

Approximation Algorithms for Demand Strip Packing

Waldo Gálvez  



Technical University of Munich, Germany

Fabrizio Grandoni 

IDSIA, USI-SUPSI, Switzerland

Afrouz Jabal Ameli 

IDSIA, USI-SUPSI, Switzerland

Kamyar Khodamoradi  

University of Würzburg, Germany

Abstract

In the Demand Strip Packing problem (DSP), we are given a time interval and a collection of tasks, each characterized by a processing time and a demand for a given resource (such as electricity, computational power, etc.). A feasible solution consists of a schedule of the tasks within the mentioned time interval. Our goal is to minimize the peak resource consumption, i.e. the maximum total demand of tasks executed at any point in time.

It is known that DSP is NP-hard to approximate below a factor $3/2$, and standard techniques for related problems imply a (polynomial-time) 2-approximation. Our main result is a $(5/3 + \varepsilon)$ -approximation algorithm for any constant $\varepsilon > 0$. We also achieve best-possible approximation factors for some relevant special cases.

2012 ACM Subject Classification Theory of computation \rightarrow Packing and covering problems; Theory of computation \rightarrow Scheduling algorithms

Keywords and phrases Strip Packing, Two-Dimensional Packing, Approximation Algorithms

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2021.20

Related Version A full version of the paper is hosted on arXiv.

Full Version: <https://arxiv.org/pdf/2105.08577.pdf> [17]

Funding *Waldo Gálvez:* Supported by the European Research Council, Grant Agreement No. 691672, project APEG.

Fabrizio Grandoni: Partially supported by the SNF Excellence Grant 200020B_182865.

Afrouz Jabal Ameli: Partially supported by the SNF Excellence Grant 200020B_182865.

Kamyar Khodamoradi: Partially supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project number 399223600. This project was carried out in part when the author was a postdoctoral researcher at IDSIA, USI-SUPSI, Switzerland.



© Waldo Gálvez, Fabrizio Grandoni, Afrouz Jabal Ameli, and Kamyar Khodamoradi; licensed under Creative Commons License CC-BY 4.0

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2021).

Editors: Mary Wootters and Laura Sanità; Article No. 20; pp. 20:1–20:23



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

Consider the following scenario: we are given a time interval and a collection of tasks, where each task is characterized by a processing time (no longer than the time interval) and a demand for a given resource. A feasible solution consists of a schedule of all the tasks within the mentioned time interval, and our goal is to minimize the peak resource consumption, i.e. the maximum total demand of tasks scheduled at any point in time. It is easy to imagine concrete applications of this scenario; for example, the considered resource might be electricity, bandwidth along a communication channel, or computational power.

The above scenario can be naturally formalized via the following DEMAND STRIP PACKING problem (DSP). We interpret the time interval as a path graph $G = (V, E)$ with W edges, where each edge is interpreted as a time slot where we can start to process a task. Let $\mathcal{I} = \{1, \dots, n\}$ be the set of tasks, where task i has integer *processing time* (or *width*) $w(i) \in [1, W]$ and integer *demand* (or *height*) $h(i) \geq 0$. A feasible solution (or *schedule* of the tasks) consists of a subpath $P(i)$ of G for each $i \in \mathcal{I}$ containing precisely $w(i)$ edges. Our goal is to minimize the peak resource consumption (or simply *peak*) which is defined as

$$\max_{e \in E} \sum_{i \in \mathcal{I}: e \in P(i)} h(i).$$

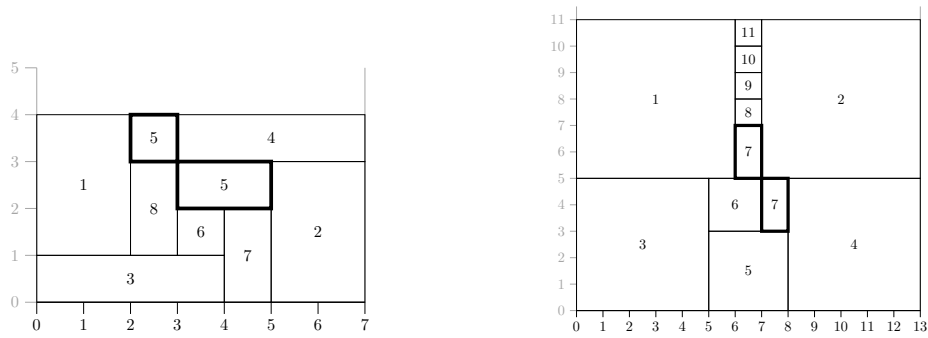
A problem closely related to DSP is the GEOMETRIC STRIP PACKING problem (GSP)¹, which can be interpreted as a variant of DSP with an extra geometric packing constraint. Here we are given an axis-aligned half-strip of integer width W (and unbounded height) and a collection of open rectangles (or tasks), where each rectangle i has integer width $w(i) \in [1, W]$ and integer height $h(i) \geq 0$. Our goal is to find an axis-aligned non-overlapping packing of all the rectangles within the strip that minimizes the peak height, i.e. the maximum height spanned by any rectangle. Notice that one can reinterpret DSP as a variant of GSP, where the processing time and demand of each task correspond to the width and height of a rectangle, resp. (this also motivated our notation). A critical difference w.r.t. GSP however is that DSP does not require to pack such rectangles geometrically².

Obviously, a feasible solution to GSP induces a feasible solution to DSP of no larger peak. The converse is however not true (see Figure 1), and consequently it makes sense to design algorithms specifically for DSP. We remark that there are applications that are better formalized by GSP than by DSP. In particular, this happens when each task requires a contiguous and fixed portion of the considered resource. For example, we might need to allocate consecutive frequencies or memory locations to each task: changing this allocation over time might be problematic. Another natural application of GSP is cutting rectangular pieces from a roll of some raw material (e.g., paper, metal, or leather). However, for other applications, the geometric constraint in GSP does not seem to be necessary, and hence it makes sense to drop it (i.e., to rather consider DSP): this might lead to better solutions, possibly via simpler and/or more efficient algorithms. Consider for example the minimization of the peak energy consumption in smart-grids [31, 44, 39].

A straightforward reduction to the NP-complete PARTITION problem (similar to the one known for GSP, see also [43]) shows that DSP is NP-hard to approximate below a factor

¹ GSP is usually simply called STRIP PACKING in the literature. We added the word “geometric” to better highlight the differences between the two problems.

² Or, equivalently, we can split such rectangles into unit-width vertical slices, and then pack such slices geometrically so that slices of the same rectangle appear consecutively in a horizontal sense.



(a) DSP solution of peak 4 whose corresponding optimal GSP solution has peak 5.

(b) SQUARE-DSP solution of peak 11 whose corresponding optimal GSP solution has peak at least 12.

■ **Figure 1** Gap instances between DSP and GSP.

$3/2$. Constant approximation algorithms for DSP are given in [43, 45]. However, a better 2-approximation can be obtained by applying an algorithm by Steinberg [42] which was developed for GSP: the reason is that Steinberg uses area-based lower bounds that extend directly from GSP to DSP.

1.1 Our Results and Techniques

Our main result is as follows³.

► **Theorem 1.** *For any constant $\varepsilon > 0$, there is a polynomial-time deterministic $(5/3 + \varepsilon)$ -approximation algorithm for DSP.*

The above approximation ratio matches the best-known result for GSP from Harren et al. [23], achieved using dynamic programming based techniques to place almost all the rectangles except for a set of very small total area, followed by a careful and quite involved case distinction to pack these remaining rectangles. However, we remark that our algorithm is entirely different, and in particular it does not compute a geometric packing of tasks/rectangles. Furthermore, our analysis is substantially simpler. Notice that the result in [23] does *not* imply a $(5/3 + \varepsilon)$ -approximation for DSP since some lower bounds used in their proofs do not hold necessarily for DSP.

We also achieve improved approximation algorithms for relevant special cases of DSP. We obtain a PTAS for the special case where the demand of each task is much lower than the optimal peak OPT . This captures applications where each job consumes a relatively small amount of the available resource (think about the electricity consumption of large-scale systems such as cities or countries).

► **Theorem 2.** *Given $\varepsilon > 0$ small enough, there exists $\delta > 0$ and a polynomial time algorithm such that, given an instance of DSP with optimal value OPT and consisting solely of tasks having height at most $\delta \cdot OPT$, it computes a $(1 + O(\varepsilon))$ -approximate solution.*

Motivated by the special case of GSP and related problems where all rectangles are squares, we also study the special case of DSP where $h(j) = w(j)$ for all tasks (the SQUARE-DSP problem). The $3/2 - \varepsilon$ hardness of approximation extends to this case (see Appendix A),

³ The same result as in Theorem 1 was achieved independently in [15]; their approach is however substantially different from ours.

and we are still able to show that there is a gap between DSP and GSP (see Figure 1b and Section 5). However, in this case, we are able to provide an optimal $3/2$ -approximation. We defer the proof of the following theorem to the full version of the paper [17].

► **Theorem 3.** *There is a deterministic polynomial-time $3/2$ -approximation for SQUARE-DSP.*

At a high level, our approach is based on a classification of tasks into groups depending on their heights and widths. We carefully schedule some groups first, so that their *demand profile* has a convenient structure. Here, by demand profile we simply mean the total demand of the already scheduled tasks over each edge. The structure of the demand profile allows us to pack the remaining groups (intuitively, on top of such profile) in a convenient way. We critically exploit the fact that, differently from GSP, we only care about the total demand on each edge. This allows us to adapt techniques from Bin Packing or Makespan Minimization (see Lemmas 6 and 12).

1.2 Related Work

GSP generalizes famous problems such as Makespan Minimization on identical machines [12] (here all the rectangles have width 1 and W corresponds to the number of processors) or Bin Packing [13] (here all the rectangles have height 1 and the height of the solution corresponds to the number of bins). Consequently, it is known that for any $\varepsilon > 0$, there is no $(3/2 - \varepsilon)$ -approximation for the problem unless $P=NP$. The first non-trivial approximation algorithm for GSP, with an approximation ratio of 3, was given by Baker, Coffman, and Rivest [5]. After a series of very technical and involved refinements [14, 41, 40, 42, 24], the current best approximation factor for the problem is $(5/3 + \varepsilon)$ due to Harren et al. [23]. GSP has been also studied in the pseudopolynomial setting, i.e., when $W = n^{O(1)}$ [30, 38, 1, 20, 25, 28, 27] and in the asymptotic setting, i.e. when the optimal value is assumed to be large [32, 29]. In both cases, approximation algorithms and almost matching lower bounds have been developed.

There is a very rich line of research on generalizations and variants of DSP such as online versions [34, 35], tasks with availability constraints or time windows [45, 44, 31], a mixture of preemptable and non-preemptable tasks [39] or generalized cost functions based on the demand at each edge [10, 35]. The variant of DSP with the extra feature of interrupting the tasks is known as STRIP PACKING WITH SLICING, for which there exists an FPTAS [3]; on the other hand, the case of DSP is still hard to approximate by a factor better than $3/2$ as noted by Tang et al. [43].

Another problem closely related to DSP is PARALLEL JOB SCHEDULING. Here we are given a set of jobs and m machines, where each job is characterized by a processing time and a number of machines where the job must be processed simultaneously (these machines do not need to be contiguous), and the goal is to minimize the makespan. The same hardness of approximation applies in this case, but interestingly an almost tight $(3/2 + \varepsilon)$ -approximation algorithm has been developed [26] and also a pseudopolynomial $(1 + \varepsilon)$ -approximation is known [30]. See [16] for a comprehensive survey on the problem and its many variants.

It is also worth mentioning another case where the distinction between geometric and demand-based packing plays a substantial role: the UNSPLITTABLE FLOW ON A PATH problem (UFP) [4, 22, 7, 21] and the STORAGE ALLOCATION problem (SAP) [36, 37]. In both problems, we are given a path graph with edge capacities, and tasks specified by a subpath, a demand (or height), and a profit. In both problems, the goal is to select a maximum profit subset of tasks that can be *packed* while respecting edge capacities. For UFP, analogously to DSP, we require that the total demand of the selected tasks on each

edge e is at most the capacity of e . For SAP, analogously to GSP, we interpret each task as a rectangle (with the width given by its number of edges) and, intuitively, we need to pack such rectangles non-overlappingly below the capacity profile. Notice that, differently from DSP and GSP, here the path associated with each task is fixed in the input. Furthermore, not all the tasks need to be packed.

Finally, in the DYNAMIC STORAGE ALLOCATION problem (DSA) the setting is analogous to SAP but, similarly to GSP, we are asked for an embedding of all the rectangles minimizing the peak height, i.e. the maximum height reached by any rectangle (in particular, there are no edge capacities). Notice that in DSA a lower bound is provided by the peak demand, i.e. the maximum over the edges e of the sum of the heights of rectangles whose path uses e . Buchsbaum et al. [9] studied in detail the relation between the optimal peak height and the peak demand, providing examples where these values differ by a constant factor. The authors also present a $(2 + \varepsilon)$ -approximation for DSA that provides guarantees even when compared with the peak demand.

1.3 Organization

We start by introducing in Section 2 some useful definitions and known results. As a warm-up, in Section 3 we present a very simple 2-approximation for DSP that allows us to illustrate part of our ideas. Then in Section 4 we present our main result, namely a $(5/3 + \varepsilon)$ -approximation for DSP. Finally, in Section 5 we provide details about the gap instances in Figure 1. The results for special cases of DSP (Theorems 2 and 3) can be found in Appendix B and the full version of this paper [17] respectively.

2 Preliminaries

Let e_1, \dots, e_W be the edges of G from left to right. Recall that a feasible solution or schedule $P(\cdot)$ specifies a subpath $P(i)$ of G of length $w(i)$ for each task i . Sometimes it is convenient to consider a *partial* schedule $P(\cdot)$ which specified the path of a subset \mathcal{I}' of tasks only (it is convenient to consider $P(i)$ as an *empty path* for the remaining tasks). We call this a schedule of \mathcal{I}' .

Let us define, for a given subset \mathcal{I}' of tasks, $h_{\max}(\mathcal{I}') := \max_{i \in \mathcal{I}'} h(i)$ and $h(\mathcal{I}') := \sum_{i \in \mathcal{I}'} h(i)$. We define analogously $w_{\max}(\mathcal{I}')$ and $w(\mathcal{I}')$ w.r.t. widths. Let also $a(\mathcal{I}') := \sum_{i \in \mathcal{I}'} a(i)$, where $a(i) := h(i) \cdot w(i)$ corresponds to the *area* of task i . We will start by showing a couple of simple lower bounds for the optimal peak OPT that will be used extensively along this work.

► **Proposition 4.** $OPT \geq \max\{h_{\max}(\mathcal{I}), \sum_{i \in \mathcal{I}: w(i) > W/2} h(i), a(\mathcal{I})/W\}$.

Proof. Since the total demand of any edge used by the task of largest height is at least $h_{\max}(\mathcal{I})$, it holds that $OPT \geq h_{\max}(\mathcal{I})$. Also notice that, in any scheduling, the tasks of width larger than $W/2$ use the edge $e_{\lceil W/2 \rceil}$, being then the total demand of this edge (and consequently OPT) at least $\sum_{i \in \mathcal{I}: w(i) > W/2} h(i)$. Finally, the last bound follows from an averaging argument and the fact that the sum over the edges of the total demand on each edge is equal to $a(\mathcal{I})$. ◀

2.1 Demand Profile and Left-Pushing

Consider a schedule $P(\cdot)$ of $\mathcal{I}' \subseteq \mathcal{I}$. We define the *demand profile* $h(P)$ of $P(\cdot)$ as the vector that stores for each edge e the total demand $\sum_{i \in \mathcal{I}': e \in P(i)} h(i)$ of the tasks whose path

contains e (if the path of i is not specified, then i does not contribute to the demand profile). Since W can be exponential in n , we need to store the demand profile in a more efficient way. This can be done by noticing that the number of times the total demand changes from an edge to the next one is at most $2n$ (when a task starts or finishes). Hence we just need to store the edges where the demand profile changes value and the corresponding demand. In particular, we can efficiently store the demand profile. Furthermore, we can efficiently update it, e.g., when augmenting an existing schedule by specifying the path $P(i)$ of one more task i , or when we modify the value of some $P(i)$ by *shifting* tasks as we will discuss later.

Given a schedule $P(\cdot)$ of $\mathcal{I}' \subseteq \mathcal{I}$ and $i \in \mathcal{I}'$, a *left-shifting* of i in $P(\cdot)$ is the operation of replacing $P(i)$ with the path $P'(i)$ of length $w(i)$ that starts one edge to the left of $P(i)$. Clearly, this operation is allowed only if $P(i)$ does not start at the leftmost edge of G . Consider a schedule $P(\cdot)$ with peak π , and let $\pi' \geq \pi$. A π' -*left-pushing* of $P(\cdot)$ is the operation of iteratively performing left-shiftings in any order until it is not possible to continue while guaranteeing that the peak is always at most π' . We will critically use left-pushings in our algorithms. Notice that a left-pushing can be computed in polynomial time (see Appendix B for some more details).

Intuitively, left-pushing accumulates the demand over the first edges while inducing a non-increasing demand profile to the right. For a node t^* of the path and a value $Q \geq 0$, we will say that a (possibly partial) schedule $P(\cdot)$ is (Q, t^*) -sorted if the corresponding demand on the edges to the left of t^* is at least Q and on the edges to the right of t^* the demand profile is non-increasing (see Figure 2); if t^* is the leftmost node we just say that the schedule is sorted. Our algorithms will first schedule some tasks and then perform a left-pushing. After that, it will be possible to schedule the remaining tasks in a convenient way thanks to the properties of the resulting demand profile.

2.2 Container-based Scheduling

Similar to recent work on related rectangle packing problems (e.g., [18, 6]), we will exploit a *container-based* scheduling approach. A *container* C can be interpreted as an artificial task, with its own width $w(C)$ (i.e. a number of edges) and height $h(C)$. Furthermore, it is classified as *vertical* or *horizontal*, with a meaning which is explained later. The containers are scheduled as usual tasks in a DSP instance (in particular by defining a path $P(C)$ for each container C), with the goal of minimizing the peak π . We also define a packing of tasks into containers C respecting the following constraints: if C is vertical, the tasks $\mathcal{I}(C)$ packed into C must have height at most $h(C)$ and total width at most $w(C)$; if C is horizontal, tasks $\mathcal{I}(C)$ must have width at most $w(C)$ and total height at most $h(C)$. Intuitively, the tasks packed into a vertical (resp., horizontal) container induce a geometric packing of the rectangles associated with each task into the rectangle corresponding to the container, where the task rectangles are packed non-overlappingly one next to the other (resp., one on top of the other). Any such packing and scheduling of containers naturally induces a schedule of the tasks: if C is horizontal, tasks $\mathcal{I}(C)$ are all scheduled starting on the leftmost edge of $P(C)$. Otherwise, tasks $\mathcal{I}(C)$ are scheduled one after the other starting at the leftmost edge of $P(C)$. It is hopefully clear to the reader that the demand profile of such a schedule of the tasks is dominated by the demand profile of the containers' schedule. In particular, if the latter has peak π , then the corresponding schedule of the tasks has a no larger peak.

The general strategy is then as follows: we first show that there exists a convenient packing of tasks into a constant number of containers and that there exists a scheduling $P^*(\cdot)$ of these containers with a small peak π . We also require that these containers are *guessable*,

meaning that we can guess their sizes by exploring a polynomial number of options. Once we guessed the correct set of containers, a π -left-pushing of $P^*(\cdot)$ can be computed by brute force (since they are constantly many tasks). Finally, we pack tasks into containers, inducing a schedule of the tasks with peak π .

This final step can be performed (almost completely) via a reduction to the GENERALIZED ASSIGNMENT problem (GAP). Recall that in GAP we are given a set of k bins, where each bin j has an associated capacity $C_j \geq 0$, and a set of n items. For each item i and bin j , the input specifies a size $s_{ij} \geq 0$ and a profit $p_{ij} \geq 0$ of item i w.r.t. bin j . A feasible solution assigns each item to some bin so that the total size of the items assigned to each bin j is at most C_j . Our goal is to maximize the total profit associated with this assignment. GAP admits a PTAS in the case of a constant number of bins (see Section E.2 in [19]).

► **Lemma 5.** *For any constant $\varepsilon' > 0$, given a set of tasks \mathcal{I}' that can be packed into a given set of containers of constant cardinality, there is a polynomial-time algorithm to pack $\mathcal{I}'' \subseteq \mathcal{I}'$ with $a(\mathcal{I}'') \geq (1 - \varepsilon')a(\mathcal{I}')$ into the mentioned containers.*

Proof. We define a GAP instance as follows: we create one bin per container, where the capacity of the bin is equal to the width of the container if it is vertical or the height of the container if it is horizontal. For each task i we define an item that has uniform profit equal to its area $a(i)$ over all the bins. Given a task i and a vertical (resp., horizontal) container j , the size s_{ij} of i into bin j is set to $w(i)$ (resp., $h(i)$) if task i can be packed into container j according to the mentioned rules. Otherwise we set $s_{ij} = +\infty$. The claim follows by applying the aforementioned PTAS for GAP with parameter ε' . ◀

Notice that the above lemma allows us to pack all the tasks but a subset of small total area, hence we need to schedule somehow such *leftover* tasks. This is not necessarily a trivial task; indeed, such tasks, though of small area, might have large height, and hence scheduling them on top of the rest might substantially increase the peak. To circumvent this issue we will identify special containers reserved for tasks of large height where we will be able to pack all such tasks with no leftovers. We will then apply the PTAS from Lemma 5 only to the remaining tasks and containers.

3 A Simple 2-Approximation for DSP

In order to introduce part of our ideas, in this section, we present a simple 2-approximation for DSP. As mentioned before, a 2-approximation can also be achieved via Steinberg's algorithm [42], however, that algorithm is substantially more complex (not surprisingly since it computes a *geometric packing* of tasks interpreted as rectangles like in GSP).

The following lemma exploits a modification of Next-Fit-Decreasing [13], the well-known approximation algorithm for Bin Packing.

► **Lemma 6.** *Let $P(\cdot)$ be a sorted schedule of $\mathcal{I}' \subseteq \mathcal{I}$ with peak at most π . For $\mathcal{I}'' := \mathcal{I} \setminus \mathcal{I}'$, assume that:*

- $\pi \geq h_{\max}(\mathcal{I}'') + \max\{a(\mathcal{I})/W, h_{\max}(\mathcal{I}'')\}$,
- $w_{\max}(\mathcal{I}'') \leq W/2$, and
- $(W - w_{\max}(\mathcal{I}''))(\pi - h_{\max}(\mathcal{I}'')) + w_{\max}(\mathcal{I}'') \cdot h_{\max}(\mathcal{I}'') \geq a(\mathcal{I})$.

Then it is possible to compute in polynomial time a schedule of \mathcal{I} having peak at most π .

Proof. By slightly abusing notation, we will next use an edge label e also to denote the position i of e in the sequence e_1, \dots, e_W of edges from left to right. We do not modify the schedule of \mathcal{I}' and schedule the remaining tasks \mathcal{I}'' as follows. Let us fix an arbitrary order

for tasks in \mathcal{I}'' , and let us initially define e^{check} to be the leftmost edge. We scan completely the list of tasks \mathcal{I}'' and, if the current task i can be scheduled starting on edge e^{check} while maintaining a peak of at most π , we do that and remove i from \mathcal{I}'' ; otherwise, we keep i in \mathcal{I}'' and try with the next task. Once we consider the final task, we update e^{check} to be the leftmost edge to the right of the current e^{check} whose demand is different from the demand of the current e^{check} . We iterate the procedure on the new e^{check} until all tasks are scheduled or we identify a task i which cannot be scheduled.

First, notice that we update e^{check} at most $|\mathcal{I}'|$ times, and after each time, we iterate through at most $|\mathcal{I}''|$ tasks, so the running time is polynomial in the size of the input. It is also not difficult to see that with this procedure, the demand profile from e^{check} to its right is always non-increasing (restricted to these edges, scheduling a task is equivalent to summing up two non-decreasing profiles), and none of the remaining tasks can fit in any one of the edges to the left of e^{check} (as we actually tried to place them there but it was not possible). Notice also that, if this procedure manages to schedule all the tasks, then the claimed peak is automatically achieved. So we will assume by contradiction that this is not the case.

Let i be a task that could not be scheduled. This could only happen due to i being too wide for the current edge e^{check} where it should be scheduled (and hence for any subsequent edge). This implies that $e^{check} > W - w(i)$ and hence there are more than $W - w(i)$ edges having demand larger than $\pi - h(i)$. Thus the total area of the scheduled tasks plus task i is strictly larger than

$$A(h(i), w(i)) := (W - w(i)) \cdot (\pi - h(i)) + h(i) \cdot w(i).$$

This expression is decreasing both as a function of $h(i)$ and as a function of $w(i)$. Indeed,

$$\frac{\partial}{\partial h(i)} A(h(i), w(i)) = 2w(i) - W \leq 0, \text{ and } \frac{\partial}{\partial w(i)} A(h(i), w(i)) = 2h(i) - \pi \leq 0,$$

where we used the fact that, by assumption, $w(i) \leq \frac{W}{2}$ and $\pi \geq 2h_{\max}(\mathcal{I}'') \geq 2h(i)$. We conclude that

$$A(h(i), w(i)) \geq A(h_{\max}(\mathcal{I}''), w_{\max}(\mathcal{I}'')) = (W - w_{\max}(\mathcal{I}''))(\pi - h_{\max}(\mathcal{I}'')) + w_{\max}(\mathcal{I}'') \cdot h_{\max}(\mathcal{I}'') \geq a(\mathcal{I}),$$

where in the last inequality we used the third assumption. This is a contradiction since a subset of tasks would have area strictly larger than the total area $a(\mathcal{I})$. ◀

We are now ready to provide a simple 2-approximation.

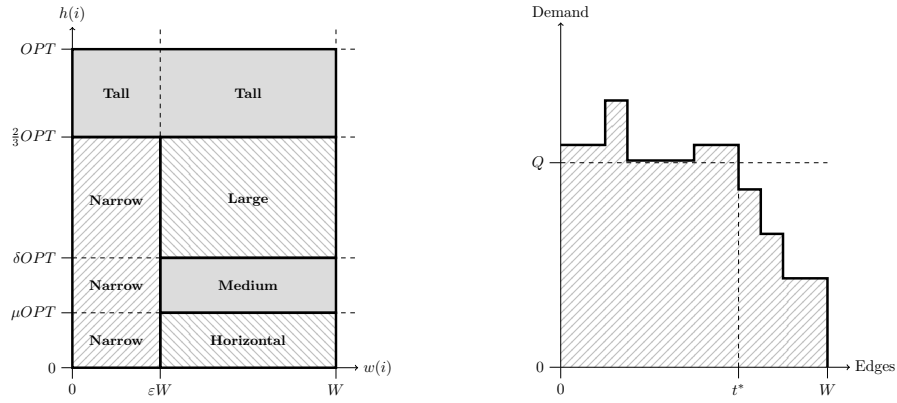
► **Corollary 7.** *There exists a deterministic 2-approximation for DSP.*

Proof. Let \mathcal{I} be an instance of DSP. We will first schedule the tasks \mathcal{I}' having width larger than $W/2$ starting on the leftmost edge. Let $\mathcal{I}'' := \mathcal{I} \setminus \mathcal{I}'$. This partial schedule is sorted and has peak $\sum_{i \in \mathcal{I}''} h(i) \leq M := \max\{h_{\max}(\mathcal{I}), \sum_{i \in \mathcal{I}'} h(i), a(\mathcal{I})/W\}$. Recall that, by Proposition 4, $M \leq OPT$. Define $\pi = 2M$, and observe that

$$(\pi - h_{\max}(\mathcal{I}''))(W - w_{\max}(\mathcal{I}'')) + h_{\max}(\mathcal{I}'') \cdot w_{\max}(\mathcal{I}'') \geq M \cdot (W/2) + M \cdot (W/2) = M \cdot W \geq a(\mathcal{I}).$$

Thus we can apply Lemma 6 with parameter $\pi = 2M$. This provides a schedule with peak at most $2M \leq 2OPT$. ◀

In the following sections, we will extend the approach in the above 2-approximation as follows. We will first compute a feasible solution of some given peak that includes all the tasks having height larger than some threshold and width larger than some threshold. Then we will left-push this schedule to add some structure to the demand profile. Finally, we schedule the remaining tasks by means of a generalization of Lemma 6 (Lemma 12) which considers (Q, t^*) -sorted schedules (rather than just sorted ones).



■ **Figure 2** (Left) Classification of the tasks according to Section 4. (Right) Representation of a (Q, t^*) -sorted demand profile.

4 A $(5/3 + \varepsilon)$ -Approximation for DSP

In this section we will prove Theorem 1. In order to attain the claimed result, we will provide first some useful definitions and preprocessing lemmas.

Let us assume that the optimal value OPT is known to the algorithm (this assumption can be dropped by approximately guessing this value, introducing an extra $(1 + \varepsilon)$ factor in the approximation). We start by classifying the tasks in the instance according to their widths and heights (see Figure 2). Let $\mu, \delta, \mu < \delta \leq \varepsilon$, be two constant parameters to be fixed later. We say that a task i is:

- *tall* if $h(i) > \frac{2}{3}OPT$,
- *large* if $h(i) \in (\delta OPT, \frac{2}{3}OPT]$ and $w(i) > \varepsilon W$,
- *horizontal* if $h(i) \leq \mu OPT$ and $w(i) > \varepsilon W$,
- *narrow* if $h(i) \leq \frac{2}{3}OPT$ and $w(i) \leq \varepsilon W$, or
- *medium* if $h(i) \in (\mu OPT, \delta OPT]$ and $w(i) > \varepsilon W$.

We denote by $\mathcal{T}, \mathcal{L}, \mathcal{H}, \mathcal{N}$ and \mathcal{M} the sets of tall, large, horizontal, narrow and medium tasks respectively. Recall that the general idea of our approach is to show that we can schedule almost every task in a constant number of containers, which in turn can be enumerated in polynomial time by brute force. In the rest of this section, we basically show the existence of such a solution for most of the previously described classes. More precisely, we argue that the optimal schedule can be rearranged so as to define constantly many containers to pack all the tasks in $\mathcal{T} \cup \mathcal{L}$ while increasing the peak of the solution by at most an additive $\frac{2}{3}OPT$ term (see Lemma 9). We then show that a subset of the tasks in \mathcal{H} can be scheduled in $\mathcal{O}(1)$ containers such that the left-over from this class has a very small area compared $W \cdot OPT$ (see Lemma 10). Then, it is not difficult to show that the left-over from \mathcal{H} can be placed in a single container with a very small height. We also prove that all the tasks in \mathcal{M} can be packed in a single container of small height by showing how to choose the parameters μ and δ (see Lemma 8). This leads in conclusion to a container-based scheduling for $\mathcal{I} \setminus \mathcal{N}$ of peak $(\frac{5}{3} + 7\varepsilon)OPT$ (see Lemma 11). Finally, we present a method to schedule every task in \mathcal{N} on top of the previous solution without increasing its peak by using a greedy approach (see Lemma 12).

We first deal with the medium tasks. As the following lemma states, it is possible to choose μ and δ in such a way that the two parameters differ by a large factor and the total height of medium tasks is small.

► **Lemma 8.** *Given a polynomial-time computable function $f : (0, 1) \rightarrow (0, 1)$, with $f(x) < x$, and any constant $\varepsilon \in (0, 1)$, we can compute in polynomial time a set Δ of $\frac{2}{\varepsilon^2}$ many positive real numbers upper bounded by ε , such that there is at least one number $\delta \in \Delta$ so that, by choosing $\mu = f(\delta)$, one has $a(\mathcal{M}) \leq \varepsilon^2 \cdot OPT \cdot W$ (hence $h(\mathcal{M}) \leq \varepsilon OPT$).*

Proof. Let $y_1 = \varepsilon$ and, for each $j \in \{1, \dots, |\Delta|\}$, define $y_{j+1} = f(y_j)$. For each $j \leq |\Delta|$, let $I_j = \{i \in \mathcal{I} : h(i) \in [y_{j+1}, y_j]\}$. Note that y_j 's are decreasing since $f(x) < x$. Observe that $I_{j'}$ is disjoint from $I_{j''}$ for every $j' \neq j''$, and the total area of tasks in $\bigcup I_j$ is at most $W \cdot OPT$. Thus, there exists a value \bar{j} such that the total area of the tasks in $I_{\bar{j}}$ is at most $\frac{2OPT \cdot W}{|\Delta|} = \varepsilon^2 \cdot OPT \cdot W$. Choosing $\delta = y_{\bar{j}}$ and $\mu = y_{\bar{j}+1}$ verifies all the conditions of the lemma as in that case $\mathcal{M} \subseteq I_{\bar{j}}$. Notice that, since every task in \mathcal{M} has width at least εW , we have that $h(\mathcal{M}) \leq \varepsilon OPT$. ◀

Function f will be given later. From now on, we will assume that μ and δ are chosen according to Lemma 8. Notice that this implies that $\mu, \delta = O_\varepsilon(1)$. The rest of this section is organized as follows. In Section 4.1 we define a container-based scheduling of $\mathcal{T} \cup \mathcal{L}$. In Section 4.2 we extend this in order to include also \mathcal{H} . In Section 4.3 we schedule the remaining tasks and prove Theorem 1.

4.1 Containers for Tall and Large Tasks

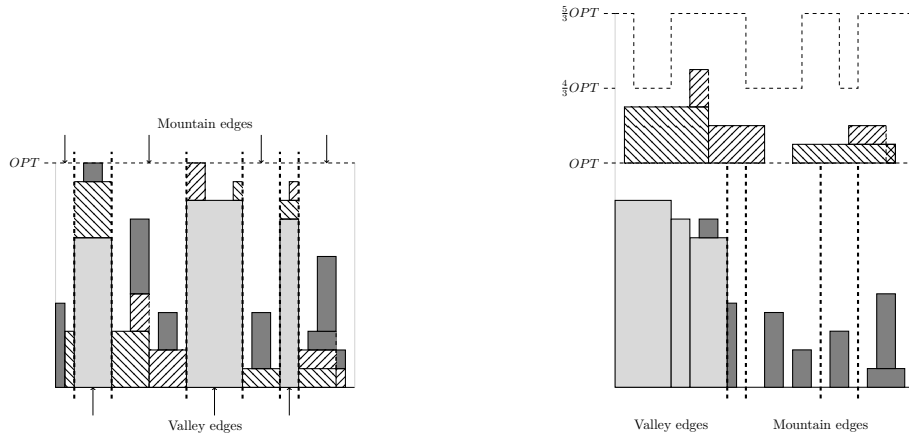
In this section, we define a packing of tall and large tasks into a constant number of guessable containers. This packing can be computed exactly, i.e. with no leftovers (in particular, we will not use Lemma 5 to compute such packing). To that aim, we will exploit the following structural result.

► **Lemma 9.** *Let $P(\cdot)$ be an optimal schedule of \mathcal{I} (hence with peak OPT). There exists a packing $P'(\cdot)$ with peak at most $\frac{5}{3}OPT$ satisfying that all the tall tasks are scheduled one after the other starting on the leftmost edge in non-increasing order of height.*

Proof. Let \mathcal{T} be the tall tasks (having height larger than $\frac{2}{3}OPT$). Notice that the paths of these tasks in $P(\cdot)$ need to be edge disjoint. Let us classify the edges into *valley edges* if some task in \mathcal{T} uses that edge in $P(\cdot)$ and *mountain edges* otherwise (see also Figure 3a). We let \mathcal{I}_{mnt} be the (*mountain*) tasks whose path in $P(\cdot)$ consists solely of mountain edges, \mathcal{I}_{vll} be the (*valley*) tasks whose path in $P(\cdot)$ consists solely of valley edges (notice that this set includes \mathcal{T}), and $\mathcal{I}_{crs} := \mathcal{I} \setminus (\mathcal{I}_{vll} \cup \mathcal{I}_{mnt})$ the remaining (*crossing*) tasks.

We next define a modified partial schedule $P'(\cdot)$ of $\mathcal{I}_{vll} \cup \mathcal{I}_{mnt}$ as follows. Let us reorder the edges of the path (and the tasks accordingly) so that valley edges appear to the left and mountain edges appear to the right in the path (maintaining their relative order). Furthermore, we rearrange the valley edges so that tasks in \mathcal{T} are scheduled from left to right in non-increasing order of height. Observe that by construction \mathcal{I}_{mnt} are scheduled on $W - w(\mathcal{T})$ edges (i.e. the total number of mountain edges). Since we temporarily removed crossing tasks, this induces a feasible packing of $\mathcal{I}_{vll} \cup \mathcal{I}_{mnt}$. The resulting packing $P'(\cdot)$ clearly has a peak of at most OPT .

Consider next the schedule $P(\cdot)$ restricted to \mathcal{I}_{crs} . We claim that this schedule has peak at most $\frac{2}{3}OPT$. Indeed, notice first that the demand of valley edges is at most $\frac{1}{3}OPT$. Consider next a mountain edge e . Let e_ℓ be the rightmost valley edge to the left of e (if any), and define e_r symmetrically to the right of e . Any task in \mathcal{I}_{crs} using e must also use e_ℓ or e_r (or both). Hence the total demand on e is at most the total demand on e_ℓ plus the total demand on e_r , thus at most $\frac{2}{3}OPT$. The claim follows by combining the schedule of \mathcal{I}_{crs} (taken from $P(\cdot)$) with the above schedule $P'(\cdot)$ of $\mathcal{I}_{vll} \cup \mathcal{I}_{mnt}$ (see also Figure 3b). ◀



(a) A scheduling of peak OPT . Light gray rectangles correspond to tasks in \mathcal{T} , dark gray ones represent mountain and valley tasks, and dashed ones the crossing tasks. (b) Structured solution having peak at most $\frac{5}{3}OPT$, where tasks in \mathcal{T} are placed one next to the other, starting at the leftmost edge and sorted non-increasingly by height.

■ **Figure 3** Depiction of the proof of Lemma 9.

We will next assume that tall tasks are scheduled as in the above lemma. By increasing the peak by εOPT (up to $(\frac{5}{3} + \varepsilon)OPT$), one can define a set of $O_\varepsilon(1)$ (tall) containers where such tasks can be packed (respecting the mentioned order and with no leftovers). Consider the demand profile of tall tasks in the considered schedule, and round it to the next multiple of εOPT . Consider the tasks \mathcal{T}_k corresponding to the value $k \cdot \varepsilon OPT$ in the rounded profile. Notice that these tasks are scheduled consecutively along some path P_k . We create a vertical container C_k of height $k \cdot \varepsilon OPT$ and width $|E(P_k)|$, pack \mathcal{T}_k into C_k , and schedule C_k on P_k . Clearly, we need to create at most $1/\varepsilon$ containers. Notice also that the dimensions of these containers can be guessed in polynomial time since there is a constant number of options for the height, and the widths correspond to the total width of a subsequence of tall tasks in the considered ordering by non-increasing height (breaking ties arbitrarily).

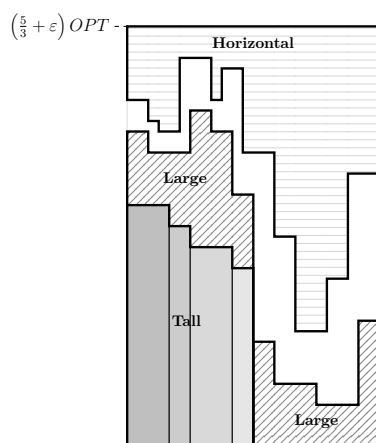
It remains to consider large tasks. Since they are at most $\frac{1}{\varepsilon\delta} = O_\varepsilon(1)$ many, it is sufficient to define a distinct (large) container for each one of them and pack the large tasks accordingly. We schedule the large containers exactly as in the solution guaranteed by Lemma 9. Clearly tall and large containers can be scheduled together with a peak of at most $(\frac{5}{3} + \varepsilon)OPT$ by the above construction.

4.2 Containers for Horizontal Tasks

In this section, we define a packing of horizontal tasks into a constant number of guessable containers. These containers can be scheduled together with the tall and large containers with small enough peaks. This will induce a convenient schedule of non-narrow tasks.

Let us focus on the schedule of tall and large containers with a peak of at most $(\frac{5}{3} + \varepsilon)OPT$ from the previous section. Consider now the demand profile of such container schedule. Since the demand profile of tall containers has at most $1/\varepsilon = O_\varepsilon(1)$ jumps, and the demand profile of large containers has at most $2/(\varepsilon\delta) = O_\varepsilon(1)$ jumps, then the overall demand profile has $O_\varepsilon(1)$ jumps.

Assume next that horizontal tasks are scheduled as in Lemma 9: notice that such tasks can be scheduled with the tall and large containers without increasing the peak. This implies that the demand profile of horizontal tasks is upper bounded (on each coordinate) by the



■ **Figure 4** The demand profile of tall and large tasks in the schedule obtained from Lemma 9 has $O_\varepsilon(1)$ jumps, bounding the profile of (sliced) horizontal tasks (on top).

difference between $(5/3 + \varepsilon)OPT$ and the demand profile of tall and large containers (see Figure 4). Under these conditions, it is possible to build containers for horizontal tasks, using the standard linear grouping technique, as the following lemma shows.

► **Lemma 10.** *Suppose there exists a schedule of \mathcal{H} such that its demand profile is upper bounded (vectorially) by a demand profile D with $O_\varepsilon(1)$ jumps. Then there exists a container packing for \mathcal{H} into $O_\varepsilon(1)$ horizontal guessable containers with demand profile upper bounded by D plus $4\varepsilon OPT$ on each coordinate.*

Proof. Let us assume by now that horizontal tasks are *horizontally sliced*, meaning that each task i , having height $h(i)$ and width $w(i)$, is replaced by $h(i)$ sibling slices, which are tasks of width $w(i)$ and height 1. The schedule of the slices is the same as for the corresponding task. In order to reduce the possible number of distinct slice widths to a constant, we will use the technique of linear grouping while increasing the final peak by at most $2\varepsilon OPT$. We start by considering all the slices in a pile, one on top of the other and sorted non-increasingly by width from bottom to top (and putting sibling slices consecutively). Since these slices have a width at least εW and total area at most $OPT \cdot W$, the pile has total height at most $\frac{1}{\varepsilon} OPT$. Starting from the bottom, we partition the pile into groups G_1, \dots, G_q of height exactly εOPT (except possibly for G_q which may have a smaller height). We remove from the solution the slices in G_1 and any slice in G_2 which used to have a sibling slice in G_1 , and temporarily remove the corresponding tasks. Observe that we are removing the tasks whose slices are fully contained in G_1 plus at most one extra task. In particular, the total height of the removed tasks is at most $(\varepsilon + \mu)OPT \leq 2\varepsilon OPT$ (here we use $\mu \leq \varepsilon$).

Next, we round up the widths of the remaining slices as follows: for $i = 2, \dots, q$, the width of slices (still) in G_i are rounded to the smallest width of any slice originally in G_{i-1} . We call this set of slices the rounded slices, and next focus on packing them. Notice that rounded slices have at most $1/\varepsilon^2$ distinct widths. This also induces a matching between each rounded slice a and a distinct original slice b , so that $w(b) \geq w(a)$. In particular, we can schedule each such a starting on the first edge of $P(b)$ without increasing the overall peak.

Now we left-shift the horizontal slices in the solution as much as possible while still obtaining a schedule whose demand profile is upper bounded by D . Let e be the starting edge of some slice S at the end of the process. Notice that one of the following cases holds: (1) D increases on edge e (including as a special case when e is the leftmost edge of G) or (2)

20:12 Approximation Algorithms for Demand Strip Packing

the edge f to the left of e is the ending edge of some other slice S' . Indeed otherwise it would be possible to left-shift S while respecting all the constraints. This implies that the possible positions for the starting edge of any slice can be obtained by considering the $O_\varepsilon(1)$ edges where D increases and then adding the total width of a few slices. Notice that there are at most $1/\varepsilon^2$ such widths, and we can sum up at most $1/\varepsilon$ of them (since horizontal slices have widths at least εW). Altogether, the number of possible starting edges for the slices is $O_\varepsilon(1)$.

Consider the leftmost possible such edge e and all the rounded slices S_e starting on e in this left-shifted schedule. We partition S_e by width w , and for each such width w and corresponding set of slices $S_{e,w}$, we construct a horizontal container of width w and height $h(S_{e,w})$ where we pack $S_{e,w}$. We repeat this procedure for each possible starting edge e , obtaining in the end $K \leq O_\varepsilon(1)$ containers in total where we packed all the rounded slices. Notice that the width of each container is the width of some rounded slice, which in turn is the width of some task. Hence the widths of the containers are guessable in the usual sense.

We next turn the above packing of rounded slices into a feasible packing of tasks (into the same containers). First of all, we repack the rounded slices as follows. We consider all the slices S_w of a given width w in any order where sibling slices appear consecutively, and all the containers C_w of that width in any order. We pack each slice $s \in S_w$ in the first container $C \in C_w$ where s still fits. Notice that $h(S_w) = h(C_w)$ by construction, hence we repack all rounded slices this way. Next we consider the tasks i whose slices are all contained in the *same* container C , and pack i into C . By construction this packing is feasible. We add the tasks which are not packed this way to the set of removed tasks defined earlier. We round up the heights of the containers to the next multiple of $\frac{\varepsilon}{K}OPT$, hence making such heights guessable. This way the peak increases at most by εOPT .

Consider the set of removed tasks. Recall that the tasks removed in the initial rounding phase have total height at most $2\varepsilon OPT$. In the following packing phase, we remove at most one task per container, hence these removed tasks have height at most $K \cdot \mu \cdot OPT \leq \varepsilon OPT$. Here we assume that $\mu \leq \varepsilon/K$: this can be achieved by choosing $f(x) = x/K$ in Lemma 8. Hence the removed tasks altogether have total height at most $3\varepsilon OPT$: we pack these tasks in one extra (guessable) horizontal container of width W and height $3\varepsilon OPT$. ◀

From the above construction it is possible to derive a schedule of non-narrow tasks of small enough peak.

► **Lemma 11.** *It is possible to compute in polynomial-time a feasible schedule of $\mathcal{I} \setminus \mathcal{N}$ with peak at most $(\frac{5}{3} + 7\varepsilon)OPT$.*

Proof. Consider the guessable containers for tall, large, and horizontal tasks and the corresponding schedule as described before. This schedule has a peak of at most $(\frac{5}{3} + \varepsilon)OPT + 4\varepsilon OPT$. By Lemma 8 the medium tasks fit into a horizontal container of width W and height εOPT . Altogether this leads to a packing of $\mathcal{L} \cup \mathcal{T} \cup \mathcal{H} \cup \mathcal{M} = \mathcal{I} \setminus \mathcal{N}$ into $O_\varepsilon(1)$ guessable containers that can be scheduled with peak at most $(\frac{5}{3} + 6\varepsilon)OPT$.

It is easy to pack $\mathcal{L} \cup \mathcal{T} \cup \mathcal{M}$ into the corresponding containers. For \mathcal{H} we apply Lemma 5 with $\varepsilon' = \varepsilon^2$ to assign the horizontal tasks to them, obtaining a set of horizontal unplaced tasks of an area at most $\varepsilon^2 \cdot W \cdot OPT$ (hence of total height at most εOPT). The latter tasks can be placed into an extra horizontal container of height εOPT and width W . The resulting set of containers can be scheduled with a peak at most $(\frac{5}{3} + 7\varepsilon)OPT$, and such a schedule can be efficiently computed as already discussed. ◀

4.3 Scheduling Narrow Tasks

At this point it just remains to schedule the narrow tasks. For this goal we need the following generalization of Lemma 6 that considers (Q, t^*) -sorted partial schedules. Recall that for a node t^* of the path and a value $Q \geq 0$, a schedule is (Q, t^*) -sorted if the demand to the left of t^* is at least Q , and to the right of t^* the demand profile is non-increasing.

► **Lemma 12.** *Let \mathcal{I} be an instance of DSP and $\alpha > 0$. Suppose we are given a $((1 + \alpha)OPT, t^*)$ -sorted schedule of $\mathcal{I}' \subseteq \mathcal{I}$. Let $\mathcal{I}'' := \mathcal{I} \setminus \mathcal{I}'$ and assume that:*

- *The peak of the schedule is at most π , with $\pi \geq (1 + \alpha)OPT + h_{\max}(\mathcal{I}'')$, and*
- *$w_{\max}(\mathcal{I}'') \leq \frac{\alpha}{2(\alpha+1)}W$.*

Then it is possible to compute in polynomial time a schedule of \mathcal{I} with peak at most π .

Proof. By overloading notation, let t^* also denote the number of edges to the left of t^* . Notice first that $W - t^* \geq 2w_{\max}(\mathcal{I}'')$. Indeed otherwise, since the input schedule is $((1 + \alpha)OPT, t^*)$ -sorted, the total area of the tasks in \mathcal{I}' would be at least

$$t^* \cdot (1 + \alpha)OPT > (W - 2w_{\max}(\mathcal{I}''))(1 + \alpha)OPT \geq W \cdot OPT$$

which is not possible. Roughly speaking, to prove the desired claim, we will apply Lemma 6 to the demand profile induced by the edges to the right of t^* . In more detail, we consider a new instance defined by a path with $\tilde{W} = W - t^*$ edges and a set of tasks $\tilde{\mathcal{I}}$ consisting of $\tilde{\mathcal{I}}'' := \mathcal{I}''$ plus a set $\tilde{\mathcal{I}}'$ of \tilde{W} tasks having width 1 and, for each edge e to the right of t^* , height equal to the total demand on edge e in the original schedule for \mathcal{I}' . To see that the required hypotheses are satisfied, notice that by scheduling the tasks in $\tilde{\mathcal{I}}'$ one next to the other sorted non-increasingly by height we obtain a sorted partial schedule of a peak at most π , where

$$\pi \geq h_{\max}(\mathcal{I}'') + (1 + \alpha)OPT \geq h_{\max}(\tilde{\mathcal{I}}'') + \max\{a(\tilde{\mathcal{I}})/\tilde{W}, h_{\max}(\tilde{\mathcal{I}}'')\}.$$

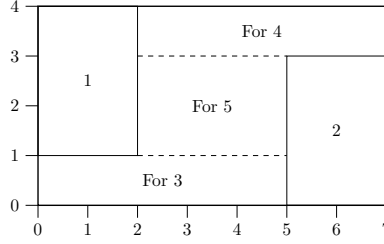
The last inequality above holds since $a(\tilde{\mathcal{I}}) \leq OPT \cdot W - a(\mathcal{I}') \leq OPT(W - (1 + \alpha)t^*)$. Finally, we notice that $w_{\max}(\tilde{\mathcal{I}}'') \leq \tilde{W}/2$, and

$$\begin{aligned} & (\tilde{W} - w_{\max}(\tilde{\mathcal{I}}''))(\pi - h_{\max}(\tilde{\mathcal{I}}'')) + w_{\max}(\tilde{\mathcal{I}}'') \cdot h_{\max}(\tilde{\mathcal{I}}'') \\ & \geq (\tilde{W} - w_{\max}(\tilde{\mathcal{I}}''))(1 + \alpha)OPT \\ & = OPT(W - t^*(1 + \alpha)) + \alpha W \cdot OPT - w_{\max}(\mathcal{I}'')(1 + \alpha)OPT \\ & \geq a(\tilde{\mathcal{I}}) + \alpha W \cdot OPT - \frac{\alpha}{2}W \cdot OPT \geq a(\tilde{\mathcal{I}}). \end{aligned}$$

Hence all the conditions of Lemma 6 apply. Given that the demand profile of the partial schedule for $\tilde{\mathcal{I}}'$ is the same as the demand profile induced by the edges to the right of t^* in the original schedule for \mathcal{I}' , we can schedule \mathcal{I}'' on top of the input schedule without exceeding the peak π . ◀

We now have all the ingredients to prove Theorem 1.

Proof of Theorem 1. Consider the schedule of $\mathcal{I} \setminus \mathcal{N}$ with peak at most $\pi := (\frac{5}{3} + 7\varepsilon)OPT$ provided by Lemma 11. We perform a π -left-pushing of this schedule, however without left-shifting any tall task. Let us prove that this partial schedule of $\mathcal{I}' := \mathcal{I} \setminus \mathcal{N}$ satisfies all the required properties of Lemma 12 with parameter π . First of all, there exists a node t^* for which (1) every edge to the left of t^* (if any) has demand larger than $(1 + 7\varepsilon)OPT$, and (2) the demand profile to the right of t^* is non-increasing. Indeed, if (1) does not hold, then there exists an edge having demand less than $(1 + 7\varepsilon)OPT$ and the following edge has



■ **Figure 5** If we assume by contradiction that some optimal solution of height 4 for the GSP instance described in Lemma 13 exists, it must have this structure.

demand larger than $(1 + 7\varepsilon)OPT$. But this means that some task which is not tall can be left-shifted (as there can be only one tall task per edge); similarly, if (2) does not hold, there is a pair of contiguous edges to the right of t^* where the demand strictly increases from left to right. But since the tall tasks are sorted non-increasingly by height, this implies that there exists a task that is not tall that can be left-shifted. In conclusion, the solution is $((1 + 7\varepsilon)OPT, t^*)$ -sorted and also $w_{\max}(\mathcal{N}) \leq \varepsilon W \leq \frac{7\varepsilon}{2(1+7\varepsilon)}W$ for ε small enough. Since $\pi \geq (1 + 7\varepsilon)OPT + h_{\max}(\mathcal{N})$, by Lemma 12 we obtain a feasible schedule of peak at most π . The claim follows by scaling ε appropriately. ◀

5 Comparison between DSP and GSP

In this Section we provide instances where a gap between the optimal values they achieve interpreted as DSP and GSP instances can be observed. First we discuss the general case and then the case of SQUARE-DSP. It is worth noticing that, for the general case, an analogous proof can be derived from the results in [8].

► **Lemma 13.** *There exists an instance of DSP with optimal peak 4 such that the corresponding GSP instance has optimal peak 5.*

Proof. Consider the following DSP instance \mathcal{I} , where $W = 7$ and the set of tasks consists of the following eight elements (see Figure 1a for a depiction):

- Two tasks of width 2 and height 3 (tasks 1 and 2 in the figure),
- Two tasks of width 4 and height 1 (tasks 3 and 4 in the figure),
- One task of width 3 and height 1 (task 5 in the figure),
- One task of width 1 and height 1 (task 6 in the figure), and
- Two tasks of width 1 and height 2 (tasks 7 and 8 in the figure).

As it is possible to see in Figure 1a, the optimal solution has peak 4 (since $OPT \geq a(\mathcal{I})/W = 4$). We will show now that there is no solution for the corresponding GSP instance of height 4, which would conclude the proof.

Suppose by contradiction that there exists a solution to the corresponding GSP instance of height 4. Let us imagine for the sake of presentation that we draw a grid of unit-size cells over the rectangular region $[0, 7] \times [0, 4]$, defining four rows of height 1 and seven columns of width 1. First of all, notice that in any feasible packing of the rectangles into the region, rectangles 1 and 2 cannot be touching the top (resp. bottom) boundary of the region at the same time. If that is the case, then the rectangles 3 and 4 do not fit in the region as they cannot be placed in the same row and none of them fits in the rows which are partially

occupied by rectangles 1 and 2. So let us assume w.l.o.g. that rectangle 1 touches the top boundary and rectangle 2 touches the bottom boundary. Since they both partially occupy the middle rows of the region, rectangles 3 and 4 must be placed one touching the bottom boundary and the other touching the top boundary. This implies that rectangle 5 has to be placed in one of the middle rows (in the other rows there is just one cell free), forcing us to place rectangles 1 and 2 one touching the left boundary and the other touching the right boundary (see Figure 5). Suppose rectangle 5 is assigned to the second row from bottom to top (the other case being symmetric). Then in the two topmost rows we have to pack two rectangles of height 2 plus a rectangle of width 4 which is not possible as their total width is larger than the space left due to rectangle 1. This contradicts the fact that there is a feasible solution for the GSP instance \mathcal{I} of height 4. ◀

Now we will prove that even for the case of square tasks, the optimum packing for the two problems of SQUARE-DSP and SQUARE-GSP can exhibit a gap.

► **Lemma 14.** *There exists an instance of SQUARE-DSP such that the optimal schedule has peak 11 but every feasible solution for the corresponding SQUARE-GSP instance has height at least 12.*

Proof. Consider a SQUARE-DSP with $W = 13$ and containing the following set \mathcal{I} of tasks (see Figure 1b for a depiction):

- Two tasks of height/width 6 (tasks 1 and 2 in the figure),
- Two tasks of height/width 5 (tasks 3 and 4 in the figure),
- One task of height/width 3 (task 5 in the figure),
- Two tasks of height/width 2 (tasks 6 and 7 in the figure), and
- Four tasks of height/width 1 (tasks 8, 9, 10 and 11 in the figure).

Since $a(\mathcal{I}) = 11 \cdot 13$, we have that $OPT \geq 11$. Figure 1b shows that the optimal peak is at most 11 and hence it is exactly 11.

Assume by contradiction that there exists a feasible packing for the corresponding SQUARE-GSP instance of height at most 11. Consider \mathcal{K} to be the region $[0, 0] \times [13, 11]$ in the plane, and let (x_i, y_i) be the coordinate of the bottom-left corner of task i in the solution. Notice that \mathcal{K} must be completely filled with tasks.

We can assume that $x_1 \leq x_2$, and since tasks 1 and 2 have height 6 and the height of \mathcal{K} is 11, it must hold that $x_1 \leq x_2 + 6$. Hence, w.l.o.g. there are two cases to consider:

- $x_1 = 0$ and $x_2 = 6$:

In this case the region $[12, y_2] \times [13, y_2 + 6]$ can only contain squares of size 1, and they cannot fill the region completely, so this case cannot happen.

- $x_1 = 0$ and $x_2 = 7$:

We show that $y_1, y_2 \in \{0, 5\}$; Assume that $y_1 \notin \{0, 5\}$. Then tasks 2, 3 and 4 must be packed inside the region $[6, 0] \times [13, 11]$ since they can not be packed above or below task 1. Since $a(j_2) + a(j_3) + a(j_4) > 77$, this is not possible, hence proving the claim.

Note that if $y_1 = y_2$ then, similarly to the previous case, the area in $[6, y_1] \times [7, y_1 + 6]$ can only contain tasks of size 1 and they cannot fill this region completely. So we can assume that $(x_1, y_1) = (0, 0)$ and $(x_2, y_2) = (7, 5)$.

Now every remaining rectangle is either packed in $[6, 0] \times [13, 5]$ or in $[0, 6] \times [7, 11]$. However, among tasks 3, 4 and 5, it is not possible to place two of them in one of the previously mentioned rectangular region together, contradicting the existence of a feasible solution of height 11. ◀

References

- 1 Anna Adamaszek, Tomasz Kociumaka, Marcin Pilipczuk, and Michal Pilipczuk. Hardness of approximation for strip packing. *ACM Trans. Comput. Theory*, 9(3):14:1–14:7, 2017. doi:10.1145/3092026.
- 2 Anna Adamaszek and Andreas Wiese. A quasi-ptas for the two-dimensional geometric knapsack problem. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1491–1505. SIAM, 2015. doi:10.1137/1.9781611973730.98.
- 3 Soroush Alamdari, Therese C. Biedl, Timothy M. Chan, Elyot Grant, Krishnam Raju Jampani, Srinivasan Keshav, Anna Lubiw, and Vinayak Pathak. Smart-grid electricity allocation via strip packing with slicing. In *13th International Symposium on Algorithms and Data Structures (WADS)*, volume 8037, pages 25–36. Springer, 2013. doi:10.1007/978-3-642-40104-6_3.
- 4 Aris Anagnostopoulos, Fabrizio Grandoni, Stefano Leonardi, and Andreas Wiese. A mazing $(2 + \varepsilon)$ -approximation for unsplittable flow on a path. *ACM Transactions on Algorithms*, 14(4):55:1–55:23, 2018. doi:10.1145/3242769.
- 5 Brenda S. Baker, Edward G. Coffman Jr., and Ronald L. Rivest. Orthogonal packings in two dimensions. *SIAM Journal on Computing*, 9(4):846–855, 1980. doi:10.1137/0209064.
- 6 Nikhil Bansal, Alberto Caprara, Klaus Jansen, Lars Prädell, and Maxim Sviridenko. A structural lemma in 2-dimensional packing, and its implications on approximability. In *Algorithms and Computation, 20th International Symposium (ISAAC)*, volume 5878, pages 77–86. Springer, 2009. doi:10.1007/978-3-642-10631-6_10.
- 7 Jatin Batra, Naveen Garg, Amit Kumar, Tobias Mömke, and Andreas Wiese. New approximation schemes for unsplittable flow on a path. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 47–58. SIAM, 2015. doi:10.1137/1.9781611973730.5.
- 8 Iwo Bladek, Maciej Drozdowski, Frédéric Guinand, and Xavier Schepler. On contiguous and non-contiguous parallel task scheduling. *J. Sched.*, 18(5):487–495, 2015. doi:10.1007/s10951-015-0427-z.
- 9 Adam L. Buchsbaum, Howard J. Karloff, Claire Kenyon, Nick Reingold, and Mikkel Thorup. OPT versus LOAD in dynamic storage allocation. *SIAM Journal on Computing*, 33(3):632–646, 2004. doi:10.1137/S0097539703423941.
- 10 Mihai Burcea, Wing-Kai Hon, Hsiang-Hsuan Liu, Prudence W. H. Wong, and David K. Y. Yau. Scheduling for electricity cost in a smart grid. *Journal of Scheduling*, 19(6):687–699, 2016. doi:10.1007/s10951-015-0447-8.
- 11 Gruia Călinescu, Amit Chakrabarti, Howard J. Karloff, and Yuval Rabani. An improved approximation algorithm for resource allocation. *ACM Transactions on Algorithms*, 7(4):48:1–48:7, 2011. doi:10.1145/2000807.2000816.
- 12 Edward Grady. Coffman and John L. Bruno. *Computer and job-shop scheduling theory / edited by E. G. Coffman, Jr. ; coauthors, J. L. Bruno ... [et al.]*. Wiley New York, 1976.
- 13 Edward G. Coffman Jr., János Csirik, Gábor Galambos, Silvano Martello, and Daniele Vigo. *Bin Packing Approximation Algorithms: Survey and Classification*, pages 455–531. Springer New York, 2013. doi:10.1007/978-1-4419-7997-1_35.
- 14 Edward G. Coffman Jr., M. R. Garey, David S. Johnson, and Robert Endre Tarjan. Performance bounds for level-oriented two-dimensional packing algorithms. *SIAM Journal on Computing*, 9(4):808–826, 1980. doi:10.1137/0209062.
- 15 Max A. Deppert, Klaus Jansen, Arindam Khan, Malin Rau, and Malte Tutas. Peak demand minimization via sliced strip packing. *CoRR*, abs/2105.07219, 2021. URL: <https://arxiv.org/abs/2105.07219>.
- 16 Pierre-François Dutot, Grégory Mounié, and Denis Trystram. Scheduling parallel tasks approximation algorithms. In Joseph Y.-T. Leung, editor, *Handbook of Scheduling - Algorithms, Models, and Performance Analysis*. Chapman and Hall/CRC, 2004.

- 17 Waldo Gálvez, Fabrizio Grandoni, Afrouz Jabal Ameli, and Kamyar Khodamoradi. Approximation algorithms for demand strip packing. *CoRR*, abs/2105.08577, 2021. URL: <https://arxiv.org/abs/2105.08577>, arXiv:2105.08577.
- 18 Waldo Gálvez, Fabrizio Grandoni, Sandy Heydrich, Salvatore Ingala, Arindam Khan, and Andreas Wiese. Approximating geometric knapsack via L-packings. In *58th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 260–271. IEEE Computer Society, 2017. doi:10.1109/FOCS.2017.32.
- 19 Waldo Gálvez, Fabrizio Grandoni, Sandy Heydrich, Salvatore Ingala, Arindam Khan, and Andreas Wiese. Approximating geometric knapsack via L-packings. *CoRR*, abs/1711.07710, 2017. URL: <http://arxiv.org/abs/1711.07710>.
- 20 Waldo Gálvez, Fabrizio Grandoni, Salvatore Ingala, and Arindam Khan. Improved pseudo-polynomial-time approximation for strip packing. In *36th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, volume 65, pages 9:1–9:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016. doi:10.4230/LIPIcs.FSTTCS.2016.9.
- 21 Fabrizio Grandoni, Salvatore Ingala, and Sumedha Uniyal. Improved approximation algorithms for unsplittable flow on a path with time windows. In *Approximation and Online Algorithms - 13th International Workshop, (WAOA)*, volume 9499, pages 13–24. Springer, 2015. doi:10.1007/978-3-319-28684-6_2.
- 22 Fabrizio Grandoni, Tobias Mömke, Andreas Wiese, and Hang Zhou. A $(5/3 + \epsilon)$ -approximation for unsplittable flow on a path: placing small tasks into boxes. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 607–619. ACM, 2018. doi:10.1145/3188745.3188894.
- 23 Rolf Harren, Klaus Jansen, Lars Prädél, and Rob van Stee. A $(5/3 + \epsilon)$ -approximation for strip packing. *Computational Geometry*, 47(2):248–267, 2014. doi:10.1016/j.comgeo.2013.08.008.
- 24 Rolf Harren and Rob van Stee. Improved absolute approximation ratios for two-dimensional packing problems. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 12th International Workshop (APPROX)*, volume 5687, pages 177–189. Springer, 2009. doi:10.1007/978-3-642-03685-9_14.
- 25 Sören Henning, Klaus Jansen, Malin Rau, and Lars Schmarje. Complexity and inapproximability results for parallel task scheduling and strip packing. *Theory of Computing Systems*, 64(1):120–140, 2020. doi:10.1007/s00224-019-09910-6.
- 26 Klaus Jansen. A $(3/2 + \epsilon)$ approximation algorithm for scheduling moldable and non-moldable parallel tasks. In *24th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pages 224–235. ACM, 2012. doi:10.1145/2312005.2312048.
- 27 Klaus Jansen and Malin Rau. Closing the gap for pseudo-polynomial strip packing. In *27th Annual European Symposium on Algorithms (ESA)*, volume 144, pages 62:1–62:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.ESA.2019.62.
- 28 Klaus Jansen and Malin Rau. Improved approximation for two dimensional strip packing with polynomial bounded width. *Theoretical Computer Science*, 789:34–49, 2019. doi:10.1016/j.tcs.2019.04.002.
- 29 Klaus Jansen and Roberto Solis-Oba. Rectangle packing with one-dimensional resource augmentation. *Discrete Optimization*, 6(3):310–323, 2009. doi:10.1016/j.disopt.2009.04.001.
- 30 Klaus Jansen and Ralf Thöle. Approximation algorithms for scheduling parallel jobs. *SIAM Journal on Computing*, 39(8):3571–3615, 2010. doi:10.1137/080736491.
- 31 Mohammad M. Karbasioun, Gennady Shaikhet, Evangelos Kranakis, and Ioannis Lambadaris. Power strip packing of malleable demands in smart grid. In *Proceedings of IEEE International Conference on Communications, (ICC)*, pages 4261–4265. IEEE, 2013. doi:10.1109/ICC.2013.6655233.

- 32 Claire Kenyon and Eric Rémila. A near-optimal solution to a two-dimensional cutting stock problem. *Mathematics of Operations Research*, 25(4):645–656, 2000. doi:10.1287/moor.25.4.645.12118.
- 33 Joseph Y.-T. Leung, Tommy W. Tam, C. S. Wong, Gilbert H. Young, and Francis Y. L. Chin. Packing squares into a square. *Journal of Parallel and Distributed Computing*, 10(3):271–275, 1990. doi:10.1016/0743-7315(90)90019-L.
- 34 Fu-Hong Liu, Hsiang-Hsuan Liu, and Prudence W. H. Wong. Greedy is optimal for online restricted assignment and smart grid scheduling for unit size jobs. In *Approximation and Online Algorithms - 17th International Workshop (WAOA)*, volume 11926, pages 217–231. Springer, 2019. doi:10.1007/978-3-030-39479-0_15.
- 35 Fu-Hong Liu, Hsiang-Hsuan Liu, and Prudence W. H. Wong. Non-preemptive scheduling in a smart grid model and its implications on machine minimization. *Algorithmica*, 82(12):3415–3457, 2020. doi:10.1007/s00453-020-00733-3.
- 36 Tobias Mömke and Andreas Wiese. A $(2+\epsilon)$ -approximation algorithm for the storage allocation problem. In *Automata, Languages, and Programming - 42nd International Colloquium (ICALP)*, volume 9134, pages 973–984. Springer, 2015. doi:10.1007/978-3-662-47672-7_79.
- 37 Tobias Mömke and Andreas Wiese. Breaking the barrier of 2 for the storage allocation problem. In *47th International Colloquium on Automata, Languages, and Programming (ICALP)*, volume 168, pages 86:1–86:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.ICALP.2020.86.
- 38 Giorgi Nadiradze and Andreas Wiese. On approximating strip packing with a better ratio than $3/2$. In Robert Krauthgamer, editor, *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1491–1510. SIAM, 2016. doi:10.1137/1.9781611974331.ch102.
- 39 Anshu Ranjan, Pramod P. Khargonekar, and Sartaj Sahni. Offline first fit scheduling in smart grids. In *2015 IEEE Symposium on Computers and Communication (ISCC)*, pages 758–763. IEEE Computer Society, 2015. doi:10.1109/ISCC.2015.7405605.
- 40 Ingo Schiermeyer. Reverse-fit: A 2-optimal algorithm for packing rectangles. In Jan van Leeuwen, editor, *Algorithms (ESA) - Second Annual European Symposium*, volume 855, pages 290–299. Springer, 1994. doi:10.1007/BFb0049416.
- 41 Daniel Dominic Sleator. A 2.5 times optimal algorithm for packing in two dimensions. *Information Processing Letters*, 10(1):37–40, 1980. doi:10.1016/0020-0190(80)90121-0.
- 42 A. Steinberg. A strip-packing algorithm with absolute performance bound 2. *SIAM Journal on Computing*, 26(2):401–409, 1997. doi:10.1137/S0097539793255801.
- 43 Shaojie Tang, Qiuyuan Huang, Xiang-Yang Li, and Dapeng Wu. Smoothing the energy consumption: Peak demand reduction in smart grid. In *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, pages 1133–1141. IEEE, 2013. doi:10.1109/INFCOM.2013.6566904.
- 44 Sean Yaw and Brendan Mumey. Scheduling non-preemptible jobs to minimize peak demand. *Algorithms*, 10(4):122, 2017. doi:10.3390/a10040122.
- 45 Sean Yaw, Brendan Mumey, Erin McDonald, and Jennifer Lemke. Peak demand scheduling in the smart grid. In *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 770–775. IEEE, 2014. doi:10.1109/SmartGridComm.2014.7007741.

A Hardness of Approximation for Square-DSP

For several rectangle packing problems it is usually the case that they are NP-hard (or APX-hard) even when restricted to instances consisting solely of squares [33]. This holds also for DSP, as the following theorem shows.

► **Theorem 15.** *For any $\varepsilon > 0$, there exists no polynomial-time $(3/2 - \varepsilon)$ -approximation algorithm for SQUARE-DSP unless $NP = P$.*

In order to prove this result, we show a gap-producing reduction from the NP-complete BALANCED PARTITION problem, formally defined as follows.

► **Definition 16** (BALANCED PARTITION). *In an instance of the BALANCED PARTITION problem, we are given a set of $2n$ positive integers $A = \{a_1, a_2, \dots, a_{2n}\}$. The goal is to decide whether there exists a partitioning of A into A_1 and A_2 such that $|A_1| = |A_2| = n$, and the sum of the numbers in each of the two sets is equal to a target value $B = \left(\sum_{j=1}^{2n} a_j\right)/2$.*

We start first by proving that the BALANCED PARTITION problem is NP-complete. This result is folklore by now but, for the sake of completeness, we bring a complete proof.

► **Theorem 17.** *BALANCED PARTITION is NP-complete.*

Proof. We will reduce the PARTITION problem to the balanced variant in polynomial time. Given an instance \mathcal{I} of the PARTITION problem with n numbers a_1, a_2, \dots, a_n , we construct an instance \mathcal{I}'_k of the BALANCED PARTITION problem for each $k \in \{1, 2, \dots, \lfloor n/2 \rfloor\}$. Let C be $(\sum_{j=1}^n a_j) + 1$. For each k , the instance \mathcal{I}'_k is defined as follows. Let \mathcal{I}'_k have all the initial numbers, a_1, a_2, \dots, a_n . Add the set **dummy** of $n - 2k + 2$ extra numbers where **dummy** = $\{\alpha, \beta_1, \beta_2, \dots, \beta_{n-2k+1}\}$ in which $\alpha = (n - 2k + 1)C$ and $\beta_i = C$ for each $i \in [n - k + 1]$. We claim that \mathcal{I} is a *Yes* instance of the PARTITION problem if and only if at least one \mathcal{I}'_k is a *Yes* instance of the BALANCED PARTITION problem.

Completeness: Assume that \mathcal{I} is a *Yes* instance of the PARTITION problem. Let S_1 and S_2 be the two sets of equal sum, say B . These sets may not necessarily have the same cardinality. With sum renumbering, assume that $S_1 = \{a_1, a_2, \dots, a_k\}$ and $S_2 = \{a_{k+1}, a_{k+2}, \dots, a_n\}$ for some $k \in [\lfloor n/2 \rfloor]$. It is easy to see that the instance \mathcal{I}'_k of the BALANCED PARTITION problem is a *Yes* instance, since we can make the sets $S'_1 = \{a_1, a_2, \dots, a_k, \beta_1, \beta_2, \dots, \beta_{n-2k+1}\}$ and $S'_2 = \{a_{k+1}, a_{k+2}, \dots, a_n, \alpha\}$ both with the sum $B + (n - 2k + 1)C$ and cardinality $n - k + 1$.

Soundness: Now assume that \mathcal{I} is a *No* instance of the PARTITION problem. We claim that no \mathcal{I}'_k can be a *Yes* instance of the BALANCED PARTITION problem either. For the sake of contradiction, assume \mathcal{I}'_k is a *Yes* instance with two partitions S'_1 and S'_2 of the same sum and cardinality. Note that if no β_i is placed in the same set as α , we reach a contradiction since we then can find two sets $S_1 = \{a_1, a_2, \dots, a_k\}$ and $S_2 = \{a_{k+1}, a_{k+2}, \dots, a_n\}$ of the same sum of the original PARTITION instance \mathcal{I} . So with some renumbering, we can assume we have $S'_1 = \{a_1, a_2, \dots, a_{k'}, \alpha, \beta_1, \beta_2, \dots, \beta_\ell\}$ and $S'_2 = \{a_{k'+1} + a_{k'+2}, \dots, a_n, \beta_{\ell+1}, \beta_{\ell+2}, \dots, \beta_{n-2k+1}\}$ for some k' and ℓ in which:

$$\sum_{j=1}^{k'} a_j + (n - 2k + 1)C + \ell \cdot C = \sum_{j'=k'+1}^n a'_j + (n - 2k + 1 - \ell)C.$$

This implies that

$$2\ell \cdot C = \sum_{j'=k'+1}^n a'_j - \sum_{j=1}^{k'} a_j \leq \sum_{j'=k'+1}^n a'_j < C,$$

which is a contradiction. \blacktriangleleft

Proof of Theorem 15. Assume an instance \mathcal{I} of the BALANCED PARTITION problem is given. Based in this instance, we define an instance \mathcal{I}' of SQUARE-DSP. Let a_{max} denote the maximum value among the integers in A . Define C as $1/\varepsilon \cdot \sum_{j=1}^{2n} a_j$, where ε is chosen such that $1/\varepsilon$ is a large but constant integer. Note that $C > 1/\varepsilon \cdot a_{max}$. Let \mathcal{I}' have $2n$ tasks, where each task i has width and height $C+a_i$ for $i \in [2n]$. Our goal is to schedule the $2n$ tasks into a path of $W = n \cdot C + B$ edges while minimizing the peak. Based on the hardness of the BALANCED PARTITION problem, we show that it is hard to distinguish between the case where an schedule with peak $2C(1+\varepsilon)$ exists and the case where the minimum peak is larger than $3(C+1)-\varepsilon$.

Completeness: Assume that \mathcal{I} is a Yes instance, meaning that a partitioning $A = A_1 \dot{\cup} A_2$ exists that satisfies the cardinality and sum constraints. Define two shelves of squares, $S_i = \{j | a_j \in A_i\}$ for $i = 1, 2$, and schedule them starting at the leftmost edge. The width of each shelf is equal to $n \cdot C + B$ and the peak is at most $2 \cdot (C + a_{max}) < 2C(1 + \varepsilon)$.

Soundness: Now, consider a No instance \mathcal{I} . We claim that, for the corresponding SQUARE-DSP instance, no schedule with peak smaller than or equal to $3(C+1) - \varepsilon$ exists. For the sake of contradiction, assume that it is the case. Since the size of each task is at least $C+1$, it means that in the optimal solution for \mathcal{I}' , no three tasks use the same edge. Also, the total width of the tasks is $2W = 2(nC + b)$, so no edge can have less than two tasks. This allows us to split the tasks S into two sets S_1 and S_2 . We start at the leftmost edge and pick one of the two tasks placed on this edge arbitrarily and put in S_1 . Since every edge has exactly two tasks, immediately to the right of this task at least one another task must start. We put it in S_1 as well and proceed until we reach the rightmost edge, breaking ties arbitrarily along the way. We set $S_2 = S \setminus S_1$. It remains to show that each set has exactly n tasks. Assume otherwise; let S_1 be composed of tasks s_1, \dots, s_{n+k} , and S_2 be the tasks s_{n+k+1}, \dots, s_{2n} for some k , $1 \leq k \leq n$. Since the tasks are placed one next to the other in each shelf, we have that $(n+k)C + \sum_{j=1}^{n+k} a_j = (n-k)C + \sum_{j'=n+k+1}^{2n} a_{j'}$. Therefore $\sum_{j'=n+k+1}^{2n} a_{j'} \geq 2k \cdot C$, which is a contradiction for any value of $k > 0$ by our choice of C .

As a result, assuming that $\text{NP} \neq \text{P}$, no polynomial-time algorithm can approximate the SQUARE-DSP problem within a factor of $\frac{3(C+1)-\varepsilon}{2C(1+\varepsilon)} = 3/2 - \varepsilon'$, for some $\varepsilon' = O(\varepsilon)$. \blacktriangleleft

B A PTAS for DSP with short tasks

In this section we will prove Theorem 2 restated below.

► **Theorem 2.** *Given $\varepsilon > 0$ small enough, there exists $\delta > 0$ and a polynomial time algorithm such that, given an instance of DSP with optimal value OPT and consisting solely of tasks having height at most $\delta \cdot OPT$, it computes a $(1 + O(\varepsilon))$ -approximate solution.*

Before proving the result in detail we provide a couple of required technical lemmas regarding the computation of π -left-pushing of a given schedule $P(\cdot)$. First of all, we prove that such a solution can be indeed computed efficiently.

► **Lemma 18.** *Given a feasible schedule $P(\cdot)$ with peak π for an instance \mathcal{I} , one can compute a π' -left-pushing of $P(\cdot)$, with $\pi' \geq \pi$, in polynomial time.*

Proof. Let $1, \dots, n$ be the tasks sorted according to their starting edge in $P(\cdot)$ from left to right. Let S_i be the starting edge of task i . First, inductively, we compute a π' -left-pushing of $\mathcal{I} \setminus \{n\}$ and do not left-shift task n . Since we only left-shifted the tasks, the demand on the edges from S_n to e_W cannot increase. Thus, we reach a feasible solution such that its peak does not exceed π' . Now we compute the starting time of task n , s^* , if we left-shift this task as much as possible. Note that s^* can only be either the leftmost edge or some edge e such that some previous task finishes next to the left of e , as otherwise at least one more unit of left-shifting is possible for task n . Now, using this fact, we have at most n possibilities for s^* and we can compute this value in polynomial time. Note that if we call the obtained schedule as $P'(\cdot)$, then $P'(\cdot)$ is indeed a π' -left-pushing of $P(\cdot)$. ◀

The following lemma summarizes the useful properties we can get when computing a left-pushing.

► **Lemma 19.** *Given a feasible schedule $P(\cdot)$ with peak π for an instance \mathcal{I} , the π' -left-pushing of $P(\cdot)$ for $\pi' \geq \pi$, let us say $P'(\cdot)$, satisfies the following properties:*

1. *There exists a node t^* such that $P'(\cdot)$ is $(\pi' - h_{\max}(\mathcal{I}), t^*)$ -sorted, and*
2. *every $i \in \mathcal{I}$ has a starting edge in \mathcal{E}' of the form $\sum_{j \in \mathcal{I}'} w(j)$ for some $\mathcal{I}' \subseteq \mathcal{I} \setminus \{i\}$ (0 if \mathcal{I}' is empty).*

Proof. We now show a proof of the two properties:

1. Suppose that there exists a node k such that the demand on the edge to the left of k is smaller than $\pi' - h_{\max}(\mathcal{I})$ and the demand on the edge to the right of k is larger than $\pi' - h_{\max}(\mathcal{I})$. This implies that some task starts at the edge to the right of k , but then it is possible to left-shift this task without surpassing the threshold of π' which is a contradiction. At this point we know that there exists k' such that every edge to the left of k' has demand larger than $\pi' - h_{\max}(\mathcal{I})$, and let k^* be the rightmost such node. Similarly to the previous case, if after k^* there exists a node k such that the load to the left of k is smaller than the demand to the right of k , then again there must exist a task starting to the right of k and, since their demands are at most $\pi' - h_{\max}(\mathcal{I})$, left-shifting such task does not violate the threshold of π' which is a contradiction.

2. Suppose there exists a task not satisfying the claim, and let i be the leftmost such task in $P'(\cdot)$. It is easy to see that i cannot start at the leftmost edge and also that the demand on the edge just to the left of $P'(i)$ is larger than $\pi' - h(i)$ as otherwise a left-shifting of i is possible. Due to i being the leftmost task, no task i' can finish just to the left of $P'(i)$, as otherwise the number of edges before $P'(i)$ would be the sum of some widths in \mathcal{I} plus $w(i')$, thus fulfilling the claim for i . This implies that every task using the edge just to the left of $P'(i)$ must also use edge $P'(i)$. But then the total demand just to the left of $P'(i)$ would be at most the total demand on $P'(i)$ minus $h(i)$, which is at most $\pi' - h(i)$. ◀

We can now proceed with the proof of Lemma 2, where at some point in the proof we will make use of the following concentration bound which was proved in [11].

► **Lemma 20.** *[Calinescu et al. [11]] Let X_1, X_2, \dots, X_n be independent random variables and let $0 \leq \beta_1, \beta_2, \dots, \beta_n \leq 1$ be real numbers, where for each $i = 1, 2, \dots, n$, $X_i = \beta_i$ with probability p_i and $X_i = 0$ otherwise. Let $X = \sum_{i=1}^n X_i$ and $\mu = \mathbb{E}[X]$. Then*

20:22 Approximation Algorithms for Demand Strip Packing

1. The variance of X , $\sigma^2(X)$, is at most μ , and
2. For any $0 < \lambda < \sqrt{\mu}$, $\mathbb{P}[X > \mu + \lambda\sqrt{\mu}] < e^{-\frac{\lambda^2}{2}(1-\lambda/\sqrt{\mu})}$.

Proof of Theorem 2. Let $\delta > 0$ be a constant that we will specify later. We will partition the tasks into two sets according to their widths: we will say that a task i is *horizontal* if $w(i) > \delta \cdot W$ and otherwise we will say it is *narrow*. Consider by now only the horizontal tasks in \mathcal{I} , and assume that the value OPT is known. Thanks to Lemma 19, by computing an OPT -left-pushing of the optimal solution, we know there exists a set $\mathcal{E}_H \subseteq E$ that can be computed in polynomial time such that the starting edge of every task belongs to \mathcal{E}_H . Indeed, edges in \mathcal{E}_H correspond to the sum of widths of some horizontal tasks, implying that the number of widths in the sum must be at most $\frac{1}{\delta}$. Hence, all the possible starting edges are of the form $\sum_{i \in \mathcal{I}'} w(i)$ where $|\mathcal{I}'| \leq \frac{1}{\delta}$. The set \mathcal{E}_H consisting of these edges has size at most $n^{1/\varepsilon-1}$ and can clearly be computed in polynomial time.

With the following integer program we can compute a feasible solution corresponding to a OPT -left-pushing of some scheduling for these tasks. We define a variable $x_{i,k}$ for each task i and starting edge k in the previously computed set \mathcal{E}_H (if task i cannot be scheduled starting at edge k this variable is not considered):

$$\begin{aligned}
 \min \quad & \lambda \\
 \text{s.t.} \quad & \sum_{k \in \mathcal{E}_H} x_{i,k} = 1 \quad \forall i \text{ horizontal} \\
 & \sum_{i \text{ hor.}} \sum_{k' \in \mathcal{E}_H(i,q)} h(i) \cdot x_{i,k'} \leq OPT \quad \forall q \in \mathcal{E}_H \\
 & x_{i,k} \in \{0,1\} \quad \forall i \text{ horizontal}, k \in \mathcal{E}_H,
 \end{aligned}$$

where, given $i \in \mathcal{I}$ and $q \in \{1, \dots, W\}$, $\mathcal{E}_H(i, e)$ is the set of edges $k \in \mathcal{E}_H$ such that, if i has k as starting edge, then it uses edge e . In other words, the second family of constraints is ensuring that the total demand of the constructed solution is at most OPT in every edge (which can be done with polynomially many constraints thanks to the size of \mathcal{E}_H).

We will consider the canonical linear relaxation of the formulation, and let \vec{x} be an optimal solution to this LP (which can be computed in polynomial time). In order to derive a feasible solution we will use *Randomized Rounding with Alterations*, a technique previously used in similar settings for Packing and Scheduling problems [11, 36, 2]. In a first stage, for each task i , we will sample one starting edge k according to the probability distribution induced by $\{x_{i,k}\}_{k \in \mathcal{E}_H}$. Now, in a second stage, we scan the starting edges k from left to right, and the sampled tasks i starting at node k according to the sample in any order, and we add i to the current solution as long as the obtained peak is no more than $(1 + \varepsilon)OPT$. Observe that this is a dependent rounding where each task i is finally scheduled in the solution with marginal probability at most $x_{i,k}$.

Suppose we are applying the previous procedure, and let k be a fixed edge in that order. Let $\tilde{X}_{i,k} \in \{0,1\}$ be equal to 1 if and only if i is scheduled starting at edge k in the first stage, and similarly we define $\tilde{Y}_{i,k}$ to be 1 if and only if i is scheduled starting at edge k in the second stage. Notice that $\tilde{Y}_{i,k} \leq \tilde{X}_{i,k}$ deterministically. By stochastic domination, we have that

$$\mathbb{P} \left[\sum_{i \text{ hor.}} \sum_{k \in \mathcal{E}_H(i,q)} \tilde{Y}_{i,k} \cdot h(i) > (1 + \varepsilon)OPT \right] \leq \mathbb{P} \left[\sum_{i \text{ hor.}} \sum_{k \in \mathcal{E}_H(i,q)} \tilde{X}_{i,k} \cdot h(i) > (1 + \varepsilon)OPT \right].$$

To upper bound the latter quantity we will consider two cases:

- If $\mu \leq \frac{3}{4\delta}$, then we can use Chebyshev's inequality for the variable $Z := \sum_{i \text{ hor.}} \frac{\tilde{X}_{i,k} \cdot h(i)}{\delta OPT}$ (notice that thanks to Lemma 20 it holds that $\sigma(Z) \leq \sqrt{\mu}$), from where we obtain that

$$\begin{aligned} \mathbb{P} \left[\sum_{i \text{ hor.}} \sum_{k \in \mathcal{E}_H(i,q)} \tilde{X}_{i,k} \cdot h(i) > (1 + \varepsilon)OPT \right] &= \mathbb{P} \left[Z > \frac{1 + \varepsilon}{\delta} \right] \\ &\leq \mathbb{P} \left[|Z - \mu| > \left(\frac{1 + \varepsilon}{\delta} - \frac{3}{4\delta} \right) \cdot \frac{\sigma(Z)}{\sqrt{\mu}} \right] \\ &\leq \frac{16\mu\delta^2}{(1 + 4\varepsilon)^2} \leq \varepsilon \end{aligned}$$

for $\delta \leq \frac{\varepsilon}{4}$.

- If $\mu > \frac{3}{4\delta}$, we first set $\lambda = \frac{1+\varepsilon-\mu\delta}{\delta\sqrt{\mu}}$ so that $\mu + \lambda\sqrt{\mu} = \frac{1+\varepsilon}{\delta}$. Notice that $\mu = \sum_{i \text{ hor.}} \frac{x_{i,k} \cdot h(i)}{\delta OPT} \leq \frac{1}{\delta}$ due to the constraints in the LP.

Now, it is not difficult to see that λ is decreasing as a function of μ , implying that $\lambda \geq \frac{1+4\varepsilon}{\sqrt{12\delta}}$. Furthermore, we have that $1 - \frac{\lambda}{\sqrt{\mu}} = 2 - \frac{1+\varepsilon}{\delta\mu} \geq \frac{2}{3}$, and thus also $\lambda < \sqrt{\mu}$. Now we can use Lemma 20 applied to the variables $\{X_{i,k}h(i)/(\delta OPT)\}_{i \text{ hor.}}$ and their sum Z and obtain

$$\begin{aligned} \mathbb{P} \left[\sum_{i \text{ hor.}} \sum_{k \in \mathcal{E}_H(i,q)} \tilde{X}_{i,k} \cdot h(i) > (1 + \varepsilon)OPT \right] &= \mathbb{P}[Z > \mu + \lambda\sqrt{\mu}] \\ &< e^{-\frac{\lambda^2}{2}(1-\lambda/\sqrt{\mu})} \\ &< e^{-\frac{2}{9} \frac{(1+4\varepsilon)^2}{12\delta}} \leq \varepsilon \end{aligned}$$

for $\delta \leq \frac{(1+4\varepsilon)^2}{54} \ln \frac{1}{\varepsilon}$.

This implies that we get a solution with peak at most $(1 + \varepsilon)OPT$ and the probability that a task is not scheduled is at most ε . As a consequence, in expectation the total area of tasks that were not placed is at most $\varepsilon W \cdot OPT$, and hence using Markov's inequality we get that the probability that these tasks have area larger than $2\varepsilon W \cdot OPT$ is at most $\frac{1}{2}$. Thus, if the area of these tasks is at most $2\varepsilon W \cdot OPT$ and since their heights are at most $\delta \cdot OPT$, we can place them into a rectangular region of height $4\varepsilon OPT$ and width W using Corollary 7. If the area guarantee is not satisfied then we repeat the whole process to ensure it as, in expectation, a constant number of times only is required.

Now we will include the set \mathcal{N} of narrow tasks into the solution by applying Lemma 12 with parameter $\pi = (1 + 5\varepsilon)OPT$. Consider a π -left-pushing of the solution. Thanks to Lemma 19, there exists a node t^* such that the obtained schedule is (π, t^*) -sorted. Furthermore, it is not difficult to see that $\pi \geq (1 + 4\varepsilon)\frac{\alpha(\mathcal{I})}{W} + h_{\max}(\mathcal{I}')$ and $w_{\max}(\mathcal{N}) \leq \varepsilon W \leq \frac{4\varepsilon}{2(4\varepsilon+1)}$ for $\varepsilon \leq 1/4$, hence satisfying the requirements of the lemma. This way, we obtain a feasible scheduling with peak at most $(1 + 5\varepsilon)OPT$.

Finally, in order to avoid knowing the value of OPT , we can approximately guess it using any constant approximation (such as Corollary 7) and define a constant number of candidates. \blacktriangleleft