectangles. In: Proceedings of the Twenty Sixth Canadian Conference on Computational Geometry (2014)
3. Aronov, B., Ezra, E., Sharir, M.: Small-size  $\varepsilon$ -nets for axis-parallel rectangles and boxes. *SIAM J. Comput.* **39**(7), 3248–3282 (2010)
4. Asano, T.: Difficulty of the maximum independent set problem on intersection graphs of geometric objects. In: 6th ICTAG (1991)
5. Caraballo, L.E., Ochoa, C., Pérez-Lantero, P., Rojas-Ledesma, J.: Matching colored points with rectangles. <http://arxiv.org/abs/1309.3696> (2014)
6. Catanzaro, D., Chaplick, S., Felsner, S., Halldórsson, B.V., Halldórsson, M.M., Hixon, T., Stacho, J.: Max point-tolerance graphs. <http://www.user.tu-berlin.de/chaplick/mpt.pdf> (2013). Accessed Apr 2014
7. Chalermsook, P., Chuzhoy, J.: Maximum independent set of rectangles. In: Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 892–901. Society for Industrial and Applied Mathematics, Philadelphia (2009)
8. Chan, T.M., Har-Peled, S.: Approximation algorithms for maximum independent set of pseudo-disks. *Discrete Comput. Geom.* **48**(2), 373–392 (2012)
9. Chepoi, V., Felsner, S.: Approximating hitting sets of axis-parallel rectangles intersecting a monotone curve. *Comput. Geom.* **46**(9), 1036–1041 (2013)
10. Cibulka, J., Hladký, J., Kazda, A., Lidický, B., Ondráčková, E., Tancer, M., Jelínek, V.: Personal Communication (2011)

11. Fon-Der-Flaass, D.G., Kostochka, A.V.: Covering boxes by points. *Discrete Math.* **120**(1), 269–275 (1993)
12. Fowler, R.J., Paterson, M.S., Tanimoto, S.L.: Optimal packing and covering in the plane are NP-complete. *Inf. Process. Lett.* **12**(3), 133–137 (1981)
13. Golumbic, M.C.: *Algorithmic Graph Theory and Perfect Graphs*. *Annals of Discrete Mathematics*, vol. 57. North-Holland, Amsterdam (2004)
14. Gyárfás, A., Lehel, J.: Covering and coloring problems for relatives of intervals. *Discrete Math.* **55**(2), 167–180 (1985)
15. Kong, H., Ma, Q., Yan, T., Wong, M.: An optimal algorithm for finding disjoint rectangles and its applications to pcb routing. In: *Proceedings of the Forty Seventh ACM/EDAC/IEEE Design Automation Conference*, pp. 212–217 (2010)
16. Hixon, T.S.: *Hook graphs and more: some contributions to geometric graph theory*. Masters thesis, Technische Universität Berlin, Berlin (2013)
17. Hsiao, J.Y., Tang, C.Y., Chang, R.S.: An efficient algorithm for finding a maximum weight 2-independent set on interval graphs. *Inf. Process. Lett.* **43**(5), 229–235 (1992)
18. Imai, H., Asano, T.: Finding the connected components and a maximum clique of an intersection graph of rectangles in the plane. *J. Algorithms* **4**(4), 310–323 (1983)
19. Károlyi, G.Y.: On point covers of parallel rectangles. *Period. Math. Hung.* **23**(2), 105–107 (1991)
20. Knuth, D.E., Raghunathan, A.: The problem of compatible representatives. *SIAM J. Discrete Math.* **5**(3), 422–427 (1992)
21. Lichtenstein, D.: Planar formulae and their uses. *SIAM J. Comput.* **11**(2), 329–343 (1982)
22. Lubiw, A.: A weighted min-max relation for intervals. *J. Comb. Theory Ser. B* **53**(2), 151–172 (1991)
23. Mustafa, N.H., Ray, S.: Improved results on geometric hitting set problems. *Discrete Comput. Geom.* **44**(4), 883–895 (2010)
24. Soto, J.A., Telha, C.: Jump number of two-directional orthogonal ray graphs. In: *Integer Programming and Combinatorial Optimization*, pp. 389–403. Springer, Berlin (2011)
25. Soto, M., Thraves, C.: (c-)And graphs—more than intersection, more than geometric. <http://arxiv.org/abs/1306.1957> (2013)
26. Wegner, G.: Über eine kombinatorisch-geometrische frage Hadwiger und Debrunner. *Isr. J. Math.* **3**(4), 187–198 (1965)
27. Yannakakis, M.: The complexity of the partial order dimension problem. *SIAM J. Algebr. Discrete Methods* **3**(3), 351–358 (1982)