Algorithmica (2022) 84:2222–2239
https://doi.org/10.1007/s00453-022-00964-6

Minimum Hitting Set of Interval Bundles Problem: Computational Complexity and Approximability

Marinus Gottschau[1] · **Marilena Leichter**[1]

Received: 3 September 2019 / Accepted: 25 March 2022 / Published online: 23 April 2022

Abstract
The minimum hitting set of bundles problem (MHSB) is a natural generalization of the minimum hitting set problem, where instead of hitting single elements, bundles of elements are hit. More specifically, we are given a ground set of elements and a family of sets. Every set in this family contains bundles of elements, which are subsets of the ground set. The task is to find a collection of elements of minimum size such that at least one bundle of every set in the family is hit. Motivated by several applications, we consider MHSB restricted to interval and 2-dimensional interval bundles. We study the computational complexity and give polynomial-time algorithms for several classes of instances with these special structured bundles.

Keywords Hitting set · Approximation algorithms · Maintenance scheduling

1 Introduction

The MINIMUM HITTING SET OF BUNDLES PROBLEM (MHSB) was introduced by Angel et al. [1]. It is defined in the following way: Let Ω be a finite set of elements and let $\mathcal{F}$ be a family of sets. Every $F \in \mathcal{F}$ is a set of bundles, where a bundle U is a subset $U \subseteq \Omega$. A bundle U is *covered* by a set of elements $S \subseteq \Omega$ if $U \subseteq S$. We say that a set F is *hit* if at least one bundle in F is covered. We want to find a collection of elements $S \subseteq \Omega$ of minimum size such that every set $F \in \mathcal{F}$ is hit. We refer to S as a hitting set of bundles. In the following, let $\mathcal{U} := \bigcup_{F \in \mathcal{F}} F$ be the set of all bundles of an instance.

The minimum hitting set problem is the special case of MHSB, where every bundle contains exactly one element [1]. It corresponds to the optimization version of

✉ Marilena Leichter
marilena.leichter@tum.de

1 Operations Research, Department of Mathematics, Technische Universität München, Munich, Germany

one of Karp's 21 NP-complete problems, the minimum set cover problem [14]. This immediately implies NP-hardness of MHSB.

MHSB has many applications. In particular, it provides an abstract framework for several scheduling problems. In this paper, we want to highlight two particular applications. The first one is scheduling jobs (non-preemptively) on a single machine with the objective of minimizing active time. It can be framed in the following way. Here, Ω is a set of time slots and $\mathcal{F}$ represents a set of jobs. The bundles of a job refer to its feasible operation times. The goal is to schedule all jobs such that the number of occupied time slots is minimized. If all bundles of a job are time intervals of the same length and refer to possible execution times of the job, the problem is referred to as active (or busy) time minimization problem with unlimited capacities [5, 15].

The second application of MHSB we want to point out is in the area of railway maintenance scheduling. Consider a railway corridor with bidirectional traffic and maintenance jobs that need to be carried out. Here, a train path is a movement over time along the railway track. Whenever such a train path interferes with a particular maintenance job, the train needs to be canceled. To minimize the impact of the mandatory maintenance on rail traffic, the goal is to schedule all maintenance jobs such that the number of canceled trains is as small as possible. Here, Ω represents the set of train paths and $\mathcal{F}$ is the family of maintenance jobs. The bundles of a maintenance job refer to the corresponding sets of train paths that interfere with the feasible execution times of a job. Eskandarzadeh et al. [9] study a variant of this problem in which bundles are determined by job-specific release dates, deadlines, and processing times. For unidirectional train traffic, this agrees with active time minimization with unlimited capacities. In this paper, we also consider the railway maintenance scheduling problem with bidirectional traffic.

MHSB allows more general bundle structures than those studied in the context of active time minimization or railway maintenance scheduling. Still, we often use the scheduling terminology to give an intuition and a better understanding of the particular cases of MHSB that we consider in this paper.

1.1 Related Work

As mentioned before, if every bundle contains only one element, MHSB corresponds to the minimum hitting set problem and its counterpart, the minimum set cover problem. Both problems have been studied extensively. There exists a polynomial-time $\alpha_{\max}$-approximation algorithm, where $\alpha_{\max}$ refers to the maximum number of elements (in MHSB this is the maximum number of bundles) a set may contain and a polynomial-time $(\ln(m) + 1)$-approximation algorithm, where m denotes the number of elements in the minimum set cover instance. For more details on the algorithms, see, e.g., [19]. Dinur and Steurer [8] proved that no polynomial-time $(1 - o(1)) \ln(m)$-approximation algorithm exists for minimum set cover, unless $\mathrm{P} = \mathrm{NP}$. Assuming the Unique Games Conjecture, the approximation factor $\alpha_{\max}$ is best possible, due to a result by Bansal and Khot [2].

The minimum hitting set of bundles problem was introduced by Angel et al. [1]. They presented a polynomial-time $F_{\max}$-factor approximation algorithm, where $F_{\max}$

refers to the maximum number of bundles a set may contain. This approximation guarantee is achieved by considering a relaxation of an integer linear programming (ILP) formulation and a simple rounding strategy. By using randomized rounding they were able to improve the approximation guarantee to $F_{\max}(1-(1-\frac{1}{F_{\max}})^M)$. Here, M denotes the maximum number of bundles an element is contained in. Note that if the same bundle is contained in different sets, it is accounted for several times. Angel et al. [1] highlighted two applications of MHSB, the multiple-query optimization problem in database systems [18] and the min k-SAT problem [3].

Wan et al. [20] studied the minimum submodular cover problem with submodular weights. In this problem, we are given a submodular, increasing function f. A set $S \subseteq \Omega$ is a submodular cover, if $f(S) = f(\Omega)$. The objective is to find a submodular cover of minimum weight with respect to a submodular, increasing weight function w. MHSB can be formulated as submodular cover problem with submodular weights in the following way. Choose the ground set to be the set of bundles $\mathcal{U}$. Define f to be the function that maps a collection of bundles on the number of sets that are hit by at least one bundle of the collection. The weight of a collection of bundles is simply the cardinality of their union. The main result in [20] implies a $H(\gamma)$-approximation for MHSB, where $H(i)$ denotes the i-th Harmonic number and γ is the maximum number of sets a bundle U in $\mathcal{U}$ hits. Iwata and Nagano [13] studied the set cover problem with submodular weights. They derived a polynomial-time $\alpha_{\max}$-approximation algorithm, where $\alpha_{\max}$ is the maximum number of sets an element appears in.

Being a generalization of the minimum hitting set problem, MHSB is $W[2]$-hard parameterized by the solution size $|S|$. Damaschke [7] proved that MHSB, parameterized by $|\mathcal{F}|$ and the solution size $|S|$, is $W[1]$-complete.

The active time minimization problem with capacity B was introduced by Chang et al. [4]. It is a scheduling problem with job-specific release dates, deadlines, and processing times and a bound B on the number of jobs that may be executed simultaneously. Chang et al. gave a polynomial-time algorithm for $B = 2$ and proved that the problem is NP-hard for $B = 3$. More closely related to our problem is the version with unlimited capacity, i.e. $B = \infty$, studied, e.g., by Fong et al. [11]. They presented a polynomial-time algorithm for agreeable deadlines, i.e., instances where any job's deadline is prior to every other job's deadline that has a later release time. Also, Fang et al. [10] studied the problem in context of wireless sensoring and presented a polynomial-time 2-approximation algorithm for this special case. Online variants of the active time minimization problem also exist and have been studied, e.g., by Koehler and Khuller in [16].

Eskandarzadeh et al. [9] studied the maintenance scheduling in a railway corridor problem, which is an application of MHSB. They presented a polynomial-time algorithm for active time minimization with unlimited capacities if all jobs have the same processing time. More specifically, they presented an ILP formulation, tailored to their restricted set of instances, for which they proved total unimodularity of the constraint matrix. In their computational experiments, they compared different ILP formulations for the bidirectional version.

Finally, Chekuri and Kumar [6] introduced a maximization variant of MHSB, the maximum coverage problem with group budgets. In this setting, we are given costs on bundles and a budget per set. The goal is to choose bundles, respecting the budget constraints, such that the size of the union of chosen bundles is maximized. Chekuri and Kumar presented a polynomial-time constant-factor approximation algorithm.

1.2 Our Results

In this paper we study the minimum hitting set of bundles problem on instances with interval bundles and 2-dimensional interval bundles.

Minimum hitting set of interval bundles: The connection to the active time minimization problem leads to a number of applications that can be modeled by MHSB. In Sect. 2, we take advantage of the structure that many of these applications have in common, in order to obtain polynomial-time algorithms. Assuming that jobs have to be executed without preemption, we obtain bundles of consecutive elements. We refer to these bundles as interval bundles and call this special case of MHSB the MINIMUM HITTING SET OF INTERVAL BUNDLES PROBLEM (MHSIB). Motivated by applications, we define the following properties of special cases of MHSIB. We say that $\mathcal{F}$ is *convex* if for every job $F \in \mathcal{F}$, the union of all possible operating times forms an interval. We say that $\mathcal{F}$ is *a-simple* $a \in \mathbb{N}$ if all bundles have size a. In the scheduling terminology this corresponds to equal processing times. In applications it also seems reasonable to assume that, for example, the number of starting times is bounded or the time horizon of feasible operating times for a job is bounded.

We present polynomial-time algorithms for several classes of interval bundle instances. These algorithms use a graph construction and solve the problem by computing a shortest path. However, the minimum hitting set of interval bundles problem is NP-hard in general. We explore the boundary of polynomial-time solvable instances and NP-hardness that arise from the aforementioned properties and parameters.

Minimum hitting set of 2-dimensional interval bundles: In Sect. 3, motivated by the application of MHSB to railway maintenance scheduling, we study another special class of instances of MHSB. In the maintenance scheduling in a railway corridor problem presented by Eskandarzadeh et al. [9], Ω is the disjoint union of sets of train paths in opposite directions on a single railway track. More generally, we can think of Ω as being the disjoint union of two totally ordered sets. Assuming that every job has to be executed without preemption implies that every bundle is a set of consecutive elements in each of the two totally ordered sets. We call this special structure *2-dimensional interval* and denote the special case of MHSB where all bundles are 2-dimensional interval by MINIMUM HITTING SET OF 2- DIMENSIONAL INTERVAL BUNDLES PROBLEM (2- DIM MHSIB). We show that 2- DIM MHSIB remains NP-hard on convex and 1-simple instances with $F_{\max} = 2$. Furthermore, we present a polynomial-time approximation algorithm for all 1-simple instances with $F_{\max} = 2$ using a result by Hochbaum [12] with an approximation guarantee slightly better than 2. For another restricted class of instances we give a polynomial-time algorithm by making use of a problem decomposition.

Remark 1 The introduction of a weight function $w : \Omega \to \mathbb{R}_+$ on the elements may be of interest. For the ease of presentation, we only consider the unweighted MHSB. Techniques used throughout the paper also work for weighted instances, thus all of our results can be easily transferred to weighted MHSB. These results can either be transferred directly with minor modifications of inequalities, or by simple modifications of the constructed graph's edge weights.

Remark 2 Due to the structure of the problem, we assume, w.l.o.g., that

(i) all bundles are non-trivial, that is, for all $F \in \mathcal{F}$ we have that $\emptyset \notin F$;
(ii) bundles of a set are not subsets of each other, that is, for $U \in F$, if $U' \subset U$ we have that $U' \notin F$; and
(iii) for all distinct $F, F' \in \mathcal{F}$ there exists a bundle $U \in F$ such that for all $U' \in F'$ we have $U' \nsubseteq U$. If there was no such bundle $U \in F$, any set of elements hitting F also hits F'.

2 Minimum Hitting Set of Interval Bundles

In this section, we focus on bundle structures that arise in many applications. As mentioned in the introduction, Ω often corresponds to time slots if viewed as a scheduling problem. In this context, one may assume that jobs have to be executed without preemption. In the following, we make use of this particular structure and identify polynomial-time solvable special cases of MHSB.

Throughout this section we look at instances in which we are given an ordering $\prec$ of Ω. To simplify notation, we may assume that $\Omega = [n]$ for some $n \in \mathbb{N}$ with the natural ordering. In this context, we consider a special case of MHSB, the MINIMUM HITTING SET OF INTERVAL BUNDLES PROBLEM (MHSIB), where every bundle U corresponds to an interval. More specifically, if $i, j \in U$, then $k \in U$ for all $i \leq k \leq j$. In addition to the interval property, we define two other properties that a family $\mathcal{F}$ may have.

Definition 1 Let $([n], \mathcal{F})$ be an instance of MHSIB. We call the family $\mathcal{F}$

(i) *convex*, if all $F \in \mathcal{F}$ are convex, i.e. the following holds. Let $i \in U'$, $j \in U''$, for some $U', U'' \in F$. For all $k \in [n]$ with $i \leq k \leq j$ there exists a bundle U in F such that $k \in U$. In other words, the union of all interval bundles U in F is again an interval.
(ii) *a-simple* for some $a \in \mathbb{N}$ if all bundles are a-simple, i.e. $|U| = a$ for all $U \in \mathcal{U}$, where $\mathcal{U} := \bigcup_{F \in \mathcal{F}} F$.

Figure 1 is an example of an instance of the general MHSB problem and visualizes the properties we defined above. The next theorem shows that MHSIB remains NP-hard even on restricted instances in terms of Definition 1.

Theorem 1 *Let* $([n], \mathcal{F})$ *be an instance of* MHSIB *and let* $F_{\max}$ *be the maximum number of bundles a set may contain. The problem remains NP-hard, if* $\mathcal{F}$

(i) is 1*-simple and* $F_{\max} = 2$*; or*
(ii) is convex and $F_{\max} = 3$*; or*
(iii) is convex and a-simple for some a, where a is some function in n.

Proof (i) An easy reduction from vertex cover implies the statement. Given a graph $G = (V, E)$, we let $\Omega := V$ and $\mathcal{F} := \{\{\{u\}, \{v\}\} : \{u, v\} \in E\}$. Observe that a minimum vertex cover is an optimal solution to MHSIB and vice versa.

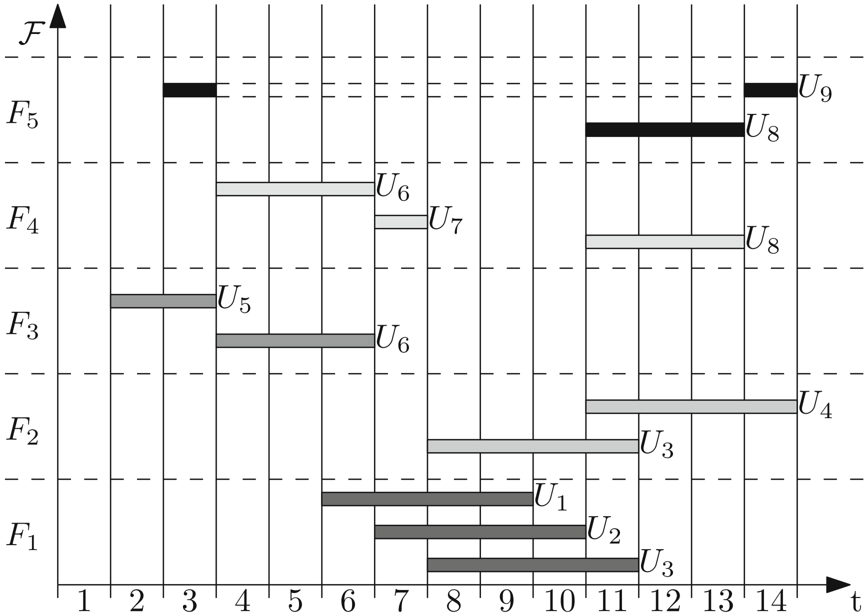

Fig. 1 Example of a family $\mathcal{F}$ to visualize the different sets and bundles properties: All bundles but U_9 have the property that they are interval. The sets F_1, F_2 and F_3 are convex, and F_1 and F_2 are in addition to that a-simple with $a = 4$. The bundle structure of the set F_1 is determined by the release date 6, deadline 11 and processing time 4. This overlapping bundle structure occurs for example in active time minimization

(ii) The statement follows again from a reduction of vertex cover to MHSIB. Let $G = (V, E)$ be a graph with an arbitrary ordering $\prec_V$ of the vertices. The main idea is to use the same construction as in i), i.e. introducing a set for every edge $\{u, v\}$ of the graph. The challenge is to guarantee that our constructed family is convex, while respecting the structure assumed in Remark 2. We will do so by constructing sets that additionally contain a large dummy bundle. Because of its size, the dummy bundle is never entirely covered by an optimal solution. A large number of dummy elements serves to ensure a one-to-one correspondence between a minimum hitting set of bundles and a minimum vertex cover.

We begin by constructing the ground set of the corresponding instance of MHSIB. We introduce d (any $d \geq |V| + 2$ will work) elements for each vertex of the graph, that is, $\Omega_v := \{\omega_i^v : i \in [d]\}$ for all $v \in V$. A vertex $v \in V$ is represented by the element ω_1^v and we refer to the $d - 1$ elements $\omega_2^v \dots \omega_d^v$ as dummy elements. The ground set is now given by $\Omega := \bigcup_{v \in V} \Omega_v$.

We use the ordering $\prec_V$ of V to obtain an ordering $\prec_\Omega$ of Ω in a natural way by

$$\omega_i^u \prec_\Omega \omega_j^v \Leftrightarrow \begin{cases} u \prec_V v; & \text{or} \\ u = v \quad \wedge \quad i < j. \end{cases}$$

Next, we define bundles and sets of the instance. As in the construction of i) for every edge $\{u, v\} \in E$ with $u \prec_V v$ we are given two bundles $\{\omega_1^u\}$ and $\{\omega_1^v\}$. In addition, we define a dummy bundle $U_{\{u,v\}}$ containing all elements in Ω that are between ω_1^u and ω_1^v. Formally, we define

$$U_{\{u,v\}} := \Omega_u \setminus \{\omega_1^u\} \cup \bigcup_{u \prec_V w \prec_V v} \Omega_w.$$

Note, that all constructed bundles are indeed interval with respect to the ordering of Ω induced by $\prec_\Omega$. The sets are then formally given by

$$\mathcal{F} = \left\{ \left\{ \{\omega_1^u\}, \{\omega_1^v\}, U_{\{u,v\}} \right\} \ : \ \{u, v\} \in E \right\}.$$

The dummy bundle $U_{\{u,v\}}$ contains a large number of elements, i.e. at least d and, hence, is never entirely hit in an optimal solution. Its only purpose is to ensure convexity of the respective family.

Any vertex cover C corresponds to a solution of the MHSIB instance of the same size. This simply follows from the fact that the vertices in C cover every edge and therefore all sets $F \in \mathcal{F}$ contain at least one covered bundle. Additionally, a solution S of MHSIB of minimum size only contains elements that correspond to vertices of the graph, i.e. $S \subseteq \{\omega_1^v : v \in V\}$. This follows from the fact that if S covers any bundle $U_{\{u,v\}}$, this immediately implies $|S| \geq d - 1 = |V| + 1$. Also, any element in $\Omega \setminus \{\omega_1^v : v \in V\}$ can be omitted unless it was needed to cover some $U_{\{u,v\}}$.

Finally, we have that S hits all sets $F \in \mathcal{F}$ and covers the bundle $\{\omega_1^u\}$ or $\{\omega_1^v\}$ for all $\{u, v\} \in E$, with either ω_1^u or ω_1^v. Therefore, the set of vertices represented by elements in S form a vertex cover in the graph.

(iii) Again, we prove the statement by a reduction of vertex cover to MHSIB. Let $G = (V, E)$ and let $\prec_V$ an arbitrary ordering of the vertices. This time we have to ensure that the constructed family is convex and a-simple. For every edge $\{u, v\}$ we construct a corresponding set $F_{\{u,v\}}$ in the family. The set contains multiple a-simple bundles, amongst them the bundles U_u and U_v which correspond to the vertices u and v. Instead of a single large dummy bundle as in ii), the sets contain a collection of dummy bundles, each of size a. To ensure that the optimal solution to MHSIB does indeed hit every set by covering a bundle corresponding to a vertex in V, we introduce additional dummy sets F_v for every $v \in V$. The set F_v contains a single bundle, that overlaps with the bundle U_v in all but one element. Since there is only one bundle in F_v, it has to be in every feasible solution of the instance of MHSIB. This ensures that hitting a set $F_{\{u,v\}}$ by a bundle U_v or U_u only requires adding one additional element to the solution, whereas hitting $F_{\{u,v\}}$ by any dummy bundle immediately increases the solution size significantly (by at least $\frac{a-1}{2}$).

Again, we start by constructing the ground set of the corresponding instance of MHSIB. Formally, let $a := 4|V| + 1$ and let v_0 be a dummy vertex with $v_0 \prec_V v$ for all $v \in V$. For every vertex in V and the dummy vertex v_0 we introduce $3a$ elements, that is, $\Omega_v := \{\omega_i^v : i \in [3a]\}$ for all $v \in V \cup \{v_0\}$. Intuitively, the element ω_1^v corresponds to the vertex $v \in V$ and all other elements in Ω_v are dummy elements. The ground set of elements is then given by $\Omega := \bigcup_{v \in V \cup \{v_0\}} \Omega_v$. As in ii) we extend the ordering $\prec_V$ of the vertices in a natural way to an ordering $\prec_\Omega$ of the elements in Ω.

We continue by constructing the bundles and the sets of the family $\mathcal{F}$ of the corresponding instance of MHSIB. For a vertex $v \in V$ let U_v be the bundle that contains the element ω_1^v and the $a-1$ previous elements with respect to $\prec_\Omega$. Formally,

$$U_v := \{\omega_{2a+2}^{v'}, \omega_{2a+3}^{v'}, \ldots, \omega_{3a-1}^{v'}, \omega_{3a}^{v'}, \omega_1^v\}$$

where v' is the vertex preceding v in the natural ordering. Note that all bundles U_v are indeed a-simple.

Next, we define a collection B of dummy bundles, which is a partition of $\Omega \setminus \{\omega_1^{v_0}, \omega_2^{v_0}, \ldots, \omega_{\frac{5a+1}{2}}^{v_0}\}$ in a-simple, disjoint bundles. B will be later used to ensure convexity of the family $\mathcal{F}$. More explicitly B is given by

$$B := \bigcup_{v \in V} \Big\{ \underbrace{\{\omega_{\frac{5a+3}{2}}^{v'}, \ldots, \omega_{3a}^{v'}, \omega_1^v, \ldots, \omega_{\frac{a+1}{2}}^v\}}_{a}, \underbrace{\{\omega_{\frac{a+3}{2}}^v, \ldots, \omega_{\frac{3a+1}{2}}^v\}}_{a}, \underbrace{\{\omega_{\frac{3a+3}{2}}^v, \ldots, \omega_{\frac{5a+1}{2}}^v\}}_{a} \Big\},$$

where, again, v' is the vertex preceding v in the natural ordering. Note that the bundles in B are well-defined, since we chose a to be odd. We continue by constructing the sets of the family $\mathcal{F}$. For every edge $\{u, v\} \in E$ the family $\mathcal{F}$ contains a set

$$F_{\{u,v\}} = \{U_u, U_v\} \cup B.$$

In addition to that, for every vertex $v \in V$ we add a dummy set containing a single bundle with the a elements preceding ω_1^v to $\mathcal{F}$. More explicitly,

$$F_v := \{\{\omega_{2a+1}^{v'}, \omega_{2a-2}^{v'}, \ldots, \omega_{3a-1}^{v'}, \omega_{3a}^{v'}\}\}$$

where, again, v' is the vertex preceding v in the natural ordering. This concludes the construction of the family. A visualization can be seen in Fig. 2.

Let C be a minimum vertex cover in G. We obtain the corresponding hitting set of bundles in the following way. As already mentioned, every bundle of a dummy set F_v has to be part of a feasible solution. Next, we ensure that every $F_{\{u,v\}}$ is hit. Adding the element ω_1^v for every $v \in C$ to the hitting set of bundles, we cover the bundle U_v and, hence, hit the corresponding set $F_{\{u,v\}}$. All in all, we obtain a hitting set of bundles of size $a \cdot |V| + |C|$. Note, every set $F_{\{u,v\}}$ is hit by one of the bundles U_u or U_v as either u or v must be contained in C. We claim that there is no solution to MHSIB of smaller cardinality. Observe, that an optimal solution to MHSIB has to hit every set F_v, each containing only one bundle of size a. Since all these sets are pairwise disjoint, a solution must be of size at least $a \cdot |V|$. Any bundle from the collection B contains at least $a/2 > |V|$ elements not contained in the bundles of the dummy sets F_v that are covered by the solution. Again, if such a bundle from B is covered the solution is of size at least $a \cdot |V| + |V| + 1$, which cannot be optimal by the previous argument, that there always exists a solution of size at most $a \cdot |V| + |V|$. □

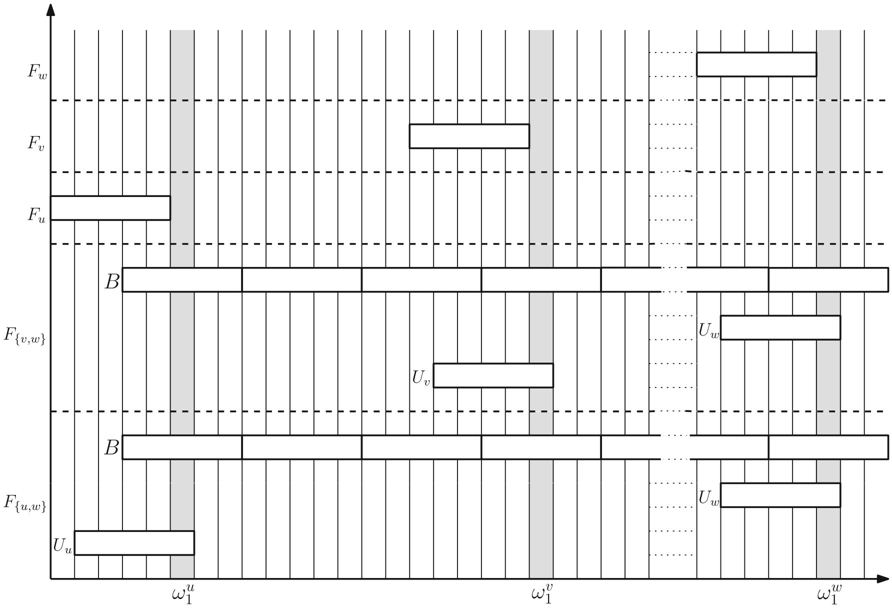

Fig. 2 Illustration of a constructed MHSIB instance in the proof of proof of Theorem 1 (iii)

Theorem 1 states that MHSIB remains NP-hard for convex families $\mathcal{F}$ with $F_{\max} \leq 3$ for all $F \in \mathcal{F}$. However, the following theorem shows that the computational complexity changes for convex instances if we reduce $F_{\max}$ by one.

Theorem 2 *Let* $([n], \mathcal{F})$ *be an instance of* MHSIB. *If* $\mathcal{F}$ *is convex and* $F_{\max} = 2$*, then it is solvable in polynomial time.*

Proof We describe a reduction to a shortest path computation. The high level idea is to construct a layered graph, such that the vertices contained in an s-t-path describe an interval decomposition of the solution.

Before explaining the construction in more details, let us take a closer look at the bundle structure of our instance to give some intuition. Since every set contains at most two bundles, w.l.o.g., we may assume that for all $F \in \mathcal{F}$, if $U, U' \in F$ then $U \cap U' = \emptyset$, as elements in the intersection have to be contained in any feasible solution. Given some $F \in \mathcal{F}$, since $\mathcal{F}$ is convex, there exist l, i and $u \in [n]$ such that $F = \{[l, i], [i+1, u]\}$. To make sure that the bundle is hit, we have to guarantee that either the set of elements $\{l, \ldots i\}$ or $\{i+1, \ldots, u\}$ is contained in the minimum hitting set of bundles.

Let $([n], \mathcal{F})$ be an instance of MHSIB. We define a graph, with layers $V_1, \ldots V_n$ corresponding to the elements $1, \ldots n$. A set V_i contains a vertex for every interval in $[1, n]$ containing i and, additionally, a vertex $v^i_\varnothing$ representing the empty set. Formally, for every $i \in [n]$ we define $V_i := \{v^i_{[a,b]} \mid 1 \leq a \leq i \leq b \leq n\} \cup \{v^i_\varnothing\}$. Note that $a, b \leq n$ and therefore the number of vertices introduced for each $i \in [n]$ is polynomial in the size of the input. Let $V_0 := \{s\}$ and $V_{n+1} := \{t\}$. Then the vertex set of the layered graph is given by $V := \bigcup_{i=0}^{n+1} V_i$.

The graph being layered, means that there only exist edges from layer V_i to the subsequent layer V_{i+1}. In particular, every shortest s-t-path contains exactly one vertex v^i from every layer V_i. Intuitively, if $v^i = v^i_\varnothing$, then the element i is not part of the corresponding minimum hitting set of bundles. If $v^i = v^i_{[a,b]}$, we add i to the solution. In particular, the solution contains all elements of the interval $[a, b]$. To ensure this, we define the edges and edge weights of the layered graph in the following way. There are four feasible types of edges:

(1) $\{v^i_\varnothing, v^{i+1}_\varnothing\}$ of weight 0,
(2) $\{v^i_\varnothing, v^{i+1}_{[i+1,b]}\}$ of weight 1,
(3) $\{v^i_{[a,b]}, v^{i+1}_{[a,b']}\}$ for some $b' \geq b$ of weight 1, and
(4) $\{v^i_{[a,i]}, v^{i+1}_\varnothing\}$ of weight 0.

Additionally, we add edges $\{s, v^1\}$ for all $v^1 \in V_1$ of weight 0 if $v^1 = v^1_\emptyset$ and of weight 1 in all remaining cases. Also, we add edges $\{v^n, t\}$ for all $v^n \in V_n$, all of weight 0. We only have positive weights of 1 on arcs of type $\{s, v^1_{[1,b]}\}$ and $\{v^i_\varnothing, v^{i+1}_{[i+1,b]}\}$ and $\{v^i_{[a,b]}, v^{i+1}_{[a,b']}\}$. These correspond to adding the element 1 and $i + 1$ to the hitting set of bundles, respectively.

In a next step, we tailor the graph to the set structure of our specific instance by deleting certain edges. This step ensures that every set is hit. Consider a set $F \in \mathcal{F}$ with $F = \{[l, i], [i+1, u]\}$. If not all elements $\{l, l+1, \dots, i\}$ are part of the solution, we have to make sure that all elements in $\{i + 1, \dots, u\}$ are added to the solution. To do so, for every $F \in \mathcal{F}$ with $F = \{[l, i], [i + 1, u]\}$, we delete all edges of the following types:

(1) the edge $\{v^i_\varnothing, v^{i+1}_\varnothing\}$,
(2) edges $\{v^i_\varnothing, v^{i+1}_{[i+1,b']}\}$ for which $b' < u$,
(3) edges $\{v^i_{[a,b]}, v^{i+1}_{[a,b']}\}$ for which $a > l$ and $b' < u$, and
(4) edges $\{v^i_{[a,i]}, v^{i+1}_\varnothing\}$ for which $a > l$.

As argued above, doing so for all sets in $\mathcal{F}$, any s-t-path gives rise to a hitting set of bundles. More formally, let $s = v^0_{I_0}, v^1_{I_1}, \dots, v^n_{I_n}, v^{n+1}_{I_{n+1}} = t$ an s-t-path (where $I_0 = I_{n+1} = \varnothing$). Then $S = \bigcup_{0 \leq i \leq n+1} I_i$ forms a hitting set of bundles of cardinality of the length of the path. To see that, let $F = \{[l, i], [i + 1, u]\} \in \mathcal{F}$ be an arbitrary set of bundles. Since $\{v^i_{I_i}, v^{i+1}_{I_{i+1}}\}$ is an edge in the graph, we have that $[l, i] \subseteq I_i$ or $[i + 1, u] \subseteq I_{i+1}$. This implies that S hits F and, hence, forms a feasible hitting set of bundles. Observe, that S is a collection of intervals and every $\{v^i_{I_i}, v^{i+1}_{I_{i+1}}\}$ of the path has weight 1 if and only if $i + 1 \in I_{i+1}$. Hence $i + 1 \in S$ for $0 \leq i \leq n - 1$ by the above construction. Note, $\{v^n_{I_n}, t\}$ has weight 0. Additionally, any hitting set of bundles S with its interval decomposition $J_1 = [l_1, u_1], \dots, J_k = [l_k, u_k]$, such that for all $j, j' \in [k]$ with $j \neq j'$ we have $J_j \cap J_{j'} = \varnothing$, gives rise to an s-t-path. I.e., let

$$I_i = \begin{cases} \varnothing & \text{if } i \notin S = \bigcup_j J_j, \\ J_j & \text{if } i \in J_j \subseteq S, \end{cases}$$

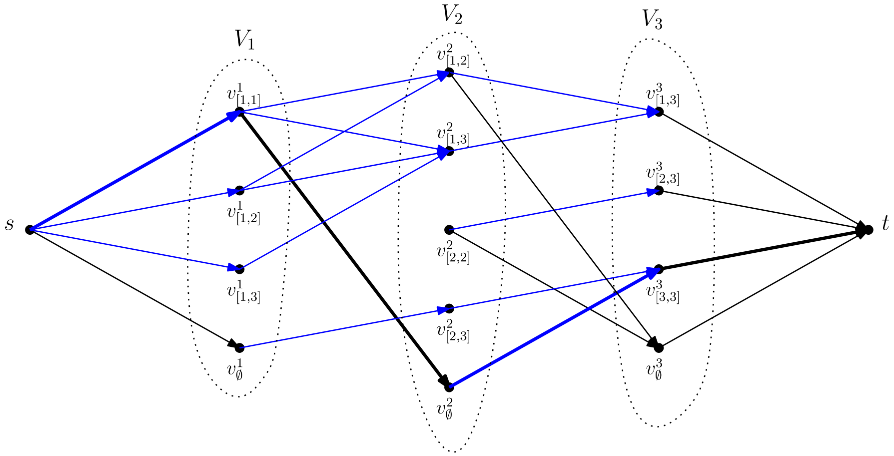

Fig. 3 Illustration of the constructed graph on an instance of 3 elements $\Omega = \{1, 2, 3\}$ and set family $\mathcal{F} = \{F_1, F_2\}$ with $F_1 = \{\{1\}, \{2, 3\}\}$ and $F_2 = \{\{2\}, \{3\}\}$. Blue edges have weight 1, black edges weight 0. A shortest s-t-path is highlighted in bold. It corresponds to the minimum hitting set of bundles containing the elements $\{1\}$ and $\{3\}$

then $s, v^1_{I_1}, \ldots, v^n_{I_n}, t$ is an s-t-path of length of the cardinality of S, as for every $F = \{[l, i], [i + 1, u]\}$ we have some J with $[l, i] \subseteq J$ or $[i + 1, u] \subseteq J$ and hence $(v^i_{I_i}, v^{i+1}_{I_{i+1}}) \in E$. A depiction of the constructed graph for a small instance is displayed in Fig. 3.

We now compute a shortest s-t-path in the graph. Correctness of the algorithm follows from the one-to-one correspondence of s-t-paths and the respective interval decomposition of a hitting set of bundles and the fact that the weight of any path equals the sum of the cardinality of the sets (intervals). $\square$

Theorem 1 also states that MHSIB remains NP-hard if $\mathcal{F}$ is a-simple and convex. The following theorem considers two different types of convex instances with an additional property. In scheduling terminology, the first type of instances corresponds to the case, where the difference of any job's release date and deadline is bounded by k. In ii) the processing time of every job is bounded by k and the number jobs intersection with time slot i is bounded by $f_{\max}$.

Theorem 3 *Let* $([n], \mathcal{F})$ *be an instance of* MHSIB. *The problem is solvable*

(i) in $O(2^{2k}n)$ *if* $|i - j| \leq k$ *for all* $i, j \in \bigcup_{U \in F} U$ *and all* $F \in \mathcal{F}$*; or*
(ii) in $O(2^{2(k+f_{\max})}n)$ *if* $|U| \leq k$ *for all* $U \in \mathcal{U}$*,* $\mathcal{F}$ *is convex and for all* $i \in [n]$*,*

$$|\{F \in \mathcal{F} \mid \exists\, U \in F : i \in U\}| \leq f_{\max}.$$

Proof (i) The key idea is to construct a weighted graph, whose shortest s-t-path gives rise to a decomposition of a solution of the corresponding MHSIB instance. Here, we make use of the fact that the size of every set is bounded by k and, thus, all subsets of elements that intersect and hit a set F can be represented by

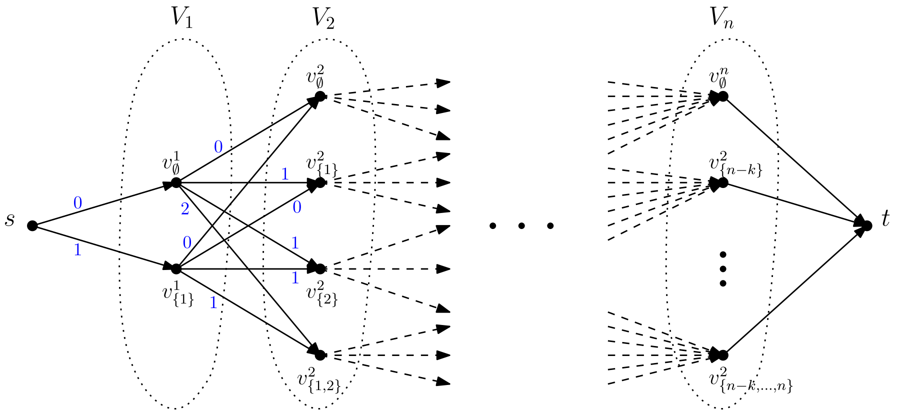

Fig. 4 Illustration of the constructed graph. Exemplary edge weights are highlighted in blue

a bounded number of sets. Note that we say that an element i intersects a set F if there exists a bundle $U \in F$ with $i \in U$.

We start by giving a formal construction of the graph $G = (V, E)$. First, define the vertex set $V := V_0 \cup \cdots \cup V_{n+1}$ with $V_i := \{v^i_S : S \subseteq \{i - \min\{i, k\}, \dots, i\}\}$, $V_0 := \{s\}$ and $V_{n+1} := \{t\}$. Note that $|V_i| \leq 2^k$. In the following, for all $i \in [n]$ we let $\mathcal{F}_i$ be the family of sets with largest element i, i.e. $i = \max\{j \in \bigcup_{U \in F} U\}$ for all $F \in \mathcal{F}_i$. The set of arcs E is then defined as follows. Our graph G only has arcs between subsequent layers, that is, there only exist arcs from V_i to V_{i+1}. More specifically, arcs from a vertex $v^i_S \in V_i$ to all other vertices in V_{i+1} exist if S hits all sets in $\mathcal{F}_i$. If $\mathcal{F}_i = \emptyset$, we include all arcs in $V_i \times V_{i+1}$. Finally, all arcs from s to V_1 are contained in the edge set.

The weight of each arc $(v^i_S, v^{i+1}_{S'})$ is given by the number of elements in $S' \setminus S$. Arcs (s, v^1_S) have weight $|S|$, and the exiting arcs (v^n_S, t) have weight 0. This concludes the construction of the graph G. For a schematic picture of such a constructed graph see Fig. 4.

Since the size of every set is bounded by k, i.e. $|i - j| \leq k$ for all $i, j \in \bigcup_{U \in F} U$, every subset of elements that hits all sets in F_i is a subset of $\{i - \min\{i, k\}, \dots, i\}\}$. Observe that any shortest s-t-path P passes through exactly one vertex of every V_i. On that path, each vertex v^i_S represents a set of elements S. Let $S^* := \bigcup_{S : v^i_S \in P} S$. We claim that S^* is an optimal solution to the corresponding Mhsib instance. Every set F has a largest element i, and only vertices v^i_S where S hits all sets in $\mathcal{F}_i$ have outgoing arcs. Thus, S^* hits every set in $\mathcal{F}$ and, hence, S^* is feasible. Observe, that by the choice of the weight function the length of any shortest s-t-path P equals $|S^*|$. On the other hand every hitting set of bundles S^* gives rise to an s-t-path of same cost in the corresponding graph G. The path is obtained by traversing the respective vertices of the sets $\{i - \min\{i, k\}, \dots, i\} \cap S^*$ for all $i \in [n]$. Note that the size of the constructed graph is in $O(2^{2k} n)$, following by the bounds on $|V_i|$ and the fact that we only have edges between subsequent layers V_i and V_{i+1}. A simple BFS finds a shortest path in linear time of the size of the graph.

(ii) Again we make use of a graph $G = (V, E)$ defined on $V := V_0 \cup \cdots \cup V_{n+1}$. Here,

$$V_i := \mathcal{P}(\{i - \min\{i, k\}, \ldots, i\}) \times \mathcal{P}(\{F \in \mathcal{F} \mid \exists U \in F : i \in U\}),$$

where $\mathcal{P}$ denotes the power set. $V_0 := \{s\}$ and $V_{n+1} := \{t\}$. In other words, every vertex in V_i for $i \in [n]$ corresponds to a tuple $(S^{\Omega}, S^{\mathcal{F}})$ with $S^{\Omega} \subseteq \{i - \min\{i, k\}, \ldots, i\}$ and $S^{\mathcal{F}} \subseteq \{F \in \mathcal{F} \mid \exists U \in F : i \in U\}$. The set $S^{\mathcal{F}}$ shall be used to encode which of the sets containing element i already have been hit. Note that by assumption, the number of vertices in each set V_i is bounded by the number of tuples, which is at most $2^k \cdot 2^{f_{\max}}$.

Given the vertex set as defined above, we only allow arcs between subsequent layers, that is, between vertices in V_i and vertices in V_{i+1}. More specifically, there exists an arc from $(S^{\Omega}, S^{\mathcal{F}}) \in V_i$ towards $(\bar{S}^{\Omega}, \bar{S}^{\mathcal{F}}) \in V_{i+1}$ if the two following conditions hold:

(a) for all $F \in \bar{S}^{\mathcal{F}} \setminus S^{\mathcal{F}}$ there exists a bundle U in F such that $U \subseteq \bar{S}^{\Omega}$. That is, $F \in \bar{S}^{\mathcal{F}} \setminus S^{\mathcal{F}}$ if $\bar{S}^{\Omega}$ covers one bundle in F; and
(b) if $F \in \mathcal{F}_i$, then $F \in S^{\mathcal{F}}$. This ensures that any s-t-path corresponds to a feasible solution of the respective instance of MHSIB.

All arcs in $\{s\} \times \{(S^{\Omega}, S^{\mathcal{F}}) \in V_1 : S^{\mathcal{F}} = \emptyset\}$ are contained in E. Let the weight function $w : E \to \mathbb{N}$ be defined as

$$w((S^{\Omega}, S^{\mathcal{F}}), (\bar{S}^{\Omega}, \bar{S}^{\mathcal{F}})) = |\bar{S}^{\Omega} \setminus S^{\Omega}|.$$

We claim, that an optimal solution to MHSIB can be obtained by computing a shortest s-t-path in the corresponding graph G. A solution to the MHSIB instance is, again, obtained by taking the union of all sets S^{Ω} of elements represented by the traversed vertices. Feasibility of the hitting set of bundles follows directly from the construction of the arcs in b). Moreover, the size of the solution equals the length of the path.

Similarly to i), by the choice of the weights, any solution to the corresponding instance of MHSIB gives rise to an s-t-path of same length. At this point we make use of the fact that the family is convex, which implies that the set of elements that is contained in a set forms an interval. Without this condition, there might exist layers i,j with $i < j - 1$, and a set F such that i and j intersect F but $j - 1$ does not. Due to condition b), this forces the set F to be hit by the elements from $\{j, \ldots, n\}$, even though in an optimal solution the set F might only be hit by elements from $\{1, \ldots, j - 1\}$.

Note that the size of the constructed graph is in $O(2^{2(k+f_{\max})}n)$, following by the bounds on $|V_i|$ and the fact that we only have edges between subsequent layers V_i and V_{i+1}. A simple BFS finds a shortest path in linear time of the size of the graph. $\square$

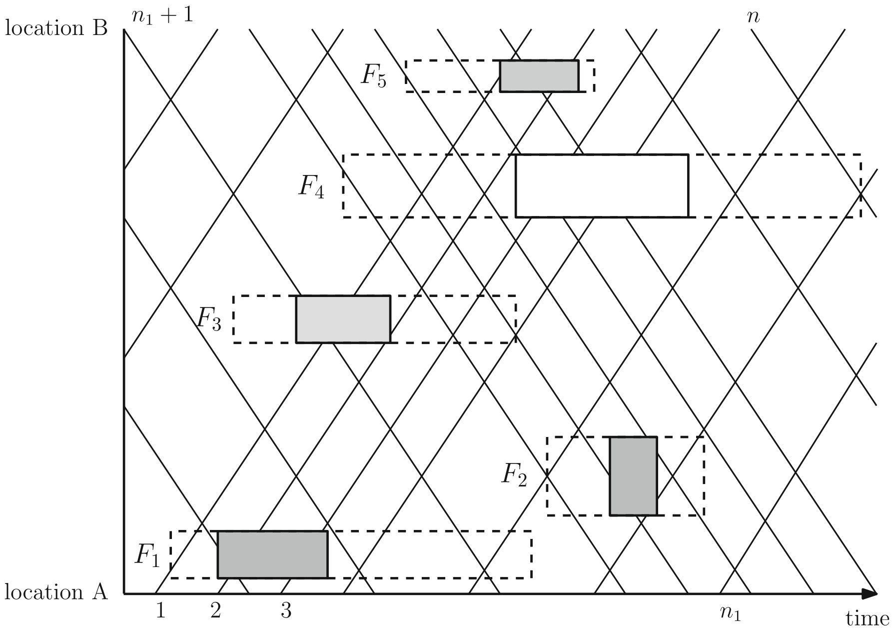

Fig. 5 Example of an instance of railway maintenance. We are given a railway corridor between location A and location B with bidirectional traffic. The parallel lines represent train paths. The paths $\{1, \ldots n_1\}$ correspond to trains from A to B and the paths $\{n_1 + 1, \ldots n\}$ correspond to trains from B to A. Dashed boxes represent maintenance jobs. In particular, the height of a box corresponds to the section of the railway corridor that requires maintenance work and the length corresponds to the time window in which a job has to be carried out. The length of a solid box represents the processing time of a job. This can be formulated as an instance of 2- DIM MHSIB in the following way: The ground set of elements is given by the set of train paths. For every dashed box we introduce a set F. The bundles in F are determined by the sets of train path that interfere with the solid box, given a certain position within the dashed box

3 Minimum Hitting Set of 2-Dimensional Interval Bundles

So far we focused on instances with a given total ordering of Ω. In this section, we consider instances with $\Omega := N_1 \dot\cup N_2$ where we are given a total ordering $\prec_1$ of the elements in N_1 and a total ordering $\prec_2$ of the elements in N_2. Throughout this section we may assume, w.l.o.g., that $N_1 := \{1, \ldots, n_1\}$ as well as $N_2 := \{n_1 + 1, \ldots, n\}$ with cardinalities $n_i := |N_i|$ for $i \in \{1, 2\}$.

As mentioned in the introduction, this setting is also motivated by an application to railway maintenance ([9] and see Fig. 5). Interpret N_1 and N_2 as two sets of train paths in opposite direction on a railway track and let $\Omega := N_1 \dot\cup N_2$. If we assume that maintenance jobs are executed without preemption, we obtain bundles with a very specific structure, so-called 2-dimensional interval bundles. More specifically, a set of bundles is *2-dimensional interval* if for all $U \in \mathcal{U}$, if $i_1, i_2 \in U$ with $i_1, i_2 \in N_1$, then for all $i' \in N_1$ with $i_1 \leq i' \leq i_2$ we also have that $i' \in U$. Analogously, if $j_1, j_2 \in U$ with $j_1, j_2 \in N_2$, then for all $j' \in N_2$ with $j_1 \leq j' \leq j_2$ we also have that $j' \in U$. We refer to this problem as the MINIMUM HITTING SET OF 2- DIMENSIONAL INTERVAL

BUNDLES PROBLEM (2- DIM MHSIB) and define the properties of being convex and a-simple as follows.

Definition 2 Let $(N_1 \dot{\cup} N_2, \mathcal{F})$ be an instance of 2- DIM MHSIB. For $i \in \{1, 2\}$, let $\mathcal{F}_{|N_i}$ be the family of sets restricted to N_i, that is, $\mathcal{F}_{|N_i} := \{(\bigcup_{U \in F}\{U \cap N_i\})\setminus\{\emptyset\} : F \in \mathcal{F}\}$. We call the family $\mathcal{F}$

(i) *2-dim-convex* if $\mathcal{F}_{|N_1}$ and $\mathcal{F}_{|N_2}$ are convex.
(ii) *a-simple* for some $a \in \mathbb{N}$ if $|U| = a$ for all $U \in \mathcal{U}$.

Remark 3 In applications of 2- DIM MHSIB, such as the maintenance scheduling problem, the property of $\mathcal{F}$ being 2-dim-convex refers to a setting where the starting time of a job can be anywhere in a given time window. a-simple is a natural property in a regular train schedule setting where, independently of the starting time, a job always interferes with the same number of train paths.

2- DIM MHSIB remains NP-hard even on 2-dim-convex families [9]. We are able to prove a slightly stronger result.

Theorem 4 *The* 2- DIM MHSIB *problem remains NP-hard if $\mathcal{F}$ is 2-dim-convex,* 1-*simple, and* $F_{\max} = 2$.

Proof We use the following graph construction to reduce vertex cover to an instance with the required properties. Given a graph $G = (V, E)$, subdivide each edge $e = \{u, v\} \in E$ into three edges $\{u, e_u\}$, $\{e_u, e_v\}$, $\{e_v, v\}$ and denote the resulting graph by $G' = (V', E')$. Note that any vertex cover S in G yields a vertex cover in G' if we add exactly one vertex e_u, e_v for every edge $e = \{u, v\} \in E$ to the cover. This new cover S' has size $S + |E|$.

Also, w.l.o.g. a minimum vertex cover S' in G' contains exactly one vertex e_u, e_v for every edge $e = \{u, v\} \in E$ (If not, it is either not a cover or we can add u and remove e_u from the cover without changing its cardinality). We claim that $S' \cap V$ is a vertex cover in G. Assume the contrary, i.e. there is an edge $e = \{u, v\}$ that is not covered. Then, since S' was a cover in G', the edges $\{e_u, e_v\}$ must be covered by S' by either e_u or e_v. Assume it is covered by e_u, then it immediately follows by assumption that $v \in S'$. A contradiction. Note that the cover $S = S' \cap V$ has cardinality $|S'| - |E|$.

Now, it is easy to see that $\mathcal{F} := \{\{\{u\}, \{v\}\} : \{u, v\} \in E'\}$ is an instance which is 1-simple, $F_{\max} = 2$ and 2-dim convex. The latter follows from the fact that we do not have edges in V and edges $\{e_u, e_v\}$ only. Thus, all sets are convex using an arbitrary ordering on V and an arbitrary ordering on the set of edge-vertices $V' \setminus V$ as long as e_v, e_u follow one after another for all $e \in E$. Figure 6 shows the construction and indicates the partition of elements into V and $V' \setminus V$. □

In the following, we outline a polynomial-time approximation algorithm for 2- DIM MHSIB where the family $\mathcal{F}$ is 1-simple and $|F| \leq 2$ for all $F \in \mathcal{F}$. Since $F_{\max} \leq 2$ in this setting, there already exists a 2-factor approximation algorithm [1]. We improve the approximation guarantee slightly by using techniques from approximation algorithms for the vertex cover problem.

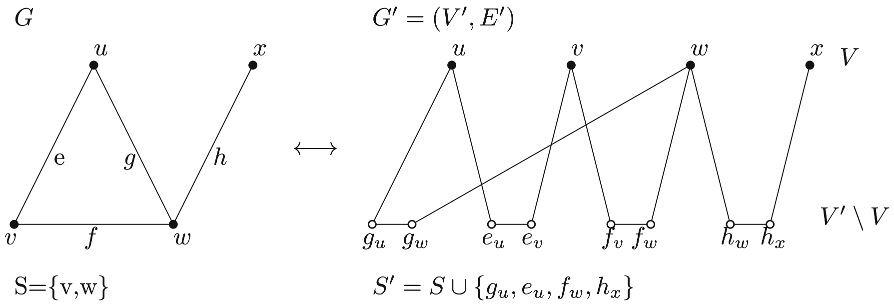

Fig. 6 Construction of G' from G and corresponding vertex covers S and S' in G and G', respectively

Theorem 5 *Let* $(N_1 \dot\cup N_2, \mathcal{F})$ *be an instance of* 2- DIM MHSIB*. If* $\mathcal{F}$ *is* 1*-simple and* $F_{\max} = 2$*, then there exists a polynomial-time* $\left(2 - \frac{1}{k+1}\right)$*-factor approximation algorithm with*

$$k := \max\Big\{|i - j| : \{\{i\}, \{j\}\} \in \mathcal{F} \text{ with either } i, j \in N_1 \text{ or } i, j \in N_2\Big\}.$$

Proof The proof of the theorem is based on a result by Hochbaum [12] for which we are going to convert our instance of 2- DIM MHSIB into a vertex cover instance. The theorem is stated below.

Theorem 6 (Hochbaum [12]) *Let G be a k-colorable graph. There exists a polynomial-time algorithm that finds a vertex cover with a size of at most* $\left(2 - \frac{2}{k}\right)$ *times the size of an optimal vertex cover.*

Given a 1-simple instance $(N_1 \dot\cup N_2, \mathcal{F})$ of 2- DIM MHSIB with $F_{\max} = 2$, we construct a graph $G = (V, E)$ with $V := N_1 \dot\cup N_2$ and $E := \{\{i, j\} : \{\{i\}, \{j\}\} \in \mathcal{F}\}$. Then, an optimal solution of 2- DIM MHSIB corresponds to a minimum vertex cover in the graph G and vice versa.

Consider the subgraphs of $G[N_1]$ and $G[N_2]$ and observe that by the construction of G and the definition of k it holds that $\Delta(G[N_1]) \leq k$ and $\Delta(G[N_2]) \leq k$. Therefore, by Brooks' Theorem both induced subgraphs admit a $(k+1)$-coloring. These colorings can be extended to a $(2k + 2)$-coloring of G.

Applying Theorem 6, a chromatic number of at most $2k+2$ immediately guarantees a polynomial-time algorithm that computes a vertex cover whose size is at most $(2 - \frac{1}{k+1})$ times the size of an optimal vertex cover. □

Finally, we present a first polynomial-time algorithm for a special case of 2- DIM MHSIB. It is based on a decomposition of the problem.

Theorem 7 *Let* $(N_1 \dot\cup N_2, \mathcal{F})$ *be a 2-dim-convex instance of* 2- DIM MHSIB *with* $F_{\max} \leq 2$*. The problem is solvable in polynomial time if for all* $F \in \mathcal{F}$ *and* $U, U' \in F$*, the symmetric difference only contains two elements, i.e.* $|U \triangle U'| = 2$ *and either*

(i) those elements belong to N_1 *or* N_2*, i.e.* $U \triangle U' \subset N_1$ *or* $U \triangle U' \subset N_2$*; or*
(ii) the symmetric difference always contains one element from N_1 *and one element from* N_2*, i.e.* $U \triangle U' \cap N_1 \neq \emptyset$ *and* $U \triangle U' \cap N_2 \neq \emptyset$*.*

Proof First, note that for any $F = \{U, U'\}$ all elements in $U \cap U'$ must be contained in any hitting set of bundles. Let $(N_1' \dot\cup N_2', \mathcal{F}')$ be the instance after removing all elements in the respective intersections. Note that there is a one-to-one correspondence between a minimum hitting set of bundles of $\mathcal{F}'$ and a minimum hitting set of bundles of $\mathcal{F}$ by adding back or removing the aforementioned elements of the intersections. Due to 2-dim-convexity and a 2-dimensional interval family $\mathcal{F}$, we have that the reduced family $\mathcal{F}'$ on the reduced set of elements $N_1' \dot\cup N_2'$ is again 2-dim-convex and 2-dimensional interval. Furthermore, w.l.o.g., $|F| = 2$ for all $F \in \mathcal{F}'$. (If $|F| = 1$ all elements in $U \in F$ are contained in any hitting set of bundles and can be removed to reduce the instance even further.) Additionally, by assumption the family $\mathcal{F}'$ is 1-simple.

Observe, the graph on $N_1' \dot\cup N_2'$ with edges corresponding to the sets in $\mathcal{F}'$ is bipartite. This follows from the fact that on the one hand, if $U \triangle U' \subset N_1$ or $U \triangle U' \subset N_2$ for all $\{U, U'\} \in \mathcal{F}$, due to convexity, the graph consists of a collection of paths since each element in $N_1' \dot\cup N_2'$ can neighbor at most two others and cycles cannot occur. On the other hand, if $U \triangle U' \cap N_1 \neq \emptyset$ and $U \triangle U' \cap N_2 \neq \emptyset$ for all $\{U, U'\} \in \mathcal{F}$, edges only occur between vertices representing the sets N_1' and N_2', but not within the sets N_1' and N_2', respectively. Now, any vertex cover in the constructed graph corresponds to a solution of 2- DIM MHSIB and vice versa. Since the graph is bipartite, we can find a vertex cover in polynomial time, which follows from König's theorem and any polynomial time maximum matching algorithm [17]. Finally, a solution to $\mathcal{F}'$ can easily be lifted to a solution to $\mathcal{F}$. □

4 Conclusion

In this paper, we study the minimum hitting set of bundles problem on interval and 2-dimensional interval bundles. In the following, we provide an outlook on future research directions. In Theorem 1, we show that MHSIB remains NP-hard if the family $\mathcal{F}$ is convex and a-simple, where a is some function in n. For convex, a-simple families with constant a the complexity remains open.

The following application of MHSB gives rise to several interesting questions for future research. The universe is given by a collection of courses. Every set represents a student and the bundles of the set correspond to a feasible subset of courses, the student should take. In this case, the bundle constraints do not only express time overlaps, but can also be used to encode that some courses might only be taken in combination (e.g. lectures and corresponding exercise classes, lab experiments, etc.). In this setting, the introduction of an upper bound on the number of students per course seems natural. This corresponds to introducing an upper bound on the number of sets that are hit by a bundle, containing a specific element. If the bundles are interval, this problem is equivalent to active time minimization with capacity B.

Acknowledgements This work has been supported in part by the Alexander von Humboldt Foundation with funds from the German Federal Ministry of Education and Research (BMBF) and by the Deutsche Forschungsgemeinschaft (DFG), GRK 2201. We want to thank Andreas S. Schulz for his valuable comments and helpful discussions, and the anonymous reviewers for their useful and constructive comments, in particular for pointing out the connection to [20].

Funding Open Access funding enabled and organized by Projekt DEAL.

References

1. Angel, E., Bampis, E., Gourvès, L.: On the minimum hitting set of bundles problem. Theor. Comput. Sci. **410**(45), 4534–4542 (2009)
2. Bansal, N., Khot, S.: Inapproximability of hypergraph vertex cover and applications to scheduling problems. In: Automata. Languages and Programming, pp. 250–261. Springer, Berlin (2010)
3. Bertsimas, D., Teo, C., Vohra, R.: On dependent randomized rounding algorithms. Oper. Res. Lett. **24**(3), 105–114 (1999)
4. Chang, J., Gabow, H.N., Khuller, S.: A model for minimizing active processor time. In: European Symposium on Algorithms, pp. 289–300. Springer, Berlin (2012)
5. Chang, J., Khuller, S., Mukherjee, K.: Lp rounding and combinatorial algorithms for minimizing active and busy time. J. Sched. **20**(6), 657–680 (2017)
6. Chekuri, C., Kumar, A.: Maximum coverage problem with group budget constraints and applications. In: Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, pp. 72–83. Springer (2004)
7. Damaschke, P.: Parameterizations of hitting set of bundles and inverse scope. J. Comb. Optim. **29**(4), 847–858 (2015)
8. Dinur, I., Steurer, D.: Analytical approach to parallel repetition. In: Proceedings of the Forty-sixth Annual ACM Symposium on Theory of Computing, STOC '14, pp. 624–633. ACM, New York (2014)
9. Eskandarzadeh, S., Kalinowski, T., Waterer, H.: Maintenance scheduling in a railway corricdor. arXiv preprint arXiv:1910.10348 (2019)
10. Fang, X., Gao, H., Li, J., Li, Y.: Application-aware data collection in wireless sensor networks. In: 2013 Proceedings IEEE INFOCOM, pp. 1645–1653 (2013)
11. Fong, K.C.K., Li, M., Li, Y., Poon, S.H., Wu, W., Zhao, Y.: Scheduling tasks to minimize active time on a processor with unlimited capacity. In: Theory and Applications of Models of Computation, pp. 247–259. Springer (2017)
12. Hochbaum, D.S.: Efficient bounds for the stable set, vertex cover and set packing problems. Discrete Appl. Math. **6**(3), 243–254 (1983)
13. Iwata, S., Nagano, K.: Submodular function minimization under covering constraints. In: 2009 50th Annual IEEE Symposium on Foundations of Computer Science, pp. 671–680 (2009)
14. Karp, R.M.: Reducibility among combinatorial problems, pp. 85–103. Springer (1972)
15. Khandekar, R., Schieber, B., Shachnai, H., Tamir, T.: Minimizing busy time in multiple machine real-time scheduling. In: IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2010). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik (2010)
16. Koehler, F., Khuller, S.: Busy time scheduling on a bounded number of machines. In: Workshop on Algorithms and Data Structures, pp. 521–532. Springer (2017)
17. Korte, B.H., Vygen, J., Korte, B., Vygen, J.: Combinatorial Optimization, vol. 1. Springer, Berlin (2011)
18. Sellis, T.K.: Multiple-query optimization. ACM Trans. Database Syst. **13**(1), 23–52 (1988)
19. Vazirani, V.V.: Approximation Algorithms. Springer, New York (2013)
20. Wan, P.J., Du, D.Z., Pardalos, P., Wu, W.: Greedy approximations for minimum submodular cover with submodular cost. Comput. Optim. Appl. **45**(2), 463–474 (2010)