Non-preemptive Scheduling in a Smart Grid Model and its Implications on Machine Minimization[*]

Fu-Hong Liu[1], Hsiang-Hsuan Liu[1,2], and Prudence W.H. Wong[2]

[1]Department of Computer Science, National Tsing Hua University, 101 Kuang Fu Road, Hsinchu, Taiwan, {fhliu,hhliu}@cs.nthu.edu.tw
[2]Department of Computer Science, University of Liverpool, Liverpool, UK, {hhliu,pwong}@liverpool.ac.uk

August 7, 2017

Abstract

We study a scheduling problem arising in demand response management in smart grid. Consumers send in power requests with a flexible feasible time interval during which their requests can be served. The grid controller, upon receiving power requests, schedules each request within the specified interval. The electricity cost is measured by a convex function of the load in each timeslot. The objective is to schedule all requests with the minimum total electricity cost. Previous work has studied cases where jobs have unit power requirement and unit duration. We extend the study to arbitrary power requirement and duration, which has been shown to be NP-hard. We give the first online algorithm for the general problem, and prove that the problem is fixed parameter tractable. We also show that the online algorithm is asymptotically optimal when the objective is to minimize the peak load. In addition, we observe that the classical non-preemptive machine minimization problem is a special case of the smart grid problem with min-peak objective, and show that we can solve the non-preemptive machine minimization problem asymptotically optimally.

1 Introduction

We study a scheduling problem arising in "demand response management" in smart grid [17, 22, 23, 35, 53]. The electrical smart grid is one of the major challenges in the 21st century [15, 47, 48]. The smart grid [18, 38] is a power grid system that makes power generation, distribution and consumption more efficient through information and communication technologies against the traditional power system. Peak demand hours happen only for a short duration, yet makes existing electrical grid less efficient. It has been noted in [8] that in the US power grid, 10% of all generation assets and 25% of distribution infrastructure are required for less than 400 hours per year, roughly 5% of the time [48]. *Demand response management* attempts to overcome this problem by shifting users' demand to off-peak hours in order to reduce peak load [7, 27, 34, 37, 40, 43]. Research initiatives in the area include [25, 33, 41, 46].

The electricity grids supports demand response mechanism and obtains energy efficiency by organizing customer consumption of electricity in response to supply conditions. It is demonstrated in [35] that demand response is of remarkable advantage to consumers, utilities, and society. Effective demand load management brings down the cost of operating the grid, as well as energy generation and distribution [34]. Demand response management is not only

---

[*]A preliminary version of this paper appeared in Proceedings of the 27th International Symposium on Algorithms and Computation, ISAAC 2016 [31] and some results are improved in this version.

advantageous to the supplier but also to the consumers as well. It is common that electricity supplier charges according to the generation cost, i.e., the higher the generation cost the higher the electricity price. Therefore, it is to the consumers' advantage to reduce electricity consumption at high price and hence reduce the electricity bill [43].

The smart grid operator and consumers communicate through smart metering devices [28,38]. A consumer sends in a power request with the power requirement (cf. height of request), required duration of service (cf. width of request), and the time interval that this request can be served (giving some flexibility). For example, a consumer may want the dishwasher to operate for one hour during the periods from 8am to 11am. The grid operator upon receiving requests has to schedule them in their respective time intervals using the minimum energy cost. The *load* of the grid at each timeslot is the sum of the power requirements of all requests allocated to that timeslot. The *electricity cost* is modeled by a convex function on the load, in particular we consider the cost to be the α-th power of the load, where $\alpha > 1$ is some constant. Typically, α is small, e.g., $\alpha = 2$ [14, 44].

Previous work. Koutsopoulos and Tassiulas [27] has formulated a similar problem to our problem where the cost function is piecewise linear. They show that the problem is NP-hard, and their proof can be adapted to show the NP-hardness of the general problem studied in this paper [6]. Burcea et al. [6] gave polynomial time optimal algorithms for the case of unit height (cf. unit power requirement) and unit width (cf. unit duration). Feng et al. [19] have claimed that a simple greedy algorithm is 2-competitive for the unit case and $\alpha = 2$. However, as to be described below in Lemma 4, there is indeed a counter example that the greedy algorithm is at least 3-competitive. This implies that it is still an open question to derive online algorithms for the problem. Salinas et al. [43] considered a multi-objective problem to minimize energy consumption cost and maximize some utility. A closely related problem is to manage the load by changing the price of electricity over time [7, 16, 37, 39]. Heuristics have also been developed for demand side management [34]. Other aspects of smart grid have also been considered, e.g., communication [8, 29, 30, 32], security [32, 36]. Reviews of smart grid can be found in [17, 22, 23, 35, 53].

The main combinatorial problem we defined in this paper has analogy to the traditional load balancing problem [3] and machine minimization problem [9, 12, 13, 42] but the main differences are the objective being maximum load and jobs are unit height [9, 12, 13, 42]. Minimizing maximum load has also been looked at in the context of smart grid [1, 26, 45, 50, 51], some of which further consider allowing reshaping of the jobs [1, 26]. As to be discussed in Section 2, our problem is more difficult than minimizing the maximum load. Our problem also has resemblance to the dynamic speed scaling problem [2, 5, 49] and our algorithm has employed some techniques there.

As to be discussed, our problem is closely related to the non-preemptive machine minimization problem [12, 13], which has been claimed to be solved optimally in asymptotically sense for the online setting [42]. We provide an alternative asymptotically optimal competitive algorithm for the non-preemptive machine minimization problem. More precisely, we show that our algorithm for the smart grid problem can also solve the non-preemptive machine minimization problem with asymptotically optimal competitive ratio. A more detailed discussion is in Section 7.

Our contribution. In this paper, we consider a demand response optimization problem minimizing the total electricity cost and study its relation with other scheduling problems. We propose the first online algorithm for the general problem with worst case competitive ratio, which is polylogarithm in the max-min ratio of the duration of jobs (Theorem 25 in Section 4); and give a lower bound for any online algorithm. Interestingly, the ratio depends on the max-min width ratio but not the max-min height ratio. The algorithm is based on an $O(1)$-competitive online algorithm for jobs with uniform duration (Section 3). We also propose $O(1)$-competitive online algorithms for some special cases (Section 5). In addition, we show that the problem is fixed parameter tractable by proposing the first fixed parameter exact algorithms for the

Width	Height	Ratio
Unit	Arbitrary	$2^\alpha \cdot (8(e + e^2)^\alpha + 1)$-competitive
		$2^{\alpha+1}$-approximate
Uniform	Arbitrary	$12^\alpha \cdot (8(e + e^2)^\alpha + 1)$-competitive
Arbitrary	Arbitrary	$\Theta(\log^\alpha(\frac{w_{\max}}{w_{\min}}))$-competitive
Unit	Uniform	$\min((4\alpha)^\alpha/2 + 1, 2^\alpha \cdot (8(e + e^2)^\alpha + 1))$-competitive
Arbitrary	Uniform	$((8\alpha)^\alpha/2 + 2^\alpha)$-competitive
		agreeable deadline

Table 1: Summary of our online results or total electricity cost.

problem; and derive lower bounds on the running time (Section 6). Table 1 gives a summary of our results. Interestingly, our online algorithm and exact algorithms depend on the variation of the job widths but not the variation of the job heights.

We further show that our online algorithms and exact algorithms can be adapted to the objective of minimizing the peak electricity cost, as well as the related problem of non-preemptive machine minimization. Our online algorithms are asymptotically optimal for both problems (Section 7.1), with competitive ratio being logarithm in the max-min ratio of the job duration. In addition, we show that both problems are fixed-parameter tractable (Section 7.2).

Technically speaking, our online algorithms are based on identifying a relationship with the dynamic speed (voltage) scaling (DVS) problem [49]. The main challenge, even when jobs have uniform width or uniform height, is that in time intervals where the "workload" is low, the optimal DVS schedule may have much lower cost than the optimal GRID schedule because jobs in DVS schedules can effectively be stretched as flat as possible while jobs in GRID schedules have rigid duration and cannot be stretched. In such case, it is insufficient to simply compare with the optimal DVS schedule. Therefore, our analysis is divided into two parts: for high workload intervals, we compare with the optimal DVS schedule; and for low workload intervals, we directly compare with the optimal GRID schedule via a lower bound on the total workload over these intervals (Lemmas 6 and 30). For jobs with arbitrary width, we adopt the natural approach of classification based on job width. We then align the "feasible interval" of each job in a more uniform way so that we can use the results on uniform width (Lemma 20).

In designing exact algorithms we use interval graphs to represent the jobs and the important notion maximal cliques to partition the time horizon into disjoint windows. Such partition usually leads to optimal substructures; nevertheless, non-preemption makes it trickier and requires a smart way to handle jobs spanning multiple windows. We describe how to handle such jobs without adding a lot of overhead.

Organization of the paper. We define the problem and provide some basic observations in Section 2. The online algorithms for uniform time duration and arbitrary power requirement are developed in Section 3 and are extended for solving the general case in Section 4. The lower bound of online algorithms is provided in Section 4.3. Several special cases regarding uniform power requirement are discussed in Section 5. We design fixed-parameter exact algorithms in Section 6 and derive a lower bound for the running time in Section 6.3. In Section 7, we extend our online and exact algorithms to the objective of maximum load and the related non-preemptive machine minimization problem. We conclude the paper in Section 8.

2   Definitions and preliminaries

The input. The time is labeled from 0 to τ and we consider events (release time, deadlines) occurring at integral time. We call the unit time $[t, t+1]$ *timeslot t*. We denote by $\mathcal{J}$ a set of input jobs in which each job J comes with *release time $r(J)$, deadline $d(J)$, width $w(J)$*

representing the duration required by J, and *height* $h(J)$ representing the power required by J. We assume $r(J)$, $d(J)$, $w(J)$, and $h(J)$ are integers. The *feasible interval*, denoted by $I(J)$, is defined as the interval $[r(J), d(J))$ and we say that J is *available* during $I(J)$. We denote by $|I|$ the length of an interval I, i.e., $|I| = t_2 - t_1$ where $I = [t_1, t_2)$. We define the *density* of J, denoted by $\text{den}(J)$, to be $\frac{h(J) \cdot w(J)}{|I(J)|}$. Roughly speaking, the density signifies the average load required by the job over its feasible interval. We then define the "average" load at any time t as $\text{avg}(t) = \sum_{J:t \in I(J)} \text{den}(J)$. In our analysis, we have to distinguish timeslots with high and low average load. Therefore, for any $h > 0$, we define $\mathcal{I}_{>h}$ and $\mathcal{I}_{\leq h}$ to be set of timeslots where the average load $\text{avg}(t)$ is larger than h and at most h, respectively. Note that $\mathcal{I}_{>h}$ and $\mathcal{I}_{\leq h}$ do not need to be contiguous.

In Section 4, we consider an algorithm that classifies jobs according to their widths. To ease discussion, we let $w_{\max}$ and $w_{\min}$ be the maximum and minimum width over all jobs, respectively. We further define the max-min ratio of width, denoted by K, to be $K = \frac{w_{\max}}{w_{\min}}$. Without loss of generality, we assume that $w_{\min} = 1$. We say that a job J is in *class* C_p if and only if $2^{p-1} < w(J) \leq 2^p$ for any $0 \leq p \leq \lceil \log K \rceil$.

Feasible schedule. A *feasible* schedule S has to assign for each job J a *start time* $st(S, J) \in \mathbb{Z}$ meaning that J runs during $[st(S, J), et(S, J))$, where the *end time* $et(S, J) = st(S, J) + w(J)$, and $[st(S, J), et(S, J)) \subseteq I(J)$. Note that this means preemption is not allowed. The *load* of S at time t, denoted by $\ell(S, t)$ is the sum of the height (power request) of all jobs running at t, i.e., $\ell(S, t) = \sum_{J:t \in [st(S,J), et(S,J))} h(J)$. We drop S and use $\ell(t)$ when the context is clear. For any algorithm $\mathcal{A}$, we use $\mathcal{A}(\mathcal{J})$ to denote the schedule of $\mathcal{A}$ on $\mathcal{J}$. We denote by $\mathcal{O}$ the optimal algorithm.

The cost of a schedule S is the sum of the α-th power of the load over all time, for a constant $\alpha > 1$, i.e., $\text{cost}(S) = \sum_t (\ell(S, t))^\alpha$. For a set of timeslots $\mathcal{I}$ (not necessarily contiguous), we denote by $\text{cost}(S, \mathcal{I}) = \sum_{t \in \mathcal{I}} (\ell(S, t))^\alpha$. Our goal is to find a feasible schedule S such that $\text{cost}(S)$ is minimized. We call this the GRID problem.

Online algorithms. In this paper, we consider online algorithms, where the job information is only revealed at the time the job is released; the algorithm has to decide which jobs to run at the current time without future information and decisions made cannot be changed later. Let $\mathcal{A}$ be an online algorithm. We say that $\mathcal{A}$ is c-competitive if for all input job sets $\mathcal{J}$, we have $\text{cost}(\mathcal{A}(\mathcal{J})) \leq c \cdot \text{cost}(\mathcal{O}(\mathcal{J}))$. In particular, we consider non-preemptive algorithms where a job cannot be preempted to resume/restart later.

Special input instances. We consider various special input instances. A job J is said to be unit-width (resp. unit-height) if $w(J) = 1$ (resp. $h(J) = 1$). A job set is said to be uniform-width (resp. uniform-height) if the width (resp. height) of all jobs are the same. A job set is said to have *agreeable deadlines* if for any two jobs J_1 and J_2, $r(J_1) \leq r(J_2)$ implies $d(J_1) \leq d(J_2)$.

Relating to the speed scaling problem. The GRID problem resembles the dynamic speed scaling (DVS) problem [49] and we are going to refer to three algorithms for the DVS problem, namely, the $\mathcal{YDS}$ algorithm which gives an optimal algorithm for the DVS problem, the online algorithms called $\mathcal{BKP}$ and $\mathcal{AVR}$. We first recap the DVS problem and the associated algorithms. In the DVS problem, jobs come with release time $r(J)$, deadline $d(J)$, and a work requirement $p(J)$. A processor can run at speed $s \in [0, \infty)$ and consumes energy in a rate of s^α, for some $\alpha > 1$. The objective is to complete all jobs by their deadlines using the minimum total energy. The main differences of DVS problem to the GRID problem include (i) jobs in DVS can be preempted while preemption is not allowed in our problem; (ii) as processor speed in DVS can scale, a job can be executed for varying time duration as long as the total work is completed while in our problem a job must be executed for a fixed duration given as input; (iii) the work requirement $p(J)$ of a job J in DVS can be seen as $w(J) \times h(J)$ for the corresponding job in GRID.

With the resemblance of the two problems, we make an observation about their optimal algorithms. Let $\mathcal{O}_D$ and $\mathcal{O}_G$ be the optimal algorithm for the DVS and GRID problem, respectively.

Given a job set $\mathcal{J}_G$ for the GRID problem, we can convert it into a job set $\mathcal{J}_D$ for DVS by keeping the release time and deadline for each job and setting the work requirement of a job in $\mathcal{J}_D$ to the product of the width and height of the corresponding job in $\mathcal{J}_G$. Then we have the following observation.

Observation 1. *Given any schedule S_G for $\mathcal{J}_G$, we can convert S_G into a feasible schedule S_D for $\mathcal{J}_D$ such that $cost(S_D(\mathcal{J}_D)) \leq cost(S_G(\mathcal{J}_G))$; implying that $cost(\mathcal{O}_D(\mathcal{J}_D)) \leq cost(\mathcal{O}_G(\mathcal{J}_G))$.*

Proof. Consider any feasible schedule S_G. At timeslot t, suppose there are k jobs scheduled and their sum of heights is H. The schedule for S_D during timeslot t can be obtained by running the processor at speed H and the jobs time-share the processor in proportion to their height. This results in a feasible schedule with the same cost and the observation follows. $\square$

It is known that the online algorithm $\mathcal{AVR}$ for the DVS problem is $\frac{(2\alpha)^\alpha}{2}$-competitive [49]. Basically, at any time t, $\mathcal{AVR}$ runs the processor at a speed which is the sum of the densities of jobs that are available at t. By Observation 1, we have the following corollary. Note that it is not always possible to convert a feasible schedule for the DVS problem to a feasible schedule for the GRID problem easily. Therefore, the corollary does not immediately solve the GRID problem but as to be shown it provides a way to analyze algorithms for GRID.

Corollary 2. *For any input $\mathcal{J}_G$ and the corresponding input $\mathcal{J}_D$, $cost(\mathcal{AVR}(\mathcal{J}_D)) \leq \frac{(2\alpha)^\alpha}{2} \cdot cost(\mathcal{O}_G)$.*

The online algorithm $\mathcal{BKP}$ proposed by Bansal et al. [4] for DVS problem is $8e^\alpha$-competitive with respect to total cost. Let $\ell(\mathcal{BKP}, t)$ denote the speed of $\mathcal{BKP}$ at time t. $\ell(\mathcal{BKP}, t) = \max_{t' > t} \frac{p(t, [et - (e-1)t', t'))}{t' - t}$ where $p(t, I)$ denotes the total work of jobs J with $I(J) \subseteq I$ and $r(J) \leq t$. That is, $\mathcal{BKP}$ chooses the interval $I^\star = [t'', t']$ which has maximal released average total work and $(t' - t'') : (t' - t) = e : 1$ and uses $\frac{p(t, I^\star)}{|I^\star|/e}$ as the speed at t. By Observation 1 we have the following corollary:

Corollary 3. *For any input $\mathcal{J}_G$ and the corresponding input $\mathcal{J}_D$, $cost(\mathcal{BKP}(\mathcal{J}_D)) \leq 8e^\alpha \cdot cost(\mathcal{O}(\mathcal{J}_D)) \leq 8e^\alpha \cdot cost(\mathcal{O}(\mathcal{J}_G))$.*

Remark: One may consider the non-preemptive DVS problem as the reference of the GRID problem. However, given a job set $\mathcal{J}_G$ and the corresponding $\mathcal{J}_D$, $cost(\mathcal{O}_D(\mathcal{J}_D))$ may not necessarily lower than $cost(\mathcal{O}_G(\mathcal{J}_G))$, where $\mathcal{O}_D$ here is the optimal algorithm for non-preemptive DVS. There is an instance shows the optimal cost of GRID is smaller. The instance contains two jobs. One has release time 0, deadline 3, width 3 and height 1. The other has release time 1, deadline 2, width 1 and height 1. Both jobs can only schedule at their release time in GRID since their widths are the same as the lengths of their feasible intervals. The optimal cost of GRID is $1^\alpha + 2^\alpha + 1^\alpha = 2^\alpha + 2$. Whereas the optimal cost of non-preemptive DVS is $2^\alpha + 2^\alpha = 2 \cdot 2^\alpha$. This is because the schedule uses speed 2 and runs the longer job with 1.5 time units and the shorter job with 0.5 time units. The optimal cost of GRID is lower when $\alpha > 1$. Therefore, it is unclear how we may use the results on non-preemptive DVS problem and so we would stick with the preemptive DVS algorithms.

Relating to minimizing maximum cost. The problem of minimizing maximum cost over time (min-max) has been studied before [50]. We note that there is a polynomial time reduction of the decision version of the min-max problem to that of the min-sum problem (the GRID problem we study in this paper) for a large enough α. In particular, one can show that with $\alpha > (\tau - 1)(2 \sum_{J \in \mathcal{J}} h(J) + 1)$, the maximum load would dominate the load in other timeslots and we would be able to solve the min-max problem if we have a solution for the min-sum problem on α.

On the other hand, minimizing the maximum cost does not necessarily minimize the total cost. For example, consider an input of three jobs J_1, J_2 and J_3 where $I(J_1) = [0, 2^\alpha)$, $h(J_1) = 1$,

$w(J_1) = 2^\alpha$; $I(J_2) = [2^\alpha, 2^\alpha + 1)$, $h(J_2) = 3$, $w(J_2) = 1$; and $I(J_3) = [0, 2^{\alpha+1})$, $h(J_3) = 1$, $w(J_3) = 2^\alpha$. Note that only J_3 has flexibility where it can be scheduled. To minimize the maximum cost over time, we would schedule J_3 to start at time 0 and achieve a maximum load of 3. This gives a total cost of $2^\alpha \cdot 2^\alpha + 3^\alpha = 4^\alpha + 3^\alpha$. However, to minimize the total cost, we would schedule J_3 to start at time 2^α giving a total cost of $2^\alpha + 4^\alpha + (2^\alpha - 1) = 4^\alpha + 2^{\alpha+1} - 1$, which is smaller than $4^\alpha + 3^\alpha$ when $\alpha > 1$.

Lower bound on Greedy. In [19], the greedy algorithm that assigns a job to a timeslot with the minimum load is considered. It is claimed in the paper that the greedy algorithm is 2-competitive on the online-list model and for the case where the load of a timeslot t is $\ell(t)^2$, jobs are of unit length and height and the feasible timeslots of a job is a set of (non-contiguous) timeslots that the job can be assigned to. We show a counter-example to this claim and show that Greedy is at least 3-competitive. This implies that it is still an open question to derive online algorithms for the GRID problem.

Lemma 4. *Greedy is no better than 3-competitive for the online-list model when $\alpha = 2$.*

Proof. Let k be an arbitrarily large integer. The adversary works in k rounds and all the jobs released are of width and height 1. In the i-th round, where $1 \le i < k$, the adversary releases 2^{k-i} jobs; and in the k-th round (the final one), the adversary releases two jobs. In the first round, the feasible timeslots of each job released are $[1, 2^k]$. In the i-th round, where $2 \le i \le k$, the feasible timeslots of each job released are all the timeslots that Greedy has assigned jobs in the $(i-1)$-th round. We claim that the total cost of Greedy is $3 \cdot 2^k - 4$ and the total cost of the optimal algorithm is 2^k. Therefore, the competitive ratio of Greedy is arbitrarily close to 3 with an arbitrarily large integer k.

We first analyze Greedy. Since Greedy always assigns to a timeslot with the minimum load, in the first round, Greedy assigns jobs to 2^{k-1} timeslots with each job to a different timeslot. These 2^{k-1} timeslots will be the feasible timeslots for the 2^{k-2} jobs in the second round. Using a similar argument, we can see that in each round, the number of feasible timeslots is twice the number of jobs released in that round. In addition, before the i-th round, the load of each feasible timeslot is $i-1$ and Greedy adds a load of 1 to each timeslot that it assigns a job, making the load become i. Therefore, the total cost of Greedy is $\sum_{i=1}^{k-2}(i^2 \cdot 2^{k-i-1}) + k^2 \cdot 2 = 3 \cdot 2^k - 4$. On the other hand, we can assign jobs released in a round to the timeslots that are not feasible timeslots for later rounds since in the i-th round, the number of feasible timeslots is 2^{k-i+1} and the number of jobs released is 2^{k-i}. Therefore, in the optimal schedule, the load of each timeslot is exactly 1 and the total cost is 2^k. $\qquad\square$

3 Online algorithm for uniform width jobs

To handle jobs of arbitrary width and height, we first study the case when jobs have uniform width (all jobs have the same width $w \ge 1$). The proposed algorithm $\mathcal{UV}$ (Section 3.2) is based on a further restricted case of unit width, i.e., $w = 1$ (Section 3.1).

3.1 Unit width and arbitrary height

In this section, we consider jobs with unit width and arbitrary height. We present an online algorithm $\mathcal{V}$ which makes reference to an arbitrary feasible online algorithm for the DVS problem, denoted by $\mathcal{R}$. In particular, we require that the speed of $\mathcal{R}$ remains the same during any integral timeslot, i.e., in $[t, t+1)$ for all integers t. Note that when jobs have integral release times and deadlines, many known DVS algorithms satisfy this criteria, including $\mathcal{YDS}$, $\mathcal{BKP}$, and $\mathcal{AVR}$.

Recall in Section 2 how a job set for the GRID problem is converted to a job set for the DVS problem. We simulate a copy of $\mathcal{R}$ on the converted job set and denote the speed used by $\mathcal{R}$ at t as $\ell(\mathcal{R}, t)$. Our algorithm makes reference to $\ell(\mathcal{R}, t)$ but not the jobs run by $\mathcal{R}$ at t.

Algorithm $\mathcal{V}$. For each timeslot t, we schedule jobs to start at t such that $\ell(\mathcal{V}, t)$ is at least $\ell(\mathcal{R}, t)$ or until all available jobs have been scheduled. Jobs are chosen in an EDF manner.

Analysis. We note that since $\mathcal{V}$ makes decision at integral time and jobs have unit width, each job is completed before any further scheduling decision is made. In other words, $\mathcal{V}$ is non-preemptive. To analyze the performance of $\mathcal{V}$, we first note that $\mathcal{V}$ gives a feasible schedule (Lemma 5), and then analyze its competitive ratio (Theorem 7).

Lemma 5. $\mathcal{V}$ *gives a feasible schedule.*

Proof. Let $\ell(S, \mathcal{I})$ denote the total work done by schedule S in $\mathcal{I}$. That is, $\ell(S, \mathcal{I}) = \sum_{t \in \mathcal{I}} \ell(S, \mathcal{I})$. According to the algorithm, for all $\mathcal{I}_t = [0, t)$, $\ell(\mathcal{V}, \mathcal{I}_t) \geq \ell(\mathcal{R}, \mathcal{I}_t)$.

Suppose on the contrary that $\mathcal{V}$ has a job J_m missing deadline at t. That is, $d(J_m) = t$ but J_m is not assigned before t. By the algorithm, for all $t' \in [0, t)$, $\ell(\mathcal{V}, t') \geq \ell(\mathcal{R}, t')$ unless there are less than $\ell(\mathcal{R}, t')$ available jobs at t' for $\mathcal{V}$. Let t_0 be the last timeslot in $[0, t)$ such that $\ell(\mathcal{V}, t_0) < \ell(\mathcal{R}, t_0)$, $r(J_m) > t_0$ since all jobs released at or before t_0 have been assigned. For all $t' \in (t_0, t)$, $\ell(\mathcal{V}, t') \geq \ell(\mathcal{R}, t')$. Also, all jobs J with $r(J) \leq t_0$ are finished by $t_0 + 1$ and jobs executed in (t_0, t) are those released after t_0. Consider set $\mathcal{J}_t$ of jobs with feasible interval completely inside $\mathcal{I} = (t_0, t)$ (note that $J_m \in \mathcal{J}_t$), $\ell(S, \mathcal{I}) \geq \sum_{J \in \mathcal{J}_t} h(J)$ for any feasible schedule S. Since $\mathcal{V}$ assigns jobs in EDF manner and is not feasible, $\ell(\mathcal{V}, \mathcal{I}) < \sum_{J \in \mathcal{J}_t} h(J)$. It follows that $\sum_{J \in \mathcal{J}_t} h(J) > \ell(\mathcal{V}, \mathcal{I}) \geq \ell(\mathcal{R}, \mathcal{I})$. It contradicts to the fact that $\mathcal{R}$ is feasible. Hence, $\mathcal{V}$ finishes all jobs before their deadlines. $\square$

Let $h_{\max}(\mathcal{V}, t)$ be the maximum height of jobs scheduled at t by $\mathcal{V}$; we set $h_{\max}(\mathcal{V}, t) = 0$ if $\mathcal{V}$ assigns no job at t. We first classify each timeslot t into two types: (i) $h_{\max}(\mathcal{V}, t) < \ell(\mathcal{R}, t)$, and (ii) $h_{\max}(\mathcal{V}, t) \geq \ell(\mathcal{R}, t)$. We denote by $\mathcal{I}_1$ and $\mathcal{I}_2$ the union of all timeslots of Type (i) and (ii), respectively. Notice that $\mathcal{I}_1$ and $\mathcal{I}_2$ can be empty and the union of $\mathcal{I}_1$ and $\mathcal{I}_2$ covers the entire time line. The following lemma bounds the cost of $\mathcal{V}$ in each type of timeslots. Recall that $\text{cost}(S, \mathcal{I})$ denotes the cost of the schedule S over the interval $\mathcal{I}$ and $\text{cost}(S)$ denotes the cost of the entire schedule.

Lemma 6. *The cost of $\mathcal{V}$ satisfies the following properties.* (i) $\text{cost}(\mathcal{V}, \mathcal{I}_1) \leq 2^{\alpha} \cdot \text{cost}(\mathcal{R})$*; and* (ii) $\text{cost}(\mathcal{V}, \mathcal{I}_2) \leq 2^{\alpha} \cdot \text{cost}(\mathcal{O})$*.*

Proof. (i) By the algorithm, $\ell(\mathcal{V}, t) < \ell(\mathcal{R}, t) + h_{\max}(\mathcal{V}, t) \leq 2 \cdot \ell(\mathcal{R}, t)$ for $t \in \mathcal{I}_1$. It follows that $\text{cost}(\mathcal{V}, \mathcal{I}_1) \leq 2^{\alpha} \cdot \sum_{t \in \mathcal{I}_1} \ell(\mathcal{R}, t)^{\alpha} = 2^{\alpha} \cdot \text{cost}(\mathcal{R}, \mathcal{I}_1) \leq 2^{\alpha} \cdot \text{cost}(\mathcal{R})$.

(ii) By convexity, $\text{cost}(\mathcal{O}) \geq \sum_J h(J)^{\alpha}$. We can see that $\text{cost}(\mathcal{O}) \geq \sum_{t \in \mathcal{I}_2} h_{\max}(\mathcal{V}, t)^{\alpha}$. According to the algorithm, $\ell(\mathcal{V}, t) < \ell(\mathcal{R}, t) + h_{\max}(\mathcal{V}, t) \leq 2 \cdot h_{\max}(\mathcal{V}, t)$ for $t \in \mathcal{I}_2$. Hence, $\text{cost}(\mathcal{V}, \mathcal{I}_2) = \sum_{t \in \mathcal{I}_2} \ell(\mathcal{V}, t)^{\alpha} \leq 2^{\alpha} \cdot \sum_{t \in \mathcal{I}_2} h_{\max}(\mathcal{V}, t)^{\alpha} \leq 2^{\alpha} \cdot \text{cost}(\mathcal{O})$. $\square$

Notice that $\text{cost}(\mathcal{V}) = \text{cost}(\mathcal{V}, \mathcal{I}_1) + \text{cost}(\mathcal{V}, \mathcal{I}_2)$ since $\mathcal{I}_1$ and $\mathcal{I}_2$ have no overlap. Together with Lemma 6 and Observation 1, we obtain the competitive ratio of $\mathcal{V}$ in the following theorem.

Theorem 7. *Algorithm $\mathcal{V}$ is $2^{\alpha} \cdot (R + 1)$-competitive, where R is the competitive ratio of the reference* DVS *algorithm $\mathcal{R}$.*

There are a number of DVS algorithms that can be used as the reference algorithm. The only requirement is that the speed of the reference algorithm within any integral interval $[t, t+1)$ for some integer t should be at most the load of the resulting online algorithm at the corresponding timeslot t. Otherwise, the feasibility of $\mathcal{V}$ cannot be guaranteed. Also, since in our online algorithm we make decision at each integral time t, it means if the load of the reference algorithm at $i + \Delta$ is larger than $\ell(\mathcal{R}, i)$ for some $0 < \Delta < 1$, our online algorithm might not be feasible.

The speed of the $\mathcal{AVR}$ and $\mathcal{YDS}$ algorithm only change at release times or deadlines of the jobs so it is valid to use $\mathcal{AVR}$ or $\mathcal{YDS}$ as a reference. Note that if we use $\mathcal{YDS}$ as the reference, the algorithm $\mathcal{V}$ is an offline algorithm since $\mathcal{YDS}$ is an offline algorithm. Unlike $\mathcal{AVR}$ and

$\mathcal{YDS}$, the speed of $\mathcal{BKP}$ within a timeslot might increase. Hence, we need to modify the $\mathcal{BKP}$ algorithm such that it can be used as the reference algorithm. In Lemma 8, we show that the speed of $\mathcal{BKP}$ in $[t, t+1)$ is bounded by a constant factor times the speed at t for any time t.

Lemma 8. *For any integral time t and a constant $0 < \Delta < 1$, $\ell(\mathcal{BKP}, t+\Delta) \leq (1+e) \cdot \ell(\mathcal{BKP}, t)$ if the release times and deadlines of jobs are integral.*

Proof. Recall that the speed of $\mathcal{BKP}$ at time t, $\ell(\mathcal{BKP}, t) = \max_I e \cdot \frac{p(t,I)}{|I|}$ where $I = [t_1, t_2)$ and $(t_2 - t_1) : (t_2 - t) = e : 1$. The proof idea is, consider the interval I chosen by $\mathcal{BKP}$ corresponding to $t + \Delta$, we can transform it into another interval I' which is one of the interval candidate for t. We show that $e \cdot \frac{p(t,I')}{|I'|}$ is at least $\frac{1}{1+e}$ times of the speed of $\mathcal{BKP}$ at $t + \Delta$.

Assume that at time $t + \Delta$, $\ell(\mathcal{BKP}, t + \Delta) = e \cdot \frac{p(t,I)}{|I|}$ where $I = [t_1, t_2)$ is chosen by $\mathcal{BKP}$. We can construct $I' = [t_1', t_2)$ such that $(t_2 - t_1') : (t_2 - t) = e : 1$ by setting $t_1' = t_2 - e(t_2 - t)$. It is clear that $I \subset I'$ since the two intervals have the same right endpoint and I' is longer than I. In fact, $|I'| = e(t_2 - t) = e(t_2 - (t + \Delta)) + e\Delta = |I| + e\Delta \leq |I| + e$. Moreover, for any interval candidate, the length must be at least 1 if the release times and deadlines of the jobs are integral. Otherwise, the interval contains no jobs and the speed is 0. Hence, $|I'| \leq (1+e)|I|$. By $\mathcal{BKP}$, $\ell(\mathcal{BKP}, t) \geq e \cdot \frac{p(t,I')}{|I'|} = e \cdot \frac{p(t+\Delta, I')}{|I'|}$. The later equality holds since there is no job released between t and $t + \Delta$. Since $I \subset I'$ and $|I'| \leq (1+e)|I|$, $e \cdot \frac{p(t+\Delta, I')}{|I'|} \geq e \cdot \frac{p(t+\Delta, I)}{|I'|} \geq e \cdot \frac{p(t+\Delta, I)}{(1+e)|I|}$. Hence, $\ell(\mathcal{BKP}, t) \geq \frac{1}{1+e} \cdot \ell(\mathcal{BKP}, t + \Delta)$. $\square$

Lemma 8 implies that, although the speeds of $\mathcal{BKP}$ change within $[t, t+1)$, the speeds are bounded by $(1 + e)$ times of the speed at t. Hence, we can modify $\mathcal{BKP}$ into $\mathcal{BKP}'$ as follows: at integral time t, the speed of $\mathcal{BKP}'$, $\ell(\mathcal{BKP}', t) = (1+e)\ell(\mathcal{BKP}, t)$; at time $t' = t + \Delta$ where t is integral and $0 < \Delta < 1$, $\ell(\mathcal{BKP}', t') = \ell(\mathcal{BKP}', t)$. By the modification, the speed of $\mathcal{BKP}'$ remains the same during any integral timeslot, and $\text{cost}(\mathcal{BKP}') \leq (1 + e)^\alpha \cdot \text{cost}(\mathcal{BKP})$. As mentioned in Section 2, the $\mathcal{BKP}$ algorithm is $8 \cdot e^\alpha$-competitive. On the other hand, $\mathcal{V}$ can take an offline DVS algorithm, e.g., the optimal $\mathcal{YDS}$ algorithm, as reference and returns an offline schedule. Therefore, we have the following corollary.

Corollary 9. $\mathcal{V}$ *is* $2^\alpha \cdot (8 \cdot (e + e^2)^\alpha + 1)$-*competitive,* $2^\alpha \cdot (\frac{(2\alpha)^\alpha}{2} + 1)$-*competitive, and* $2^\alpha \cdot 2$-*approximate when the algorithm* $\mathcal{BKP}'$, $\mathcal{AVR}$, *and* $\mathcal{YDS}$ *are referenced, respectively.*

3.2 Uniform width and arbitrary height

In this section, we consider jobs with uniform width w and arbitrary height. The idea of handling uniform width jobs is to treat them as if they were unit width, however, this would mean that jobs may have release times or deadlines at non-integral time. To remedy this, we define a procedure ALIGNFI to align the feasible intervals (precisely, release times and deadlines) to the new time unit of duration w.

Let $\mathcal{J}$ be a uniform width job set. We first define the notion of "tight" and "loose" jobs. A job J is said to be *tight* if $|I(J)| < 2w$; otherwise, it is *loose*. Let $\mathcal{J}_T$ and $\mathcal{J}_L$ be the disjoint subsets of tight and loose jobs of $\mathcal{J}$, respectively. We design different strategies for tight and loose jobs. As to be shown, tight jobs can be handled easily by starting them at their release times. For any loose job, we modify it via Procedure ALIGNFI such that its release time and deadline is a multiple of w. With this alternation, we can treat the jobs as unit width and make scheduling decisions at time multiple of w.

Procedure ALIGNFI. Given a loose job set $\mathcal{J}_L$ in which $w(J) = w$ and $|I(J)| \geq 2 \cdot w$ $\forall J \in \mathcal{J}_L$. We define the procedure ALIGNFI to transform each loose job $J \in \mathcal{J}_L$ into a job J' with release time and deadline "aligned" as follows. We denote the resulting job set by $\mathcal{J}'$.

- $r(J') \leftarrow \min_{i \geq 0}\{i \cdot w \mid i \cdot w \geq r(J)\}$;

- $d(J') \leftarrow \max_{i \geq 0}\{i \cdot w \mid i \cdot w \leq d(J)\}$.

Observation 10. *For any job $J \in \mathcal{J}_L$ and the corresponding J', (i) $\frac{1}{3} \cdot |I(J)| < |I(J')| \leq |I(J)|$; (ii) $|I(J')| \geq w$; (iii) $I(J') \subseteq I(J)$.*

Notice that after ALIGNFI, the release time and deadline of each loose job are aligned to timeslot $i_1 \cdot w$ and $i_2 \cdot w$ for some integers $i_1 < i_2$. By Observation 10, a feasible schedule of J' is also a feasible schedule of J. Furthermore, after ALIGNFI all jobs are released at time which is a multiple of w. Hence, the job set $\mathcal{J}'$ can be treated as job set with unit width, where each unit has duration w instead of 1.

As a consequence of altering the feasible intervals, we introduce two additional procedures that convert associated schedules. Given a schedule S for job set $\mathcal{J}_L$, ALIGNSCH converts it to a schedule S' for the corresponding job set $\mathcal{J}'$. The other procedure FREESCH takes a schedule S' for a job set $\mathcal{J}'$ and converts it to a schedule S for $\mathcal{J}_L$.

Transformation ALIGNSCH. ALIGNSCH transforms S into S' by shifting the execution interval of every job $J \in \mathcal{J}_L$.

- $st(S', J') \leftarrow \min\{d(J') - w(J'), \min_{i \geq 0}\{i \cdot w \mid i \cdot w \geq st(S, J)\}\}$;

- $et(S', J') \leftarrow st(S', J') + w(J')$.

Observation 11. *Consider any schedule S for $\mathcal{J}_L$ and the schedule S' for $\mathcal{J}'$ constructed by ALIGNSCH. The following properties hold:* (i) *For any job $J \in \mathcal{J}_L$ and the corresponding J', $st(J') > st(J) - w$ and $et(J') < et(J) + w$;* (ii) *S' is a feasible schedule for $\mathcal{J}'$; and* (iii) *At any time t, $\ell(S', t) \leq \ell(S, t) + \ell(S, t - (w - 1)) + \ell(S, t + (w - 1))$.*

Proof. (ii) By ALIGNSCH, $st(S', J') \leq d(J') - w(J')$. Also, $|[st(S', J'), et(S', J'))| = w(J')$. Hence $[st(S', J'), et(S', J')) \subseteq I(J')$. That is, S' is a feasible schedule for both $\mathcal{J}'$ and J.

(iii) By (i), $st(J') > st(J) - w$ and $et(J') < et(J) + w$ for each J. Hence, for any timeslot t, for each job J with $[st(S, J), et(S, J)) \cap [t - (w - 1), t + (w - 1)) = \emptyset$, $t \notin [st(S', J'), et(S', J'))$. On the other hand, consider the jobs J that $[st(J), et(J)) \cap [t - (w - 1), t + (w - 1)) \neq \emptyset$. Since $|[st(J), et(J))| = w$, at least one of the timeslots $t - (w - 1)$, t, or $t + (w - 1)$ is in $[st(J), et(J))$. Hence we can capture $\ell(S', t)$ by $\ell(S, t) + \ell(S, t - (w - 1)) + \ell(S, t + (w - 1))$. $\square$

Corollary 12. *Using ALIGNSCH to generate S' given S, we have $cost(S') \leq 3^\alpha \cdot cost(S)$.*

Proof. By Observation 11 (iii), $cost(S') = \sum_t \ell(S', t)^\alpha \leq \sum_t (3 \cdot \ell(S, t))^\alpha = 3^\alpha \cdot cost(S)$. $\square$

Lemma 13. *$cost(\mathcal{O}(\mathcal{J}')) \leq 3^\alpha \cdot cost(\mathcal{O}(\mathcal{J}_L))$.*

Proof. Consider set of loose jobs $\mathcal{J}_L$ with uniform width and the corresponding $\mathcal{J}'$. Given $\mathcal{O}(\mathcal{J}_L)$, there exists schedule $S(\mathcal{J}')$ generated by ALIGNSCH. By Lemma 12, $cost(S(\mathcal{J}')) \leq 3^\alpha \cdot cost(\mathcal{O}(\mathcal{J}_L))$. Hence, $cost(\mathcal{O}(\mathcal{J}')) \leq cost(S(\mathcal{J}')) \leq 3^\alpha \cdot cost(\mathcal{O}(\mathcal{J}_L))$. $\square$

Transformation FREESCH. FREESCH transforms S' into S.

- $st(S, J) \leftarrow st(S', J')$;

- $et(S, J) \leftarrow et(S', J')$.

The feasibility of S' can be easily proved by Observation 10.

Lemma 14. *Using FREESCH, we have $cost(S) = cost(S')$.*

Proof. Since the execution intervals of J and J' are the same, $\ell(S, t) = \ell(S', t)$ for all t. Hence $cost(S) = cost(S')$. $\square$

Online algorithm $\mathcal{UV}$. The algorithm takes a job set $\mathcal{J}$ with uniform width w as input and schedules the jobs in $\mathcal{J}$ as follows. Let $\mathcal{J}_T$ be the set of tight jobs in $\mathcal{J}$ and $\mathcal{J}_L$ be the set of loose jobs in $\mathcal{J}$.

1. For any tight job $J \in \mathcal{J}_T$, schedule J to start at $r(J)$.

2. Loose jobs in $\mathcal{J}_L$ are converted to $\mathcal{J}'$ by AlignFI. For $\mathcal{J}'$, we run Algorithm $\mathcal{V}$, which is defined in Section 3.1, with $\mathcal{BKP}$ as the reference DVS algorithm. Jobs are chosen in an earliest deadline first (EDF) manner.

Note that the decisions of $\mathcal{UV}$ can be made online.

Analysis of Algorithm $\mathcal{UV}$. We analyze the tight jobs and loose jobs separately. We first give an observation.

Observation 15. *For any two job sets $\mathcal{J}_x \subseteq \mathcal{J}_y$, $cost(\mathcal{O}(\mathcal{J}_x)) \leq cost(\mathcal{O}(\mathcal{J}_y))$.*

Proof. Assume on the contrary that $cost(\mathcal{O}(\mathcal{J}_y)) < cost(\mathcal{O}(\mathcal{J}_x))$, we can generate a schedule $S(\mathcal{J}_x)$ by removing jobs from $\mathcal{O}(\mathcal{J}_y)$ which are not in $\mathcal{J}_x$. It follows that $cost(S(\mathcal{J}_x)) \leq cost(\mathcal{O}(\mathcal{J}_y)) < cost(\mathcal{O}(\mathcal{J}_x))$, contradicting to the fact that $\mathcal{O}(\mathcal{J}_x)$ is optimal for $\mathcal{J}_x$. □

In the following analysis we say that interval $I = [t_1, t_2)$ is a $\mathcal{BKP}$ *interval of* t if $t \in I$ and $(t_2 - t_1) : (t_2 - t) = e : 1$. The next lemma proves the competitive ratio separately for $\mathcal{J}_T$ and $\mathcal{J}_L$.

Lemma 16. (i) $cost(\mathcal{UV}(\mathcal{J}_T)) \leq 3^\alpha \cdot cost(\mathcal{O}(\mathcal{J}))$; (ii) $cost(\mathcal{UV}(\mathcal{J}_L)) \leq 6^\alpha \cdot (8(e + e^2)^\alpha + 1) \cdot cost(\mathcal{O}(\mathcal{J}))$.

Proof. (i) We prove that any feasible schedule S for tight jobs is 3^α-competitive. We first extend jobs $J \in \mathcal{J}_T$ to J' as follows: $r(J') = r(J)$, $d(J') = d(J)$, $w(J') = d(J) - r(J)$, and $h(J') = h(J)$. That is, every job has its width as the length of its feasible interval. We denote the resulting job set by $\mathcal{J}'$. Since each job in $\mathcal{J}'$ are not shiftable, there is only one feasible schedule for $\mathcal{J}'$ and it is optimal. Thus, $cost(S(\mathcal{J}_T)) \leq cost(\mathcal{O}(\mathcal{J}'))$ for any feasible schedule S for $\mathcal{J}_T$.

For each job in $\mathcal{J}_T$, the length of its feasible interval is at most $2w - 1$. Hence, we can bound the load at any time t of $\mathcal{O}(\mathcal{J}')$ by the loads of constant number of timeslots in $S(\mathcal{J}_T)$. Assume that at timeslot t an extended job J' is executed. That is, $t \in [r(J'), d(J'))$ since J' is not shiftable. Consider the job J corresponding to J', the execution interval of J in any feasible schedule must contains either timeslot $t - (w - 1)$, t, or $t + (w - 1)$. Hence we can upper bound the load at any time t in $\mathcal{O}(\mathcal{J}')$: $\ell(\mathcal{O}(\mathcal{J}'), t) \leq \ell(\mathcal{O}(\mathcal{J}_T), t - (w - 1)) + \ell(\mathcal{O}(\mathcal{J}_T), t) + \ell(\mathcal{O}(\mathcal{J}_T), t + (w - 1))$. Therefore, $cost(S(\mathcal{J}_T)) \leq cost(\mathcal{O}(\mathcal{J}')) \leq 3^\alpha \cdot cost(\mathcal{O}(\mathcal{J}_T))$.

(ii) For $\mathcal{J}_L$, we apply AlignFI and get $\mathcal{J}'_L$. We then run $\mathcal{V}$ and get $\mathcal{V}(\mathcal{J}'_L)$, which can be viewed as a schedule for unit width jobs. We get $S(\mathcal{J}_L) = \mathcal{V}(\mathcal{J}'_L)$ by FreeSch. Hence, $cost(\mathcal{UV}(\mathcal{J}_L)) = \sum_t \ell(\mathcal{UV}(\mathcal{J}_L), t)^\alpha = \sum_t \ell(S(\mathcal{J}_L), t)^\alpha = \sum_t \ell(\mathcal{V}(\mathcal{J}'_L), t)^\alpha = cost(\mathcal{V}(\mathcal{J}'_L))$. According to Corollary 9, $cost(\mathcal{V}(\mathcal{J}'_L)) \leq 2^\alpha \cdot (8 \cdot (e + e^2)^\alpha + 1) \cdot cost(\mathcal{O}(\mathcal{J}'_L))$ by choosing $\mathcal{BKP}$ as reference algorithm. Since $\mathcal{J}_L$ is set of loose jobs with uniform width, $cost(\mathcal{O}(\mathcal{J}'_L)) \leq 3^\alpha \cdot cost(\mathcal{O}(\mathcal{J}_L)) \leq 3^\alpha \cdot cost(\mathcal{O}(\mathcal{J}))$ by Lemma 13 and Observation 15. Hence, $cost(\mathcal{UV}(\mathcal{J}_L)) \leq 2^\alpha \cdot (8 \cdot (e + e^2)^\alpha + 1) \cdot 3^\alpha \cdot cost(\mathcal{O}(\mathcal{J}))$. □

Theorem 17. $cost(\mathcal{UV}(\mathcal{J})) \leq 12^\alpha \cdot (8(e + e^2)^\alpha + 1) \cdot cost(\mathcal{O}(\mathcal{J}))$.

Proof. By definition, $cost(\mathcal{UV}(\mathcal{J})) = \sum_t \ell(\mathcal{UV}(\mathcal{J}), t)^\alpha = \sum_t (\ell(\mathcal{UV}(\mathcal{J}_T), t) + \ell(\mathcal{UV}(\mathcal{J}_L), t))^\alpha \leq 2^{\alpha-1} \cdot \sum_t (\ell(\mathcal{UV}(\mathcal{J}_T), t)^\alpha + \ell(\mathcal{UV}(\mathcal{J}_L), t)^\alpha) = 2^{\alpha-1} \cdot (cost(\mathcal{UV}(\mathcal{J}_T)) + cost(\mathcal{UV}(\mathcal{J}_L)))$. By Lemma 16, $cost(\mathcal{UV}(\mathcal{J})) \leq 2^{\alpha-1} \cdot (3^\alpha + 6^\alpha \cdot (8(e + e^2)^\alpha + 1)) \cdot cost(\mathcal{O}(\mathcal{J})) \leq 2^\alpha \cdot 6^\alpha \cdot (8(e + e^2)^\alpha + 1) \cdot cost(\mathcal{O}(\mathcal{J}))$. □

4   Online algorithm for general case

In this section, we present an algorithm $\mathcal{G}$ for jobs with arbitrary width and height. We first transform job set $\mathcal{J}$ to a "nice" job set $\mathcal{J}^*$ (to be defined) and show that such a transformation only increases the cost modestly. Furthermore, we show that for any nice job set $\mathcal{J}^*$, we can bound $\mathrm{cost}(\mathcal{G}(\mathcal{J}^*))$ by $\mathrm{cost}(\mathcal{O}(\mathcal{J}^*))$ and in turn by $\mathrm{cost}(\mathcal{O}(\mathcal{J}))$. Then we can establish the competitive ratio of $\mathcal{G}$.

4.1   Nice job set and transformations

A job J is said to be a *nice job* if $w(J) = 2^p$, for some non-negative integer p and a job set $\mathcal{J}^*$ is said to be a *nice job set* if all its jobs are nice jobs. In other words, the nice job J is in class C_p.

Procedure CONVERT. Given a job set $\mathcal{J}$, we define the procedure CONVERT to transform each job $J \in \mathcal{J}$ into a nice job J^* as follows. We denote the resulting nice job set by $\mathcal{J}^*$. Suppose J is in class C_p. We modify its width, release time and deadline.

- $w(J^*) \leftarrow 2^p$;

- $r(J^*) \leftarrow r(J)$;

- $d(J^*) \leftarrow r(J^*) + \max\{d(J) - r(J), 2^p\}$.

Modifications to $r(J^*)$ and $d(J^*)$ are due to rounding up the width. The observation below follows directly from the definition.

Observation 18. *For any job J and its nice job J^* transformed by* CONVERT, *(i)* $I(J) \subseteq I(J^*)$; *(ii)* $I(J) \neq I(J^*)$ *if and only if* $|I(J)| < 2^p$; *in this case,* $den(J) > \frac{1}{2}$ *and* $den(J^*) = 1$.

We then define two procedures that transform schedules related to nice job sets. RELAXSCH takes a schedule S for a job set $\mathcal{J}$ and converts it to a schedule S^* for the corresponding nice job set $\mathcal{J}^*$. On the other hand, SHRINKSCH takes a schedule S^* for a nice job set $\mathcal{J}^*$ and converts it to a schedule S for $\mathcal{J}$.

Transformation RELAXSCH. RELAXSCH transforms S into S^* by moving the start and end time of every job J.

- $st(S^*, J^*) = \min\{d(J^*) - w(J^*), st(S, J)\}$

- $et(S^*, J^*) = st(S^*, J^*) + w(J^*)$.

Observation 19 asserts that the resulting schedule S^* is feasible for $\mathcal{J}^*$ while Lemmas 20 and 21 analyze the load and cost of the schedule.

Observation 19. *Consider any schedule S for $\mathcal{J}$ and the schedule S^* constructed by* RELAXSCH *for the corresponding $\mathcal{J}^*$. We have* $[st(S^*, J^*), et(S^*, J^*)] \subseteq [r(J^*), d(J^*)]$; *in other words, S^* is a feasible schedule for $\mathcal{J}^*$.*

To analyze the load of the schedule S^*, we consider partial schedule $S_p^* \subseteq S^*$ (resp. $S_p \subseteq S$) which is for all the jobs of $\mathcal{J}^*$ (resp. $\mathcal{J}$) in class C_p. Intuitively, the load of S_p^* at any time is at most the sum of the load of S_p at the current time and $2^{p-1} - 1$ timeslots before and after the current time.

Lemma 20. *At any time t,* $\ell(S_p^*, t) \leq \ell(S_p, t) + \ell(S_p, t - (2^{p-1} - 1)) + \ell(S_p, t + (2^{p-1} - 1))$.

Proof. We prove that for any job J, J^* contributes to $\ell(S_p^*, t)$ only if J contributes to either $\ell(S_p, t)$, $\ell(S_p, t - (2^{p-1} - 1))$, or $\ell(S_p, t + (2^{p-1} - 1))$, . There are two cases that J does not contribute to $\ell(S_{p,q}, t - (2^{p-1} - 1))$ nor $\ell(S_{p,q}, t + (2^{p-1} - 1))$: (i) $et(J) < t - (2^{p-1} - 1)$ or $st(J) > t + (2^{p-1} - 1)$, and (ii) $[st(J), et(J)] \subseteq (t - (2^{p-1} - 1), t + (2^{p-1} - 1))$.

Consider case (i). $et(J^*) \leq et(J) + (2^{p-1} - 1)$ and $st(J^*) \geq st(J) - (2^{p-1} - 1)$. Hence, $t \notin [st(J^*), et(J^*)]$ if $et(J) < t - (2^{p-1} - 1)$ or $st(J) > t + (2^{p-1} - 1)$. That is, J^* does not contribute to $\ell(S_p^*, t)$. Notice that if $et(J) = t - (2^{p-1} - 1)$ or $st(J) = t + (2^{p-1} - 1)$, J does not necessarily contribute to $\ell(S_p^*, t)$. We count the contribution for worst case analysis.

For case (ii), consider job J with $[st(J), et(J)] \subseteq (t - (2^{p-1} - 1), t + (2^{p-1} - 1))$. Since $2^{p-1} < w(J) \leq 2^p$, $t \in [st(J), et(J)]$. That is, J contributes to $\ell(S_p, t)$ no matter if J^* contributes to $\ell(S_p^*, t - (2^{p-1} - 1))$ or $\ell(S_p^*, t + (2^{p-1} - 1))$.

By case (i) and (ii), for any job J with $[st(J), et(J)] \cap [t - (2^{p-1} - 1), t + (2^{p-1} - 1)] = \emptyset$, J^* does not contribute to $\ell(S_p^*, t)$. And for any job J with $[st(J), et(J)] \subseteq (t - (2^{p-1} - 1), t + (2^{p-1} - 1))$, J contributes to $\ell(S_p, t)$. Hence, by assuming all jobs at timeslot $t - (2^{p-1} - 1)$ or $t + (2^{p-1} - 1)$ contribute to $\ell(S_p^*, t)$, $\ell(S_p^*, t)$ is bounded by $\ell(S_p, t) + \ell(S_p, t - (2^{p-1} - 1)) + \ell(S_p, t + (2^{p-1} - 1))$. □

Lemma 21. *Using* RELAXSCH, *we have* $cost(S_p^*) \leq 3^\alpha \cdot cost(S_p)$.

Proof. By Lemma 20, $cost(S_p^*) = \sum_t \ell(S_p^*, t)^\alpha \leq \sum_t (\ell(S_p, t) + \ell(S_p, t - (2^{p-1} - 1)) + \ell(S_p, t + (2^{p-1} - 1)))^\alpha \leq \sum_t (3 \cdot \ell(S_p, t))^\alpha = 3^\alpha \cdot cost(S_p)$. □

Transformation SHRINKSCH. On the other hand, SHRINKSCH converts a schedule S^* for a nice job set $\mathcal{J}^*$ to a schedule S for the corresponding job set $\mathcal{J}$. We set

- $st(S, J) \leftarrow st(S^*, J^*)$;

- $et(S, J) \leftarrow st(S, J) + w(J)$, therefore, $et(S, J) \leq et(S^*, J^*)$.

Observation 22 asserts that the resulting schedule S is feasible for J and Lemma 23 analyzes the cost of the schedule.

Observation 22. *Consider any schedule S^* for $\mathcal{J}^*$ and schedule S constructed by* SHRINKSCH *for the corresponding $\mathcal{J}$. For any J^* and the corresponding J, we have* (i) $[st(S, J), et(S, J)] \subseteq [st(S^*, J^*), et(S^*, J^*)]$; (ii) $[st(S, J), et(S, J)] \subseteq [r(J), d(J)]$.

By Observation 22, we have the following lemma.

Lemma 23. *Using* SHRINKSCH, *we have* $cost(S_p) \leq cost(S_p^*)$.

4.2 The online algorithm

Online algorithm $\mathcal{G}$. We are now ready to describe the algorithm $\mathcal{G}$ for an arbitrary job set $\mathcal{J}$. When a job J is released, it is converted to J^* by CONVERT and classified into one of the classes C_p. Jobs in the same class after CONVERT (being a uniform-width job set) are scheduled by $\mathcal{UV}$ independently of other classes. We then modify the execution time of J^* in $\mathcal{UV}$ to the execution time of J in $\mathcal{G}$ by Transformation SHRINKSCH. Note that all these procedures can be done in an online fashion.

Using the results in Sections 3 and 4.1, we can compare the cost of $\mathcal{G}(\mathcal{J})$ with $\mathcal{O}(\mathcal{J}_p^*)$ for each class C_p (see Theorem 25). It remains to analyze the cost of $\mathcal{O}(\mathcal{J}_p^*)$ and $\mathcal{O}(J)$ in the next observation.

Observation 24. *Consider any job set $\mathcal{J}$, its corresponding job set $\mathcal{J}^*$ and the corresponding job set of each class $\mathcal{J}_p$ and $\mathcal{J}_p^*$.* (i) $cost(\mathcal{O}(\mathcal{J}_p^*)) \leq 3^\alpha \cdot cost(\mathcal{O}(\mathcal{J}_p))$; (ii) $cost(\mathcal{O}(\mathcal{J}_p)) \leq cost(\mathcal{O}(\mathcal{J}))$.

Proof. (i) Given $\mathcal{O}(\mathcal{J}_p)$, there exists schedule $S(\mathcal{J}_p^*)$ generated by RELAXSCH. By Lemma 21 , $cost(S(\mathcal{J}_p^*)) \leq 3^\alpha \cdot cost(\mathcal{O}(\mathcal{J}_p))$. Hence, $cost(\mathcal{O}(\mathcal{J}_p^*)) \leq cost(S(\mathcal{J}_p^*)) \leq 3^\alpha \cdot cost(\mathcal{O}(\mathcal{J}_p))$.

(ii) Assume on the contrary that $cost(\mathcal{O}(\mathcal{J})) < cost(\mathcal{O}(\mathcal{J}_p))$, we can generate a schedule $S(\mathcal{J}_p)$ by removing jobs from $\mathcal{O}(\mathcal{J})$ which are not in $\mathcal{J}_p$. It follows that $cost(S(\mathcal{J}_p)) \leq cost(\mathcal{O}(\mathcal{J})) < cost(\mathcal{O}(\mathcal{J}_p))$, contradicting to the fact that $\mathcal{O}(\mathcal{J}_p)$ is optimal for $\mathcal{J}_p$. □

Theorem 25. *For any job set $\mathcal{J}$, we have $\text{cost}(\mathcal{G}(\mathcal{J})) \leq (36\lceil \log K \rceil)^\alpha \cdot (8(e + e^2)^\alpha + 1) \cdot \text{cost}(\mathcal{O}(\mathcal{J}))$, where $K = \frac{w_{\max}}{w_{\min}}$.*

Proof. By definition, $\text{cost}(\mathcal{G}(\mathcal{J})) = \sum_t \ell(\mathcal{G}(\mathcal{J}), t)^\alpha = \sum_t (\sum_{p=1}^{\lceil \log K \rceil} \ell(\mathcal{G}(\mathcal{J}_p), t))^\alpha$. The latter is at most $\lceil \log K \rceil^{\alpha-1} \sum_{p=1}^{\lceil \log K \rceil} \sum_t \ell(\mathcal{G}(\mathcal{J}_p), t)^\alpha$. For each group of jobs $\mathcal{J}_p$, we CONVERT it to $\mathcal{J}_p^*$, apply algorithm $\mathcal{UV}$ on it, and transform the schedule into a schedule for $\mathcal{J}_p$ by SHRINKSCH. Hence, $\ell(\mathcal{G}(\mathcal{J}_p), t) \leq \ell(\mathcal{UV}(\mathcal{J}_p^*), t)$ for each t. It follows that $\text{cost}(\mathcal{G}(\mathcal{J})) \leq \lceil \log K \rceil^{\alpha-1} \sum_{p=1}^{\lceil \log K \rceil} \text{cost}(\mathcal{G}(\mathcal{J}_p)) \leq \lceil \log K \rceil^{\alpha-1} \sum_{p=1}^{\lceil \log K \rceil} \text{cost}(\mathcal{UV}(\mathcal{J}_p^*))$. By Lemma 17 and Observations 24 (i) and 15, $\text{cost}(\mathcal{UV}(\mathcal{J}_p^*)) \leq 12^\alpha \cdot (8(e + e^2)^\alpha + 1) \cdot \text{cost}(\mathcal{O}(\mathcal{J}_p^*)) \leq 12^\alpha \cdot (8(e + e^2)^\alpha + 1) \cdot 3^\alpha \cdot \text{cost}(\mathcal{O}(\mathcal{J}_p)) \leq 36^\alpha \cdot (8(e + e^2)^\alpha + 1) \cdot \text{cost}(\mathcal{O}(\mathcal{J}))$. Hence $\text{cost}(\mathcal{G}(\mathcal{J})) \leq 36^\alpha \cdot \lceil \log K \rceil^{\alpha-1} \cdot (8(e + e^2)^\alpha + 1) \cdot \sum_{p=1}^{\lceil \log K \rceil} \text{cost}(\mathcal{O}(\mathcal{J})) = (36\lceil \log K \rceil)^\alpha \cdot (8(e + e^2)^\alpha + 1) \cdot \text{cost}(\mathcal{O}(\mathcal{J}))$. $\square$

Note that the logarithm in the competitive ratio comes from the number of classes defined in Section 2. Suppose we change the definition of classes such that class p includes jobs of size in the range $((1 + \lambda)^{p-1}, (1 + \lambda)^p]$ for some $\lambda > 0$ and Procedure CONVERT such that the width of jobs in class C_p is round up to $(1 + \lambda)^p$. Then, the number of classes becomes $\lceil \log_{1+\lambda} K \rceil$. In addition, the competitive ratio depends on Lemma 20 that bounds the load at any timeslot by the load of three other timeslots. This number of timeslots is also affected by the definition of classes. In summary, the following lemma states the competitive ratio for varying λ.

Lemma 26. *For $0 < \lambda \leq 0.5$, $0.5 < \lambda \leq 1$ and $\lambda > 1$, the competitive ratio of our algorithm becomes $(12 \times 2\lceil \log_{1+\lambda} K \rceil)^\alpha (8(e + e^2)^\alpha + 1)$, $(12 \times 3\lceil \log_{1+\lambda} K \rceil)^\alpha (8(e + e^2)^\alpha + 1)$, and $(12 \times (2\lambda + 1)\lceil \log_{1+\lambda} K \rceil)^\alpha (8(e + e^2)^\alpha + 1)$, respectively.*

Proof. The number of classes is $\lceil \log_{1+\lambda} K \rceil$, which replaces $\lceil \log K \rceil$ in Theorem 25. We note that this number decreases as λ increases. In Lemma 20, the load of S_p^* at any time t is bounded by the load of S_p at three timeslots when $\lambda = 1$. We observe that this property stays the same for $0.5 < \lambda \leq 1$. Using a similar argument, we can show that if λ is smaller and $0 < \lambda \leq 0.5$, then the number of timeslots involved becomes smaller and equals to 2. Furthermore, when $\lambda > 1$, the number of timeslots increases and equals to $2\lambda + 1$. $\square$

We note the competitive ratio for $\lambda < 1$ is larger than that for $\lambda = 1$, and the best competitive ratio occurs when $1 < \lambda < 2$.

4.3 Lower bound

In this section, we show lower bounds on competitive ratio for Grid problem with unit height and arbitrary width by designing an adversary for the problem. The lower bounds are immediately lower bounds for the general case of Grid problem.

The adversary constructs a set of jobs with a low cost of offline optimal schedule but a high cost of any online algorithm $\mathcal{A}$. It generates jobs one by one and assigns release times, deadlines and widths of jobs based on the previously generated jobs. The start times of jobs scheduled by $\mathcal{A}$ will be used for the job generations later. This ensures that $\mathcal{A}$ has to put a job on top of all existing jobs and results in a high energy cost. Meanwhile, the adversary will choose an appropriate feasible interval for each job such that an optimal offline algorithm can schedule the job set with low energy cost. The following is the description of the adversary.

Adversary Λ and job instance $\mathcal{J}$. Given an online algorithm $\mathcal{A}$, a constant $\alpha > 1$ and a large number x, adversary Λ outputs a set of jobs $\mathcal{J}$ consisting of $\lfloor \alpha \rfloor + 1$ jobs. Let J_i be the ith job of $\mathcal{J}$. The adversary first computes a width for each job before running $\mathcal{A}$. It sets $w(J_{\lfloor \alpha \rfloor}) = x$, $w(J_{\lfloor \alpha \rfloor+1}) = x - 1$, and $w(J_i) = 3w(J_{i+1}) + 1$ for $1 \leq i \leq \lfloor \alpha \rfloor - 1$. Then Λ releases the jobs from J_1 to $J_{\lfloor \alpha \rfloor+1}$ accordingly and computes a release time and a deadline for each job through an interaction with $\mathcal{A}$. For the first job J_1, Λ chooses any release time and deadline

such that $d(J_1) - r(J_1) \geq 3w(J_1)$. For the ith job $J_i \in \mathcal{J}$ for $2 \leq i \leq \lfloor \alpha \rfloor + 1$ accordingly, Λ sets $r(J_i) = st(\mathcal{A}, J_{i-1}) + 1$ and $d(J_i) = et(\mathcal{A}, J_{i-1})$. This limits $\mathcal{A}$ to fewer choices of start times for scheduling a new job. A job can only be scheduled in the execution interval of its previous job by $\mathcal{A}$. On the other hand, no two jobs have the same release time.

Let $w_{\max}$ and $w_{\min}$ denote by the maximum and minimum width of jobs respectively, and let $\mathcal{O}$ be an optimal offline algorithm for Grid problem. We have the following results.

Lemma 27. $cost(\mathcal{O}(\mathcal{J})) \leq x \cdot 3^{\lfloor \alpha \rfloor}$.

Proof. By the setting of Λ, we show that $\mathcal{O}$ can schedule all the jobs $\mathcal{J}$ without overlapping, and the cost of an optimal schedule is just the sum of widths of all the jobs.

For any job $J_i \in \mathcal{J}$ and $i \geq 2$, the length of its feasible interval is $d(J_i) - r(J_i) = et(\mathcal{A}, J_{i-1}) - (st(\mathcal{A}, J_{i-1}) + 1) = w(J_{i-1}) - 1 = 3w(J_i)$. This means no matter where we schedule a job, at least one of the intervals $[r(J_i), st(J_i))$ and $[et(J_i), d(J_i))$ has length at least $w(J_i)$. Algorithm $\mathcal{O}$ can schedule the subsequent jobs in the interval with length at least $w(J_i)$ such that the subsequent jobs do not overlap with J_i. This is because the sum of widths of all the subsequent jobs does not exceed $w(J_i)$. Since this argument can be applied on all the jobs, this implies that all the jobs do not overlap with each other in an optimal schedule. Thus the cost of an optimal schedule is the sum of widths of all the jobs. More precisely,

$$\begin{aligned}
cost(\mathcal{O}(\mathcal{J})) &= (x-1) + x + (3x+1) + (3(3x+1)+1) + \ldots + w_{\max} \\
&\leq 2x + 2 \cdot 3x + 2 \cdot 9x + \ldots + 2 \cdot 3^{\lfloor \alpha \rfloor - 1} x \\
&= 2x \cdot \frac{3^{\lfloor \alpha \rfloor} - 1}{2} \leq x \cdot 3^{\lfloor \alpha \rfloor} \ .
\end{aligned}$$ $\qquad \square$

Theorem 28. *For any deterministic online algorithm $\mathcal{A}$ for Grid problem with unit height and arbitrary width, adversary Λ constructs an instance $\mathcal{J}$ such that* (i) *for constant α,*

$$\frac{cost(\mathcal{A}(\mathcal{J}))}{cost(\mathcal{O}(\mathcal{J}))} \geq \left(\frac{\lfloor \alpha \rfloor + 1}{3} \right)^\alpha \ ;$$

and (ii) *for arbitrary α,*

$$\frac{cost(\mathcal{A}(\mathcal{J}))}{cost(\mathcal{O}(\mathcal{J}))} \geq \left(\frac{1}{3} \log \frac{w_{\max}}{w_{\min}} \right)^\alpha \ .$$

Proof. We first give a lower bound on $cost(\mathcal{A}(\mathcal{J}))$ and then give the lower bounds on the competitive ratio by combining $cost(\mathcal{A}(\mathcal{J}))$ with Lemma 27.

(i) By the setting of Λ, all the jobs scheduled by $\mathcal{A}$ overlap with each other. For ease of the computation for the cost of $\mathcal{A}$, we only consider the timeslots contained by the execution interval of the last job $J_{\lfloor \alpha \rfloor + 1}$. Thus $cost(\mathcal{A}(\mathcal{J})) \geq (x-1) \cdot (\lfloor \alpha \rfloor + 1)^\alpha$ and

$$\frac{cost(\mathcal{A}(\mathcal{J}))}{cost(\mathcal{O}(\mathcal{J}))} \geq \frac{(x-1) \cdot (\lfloor \alpha \rfloor + 1)^\alpha}{x \cdot 3^{\lfloor \alpha \rfloor}} \geq \left(\frac{\lfloor \alpha \rfloor + 1}{3} \right)^\alpha$$

as x to be large enough.

(ii) Assume α can be arbitrarily large. We use $w_{\max}$ and $w_{\min}$ to bound $\lfloor \alpha \rfloor + 1$. According to Lemma 27, we have $w_{\max} \leq cost(\mathcal{O}(\mathcal{J})) \leq x \cdot 3^{\lfloor \alpha \rfloor}$, and thus

$$\lfloor \alpha \rfloor \geq \log_3 \frac{w_{\max}}{x} \geq \log_3 \frac{w_{\max}}{3(x-1)} = \log_3 \frac{w_{\max}}{w_{\min}} - 1 \ .$$

Note that $x \leq 3(x-1)$ if $x \geq 2$. Therefore, $cost(\mathcal{A}(\mathcal{J})) \geq (x-1) \cdot \log^\alpha \frac{w_{\max}}{w_{\min}}$. Combining with Lemma 27, we have the lower bound on the competitive ratio

$$\frac{cost(\mathcal{A}(\mathcal{J}))}{cost(\mathcal{O}(\mathcal{J}))} \geq \frac{(x-1) \cdot \log^\alpha \frac{w_{\max}}{w_{\min}}}{x \cdot 3^{\lfloor \alpha \rfloor}} \geq \left(\frac{1}{3} \log \frac{w_{\max}}{w_{\min}} \right)^\alpha$$

as x to be large enough. $\qquad \square$

Corollary 29. *For any deterministic online algorithm for Grid problem, the competitive ratio is at least* (i) $(\frac{\lfloor \alpha \rfloor + 1}{3})^\alpha$ *for constant α; and* (ii) $(\frac{1}{3} \log \frac{w_{\max}}{w_{\min}})^\alpha$ *for arbitrary α.*

5 Online algorithm for uniform height jobs

In this section we focus on uniform-height jobs of height h and consider two special cases of the width. We first consider jobs with uniform-height and unit-width (Section 5.2) and secondly consider jobs with agreeable deadlines (Section 5.3).

To ease the discussion, we refine a notation we defined before. For any algorithm $\mathcal{A}$ for a job set $\mathcal{J}$ and a time interval $\mathcal{I}$, we denote by $\mathcal{A}(\mathcal{J}, \mathcal{I})$ the schedule of $\mathcal{A}$ on $\mathcal{J}$ over the time interval $\mathcal{I}$.

5.1 Main ideas

The main idea is to make reference to the online algorithm $\mathcal{AVR}$ and consider two types of intervals, $\mathcal{I}_{>h}$ where the average load is higher than h and $\mathcal{I}_{\leq h}$ where the average load is at most h. For the former, we show that we can base on the competitive ratio of $\mathcal{AVR}$ directly; for the latter, our load could be much higher than that of $\mathcal{AVR}$ and in such case, we compare directly to the optimal algorithm. Combining the two cases, we have Lemma 30, which holds for any job set. In Sections 5.2 and 5.3, we show how we can use this lemma to obtain algorithms for the special cases. Notice that the number $\lceil \frac{\text{avg}(t)}{h} \rceil$ is the minimum number of jobs needed to make the load at t at least $\text{avg}(t)$.

Lemma 30. *Suppose we have an algorithm $\mathcal{A}$ for a any job set $\mathcal{J}$ such that for some c and c'* (i) $\ell(\mathcal{A}, t) \leq c \cdot h \cdot \lceil \frac{\text{avg}(t)}{h} \rceil$ *for all $t \in \mathcal{I}_{>h}$, and* (ii) $\ell(\mathcal{A}, t) \leq c' \cdot h$ *for all $t \in \mathcal{I}_{\leq h}$. Then we have* $\text{cost}(\mathcal{A}(\mathcal{J})) \leq (\frac{(4c\alpha)^\alpha}{2} + c'^\alpha) \cdot \text{cost}(\mathcal{O}(\mathcal{J}))$.

Proof. We denote the speed of $\mathcal{AVR}$ at t as $\ell(\mathcal{AVR}, t)$. We are going to prove that (a) $\text{cost}(\mathcal{A}(\mathcal{J}, \mathcal{I}_{>h})) \leq \frac{(4c\alpha)^\alpha}{2} \cdot \text{cost}(\mathcal{O}(\mathcal{J}))$ and (b) $\text{cost}(\mathcal{A}(\mathcal{J}, \mathcal{I}_{\leq h})) \leq c'^\alpha \cdot \text{cost}(\mathcal{O}(\mathcal{J}))$. Hence, the total cost $\text{cost}(\mathcal{A}(\mathcal{J})) \leq (\frac{(4c\alpha)^\alpha}{2} + c'^\alpha) \cdot \text{cost}(\mathcal{O}(\mathcal{J}))$ since $\mathcal{I}_{>h}$ and $\mathcal{I}_{\leq h}$ are disjoint.

(a) We compare $\ell(\mathcal{A}, t)$ to $\ell(\mathcal{AVR}, t)$ for each timeslot t in $\mathcal{I}_{>h}$. The assumption of $\mathcal{A}$ means that $\ell(\mathcal{A}, t) \leq c \cdot h \cdot \lceil \frac{\text{avg}(t)}{h} \rceil < c \cdot h \cdot (\frac{\text{avg}(t)}{h} + 1) = c \cdot (\text{avg}(t) + h) \leq 2c \cdot \text{avg}(t)$ since $\text{avg}(t) > h$. By definition, $\text{cost}(\mathcal{A}(\mathcal{J}, \mathcal{I}_{>h})) = \sum_{t \in \mathcal{I}_{>h}} \ell(\mathcal{A}, t)^\alpha \leq \sum_{t \in \mathcal{I}_{>h}} (2c \cdot \text{avg}(t))^\alpha$. Recall that $\ell(\mathcal{AVR}, t) = \text{avg}(t)$ for each t. Hence, by Corollary 2, $\text{cost}(\mathcal{A}(\mathcal{J}, \mathcal{I}_{>h})) \leq (2c)^\alpha \cdot \text{cost}(\mathcal{AVR}(\mathcal{J}, \mathcal{I}_{>h})) \leq (2c)^\alpha \cdot \text{cost}(\mathcal{AVR}(\mathcal{J})) \leq \frac{(4c\alpha)^\alpha}{2} \cdot \text{cost}(\mathcal{O}(\mathcal{J}))$.

(b) Since only jobs which are available in $\mathcal{I}_{\leq h}$ can be scheduled at $t \in \mathcal{I}_{\leq h}$, $\text{cost}(\mathcal{A}(\mathcal{J}, \mathcal{I}_{\leq h})) \leq \sum_{J: I \cap \mathcal{I}_{\leq h} \neq \emptyset} w(J) \cdot (c'h)^\alpha \leq \sum_{J \in \mathcal{J}} w(J) \cdot (c'h)^\alpha$. By convexity, $\text{cost}(\mathcal{O}(\mathcal{J})) \geq \sum_{J \in \mathcal{J}} w(J) \cdot h(J)^\alpha = \sum_{J \in \mathcal{J}} w(J) \cdot h^\alpha$. Hence, $\text{cost}(\mathcal{A}(\mathcal{J}, \mathcal{I}_{\leq h})) \leq c'^\alpha \cdot \text{cost}(\mathcal{O}(\mathcal{J}))$.

Adding up the two cost, we have $\text{cost}(\mathcal{A}(\mathcal{J})) = \text{cost}(\mathcal{A}(\mathcal{J}, \mathcal{I}_{>h})) + \text{cost}(\mathcal{A}(\mathcal{J}, \mathcal{I}_{\leq h})) \leq (\frac{(4c\alpha)^\alpha}{2} + c'^\alpha) \cdot \text{cost}(\mathcal{O}(\mathcal{J}))$ and the theorem follows. $\qquad \square$

5.2 Uniform-height and unit-width

In this section we consider job sets where all jobs have uniform-height and unit-width, i.e., $w(J) = 1$ and $h(J) = h$ for all jobs J. Note that such case is a subcase discussed in Section 3.1. Here we illustrate a different approach using the ideas above and describe the algorithm $\mathcal{UU}$ for this case. The competitive ratio of $\mathcal{UU}$ is better than that of Algorithm $\mathcal{V}$ in Section 3.1 when $\alpha < 3.22$.

Algorithm $\mathcal{UU}$. At any time t, choose $\lceil \frac{\text{avg}(t)}{h} \rceil$ jobs according to the EDF rule and schedule them to start at t. If there are fewer jobs available, schedule all available jobs.

The next theorem asserts that the algorithm gives feasible schedule and states its competitive ratio.

15

Theorem 31. (i) *The schedule constructed by Algorithm $\mathcal{UU}$ is feasible.* (ii) *Algorithm $\mathcal{UU}$ is $(\frac{(4\alpha)^\alpha}{2} + 1)$-competitive.*

Proof. (i) The feasibility can be proved by comparing to $\mathcal{AVR}$. At any time t, the total work done by $\mathcal{AVR}$ in interval $[0, t)$ is $\sum_{t' < t} \text{avg}(t')$. On the other hand, the total work done by $\mathcal{UU}$ in the same interval is $\sum_{t' < t} h \cdot \lceil \frac{\text{avg}(t)}{h} \rceil \geq \sum_{t' < t} h \cdot \frac{\text{avg}(t)}{h} = \sum_{t' < t} \text{avg}(t)$ if there are enough available jobs in this interval. If the number of available jobs is less than $\sum_{t' < t} h \cdot \lceil \frac{\text{avg}(t)}{h} \rceil$, $\mathcal{UU}$ will execute all these jobs. The work done by $\mathcal{UU}$ within interval $[0, t)$ is at least the work done by $\mathcal{AVR}$ in both cases. Hence, $\mathcal{UU}$ is feasible since $\mathcal{AVR}$ is feasible.

(ii) We note that $\ell(\mathcal{UU}, t) \leq h \cdot \lceil \frac{\text{avg}(t)}{h} \rceil$. To use Lemma 30, we can set $c' = 1$ for $t \in \mathcal{I}_{\leq h}$ by the definition of $\mathcal{I}_{\leq h}$. Furthermore, we can set $c = 1$ for $t \in \mathcal{I}_{>h}$. $\qquad\square$

5.3   Uniform-height, arbitrary width and agreeable deadlines

In this section we consider jobs with agreeable deadlines. We first note that simply scheduling $\lceil \frac{\text{avg}(t)}{h} \rceil$ number of jobs may not return a feasible schedule.

Example 32. *Consider four jobs each job J with $r(J) = 0$, $d(J) = 5$, $h(J) = h$, $w(J) = 3$. Note that $avg(t) = 2.4 \cdot h$ for all t. If we schedule at most $\lceil \frac{avg(t)}{h} \rceil = 3$ jobs at any time, we can complete three jobs but the remaining job cannot be completed. To schedule all jobs feasibly, we need at least two timeslots where all jobs are being executed.*

To schedule these jobs, we first observe in Lemma 33 that for a set of jobs with total densities at most h, it is feasible to schedule them such that the load at any time is at most h. Roughly speaking, we consider jobs in the order of release, and hence, in EDF manner since the jobs have agreeable deadlines. We keep the current ending time of all jobs that have been considered. As a new job is released, if its release time is earlier than the current ending time, we set its start time to the current ending time (and increase the current ending time by the width of the new job); otherwise, we set its start time to be its release time. Lemma 33 asserts that such scheduling is feasible and maintains the load at any time to be at most h.

Using this observation, we then partition the jobs into "queues" each of which has sum of densities at most h. Each queue $\mathcal{Q}_i$ is scheduled independently and the resulting schedule is to "stack up" all these schedules. The queues are formed in a Next-Fit manner: (i) the current queue $\mathcal{Q}_q$ is kept "open" and a newly arrived job is added to the current queue if including it makes the total densities stays at most 1; (ii) otherwise, the current queue is "closed" and a new queue $\mathcal{Q}_{q+1}$ is created as open.

Lemma 33. *Given any set of jobs of uniform-height, arbitrary-width and agreeable deadlines. If the sum of densities of all these jobs is at most h, then it is feasible to schedule all of them using a maximum load h at any time. That is, there is no stacking up among these jobs.*

Proof. Suppose there are k jobs $J_1, J_2, \cdots, J_k$ such that $\sum_{1 \leq i \leq k} \text{den}(J_i) \leq h$. Without loss of generality, we assume that $d(J_i) \leq d(J_j)$ and $r(J_i) \leq r(J_j)$ for $1 \leq i < j \leq k$. We claim that it is feasible to set $[st(J_i), et(J_i))$ to $[\max\{r(J_i), et(J_{i-1})\}, st(J_i) + w(J_i))$ for all $1 < i \leq k$ and $[st(J_1), et(J_1)) = [r(J_1), r(J_1) + w(J_1))$.

We observe that $\text{den}(J_1) \leq h$ since $\sum_i \text{den}(J_i) \leq h$. It is feasible to set $[st(J_1), et(J_1))$ to $[r(J_1), r(J_1) + w(J_1))$ since the input is feasible. Then we have to prove that $[st(J_i), et(J_i)) = [\max\{r(J_i), et(J_{i-1})\}, st(J_i) + w(J_i)) \subseteq [r(J_i), d(J_i))$. Since $st(J_i) = \max\{r(J_i), et(J_{i-1})\}$, we have $st(J_i) \geq r(J_i)$. Assume that $\cup_{g \leq i} I(J_g)$ is a contiguous interval. Since $\sum_{g \leq i} \text{den}(J_g) \leq h$, $\sum_{g \leq i} \frac{w(J_g)}{d(J_g) - r(J_g)} \leq 1$. From the ordering of jobs we have $1 \geq \sum_{g \leq i} \frac{w(J_g)}{d(J_g) - r(J_g)} \geq \sum_{g \leq i} \frac{w(J_g)}{d(J_i) - r(J_1)}$. Hence, $\sum_{g \leq i} w(J_g) \leq d(J_i) - r(J_1)$. Therefore J_i can be finished before $d(J_i)$. On the other hand, if $\cup_{g \leq i} I(J_g)$ is not contiguous. The proof above shows that for each contiguous, each of the involving jobs can be finished by its deadline. $\qquad\square$

Width	Height	Time complexity
Arbitrary	Arbitrary	$w_{\max}{}^{2m} \cdot (W_{\max} + 1)^{4m} \cdot O(n^2)$
Arbitrary	Arbitrary	$(4m \cdot w_{\max}{}^2)^{2m} \cdot O(n^2)$
Unit	Arbitrary	$2^{O(N)}$

Table 2: Summary of our exact algorithms.

Algorithm $\mathcal{AD}$. The algorithm consists of the following components: InsertQueue, Set-StartTime and ScheduleQueue.

InsertQueue: We keep a counter q for the number of queues created. When a job J arrives, if $\operatorname{den}(J) + \sum_{J' \in \mathcal{Q}_q} \operatorname{den}(J') \leq h$, then job J is added to $\mathcal{Q}_q$; otherwise, job J is added to a new queue $\mathcal{Q}_{q+1}$ and we set $q \leftarrow q + 1$.

SetStartTime: For the current queue, we keep a current ending time E, initially set to 0. When a new job J is added to the queue, if $r(J) \leq E$, we set $st(J) \leftarrow E$; otherwise, we set $st(J) \leftarrow r(J)$. We then update E to $st(J) + w(J)$.

ScheduleQueue: At any time t, schedule all jobs in all queues with start time set at t.

By Lemma 33, the schedule returns by $\mathcal{AD}$ is feasible. We then analyze its load and hence, derive its competitive ratio. Recall the definition of $\mathcal{I}_{>h}$ and $\mathcal{I}_{\leq h}$

Lemma 34. *Using $\mathcal{AD}$, we have* (i) $\ell(\mathcal{AD}, t) \leq 2 \cdot h \cdot \lceil \frac{avg(t)}{h} \rceil$ *for* $t \in \mathcal{I}_{>h}$; (ii) $\ell(\mathcal{AD}, t) \leq 2h$ *for* $t \in \mathcal{I}_{\leq h}$.

Proof. For timeslot t, suppose there are k queues $(Q_1, Q_2, \cdots, Q_k)$ which contains jobs available at t. According to our algorithm, $\ell(\mathcal{AD}, t) \leq k \cdot h$.

Let D_i be the sum of densities of jobs in Q_i. Consider $t \in \mathcal{I}_{>h}$. By the InsertQueue procedure, $D_i + D_{i+1} > h$ for $1 \leq i < k - 1$. Therefore, if $k \geq 3$, $avg(t) = \sum_{1 \leq i \leq k} D_i > \sum_{1 \leq i \leq k-1} D_i \geq \lfloor \frac{k-1}{2} \rfloor \cdot h$. It can be shown that $k \leq 2 \cdot \lceil \frac{avg(t)}{h} \rceil$ since $avg(t) > 1$.[1] That is, $\ell(\mathcal{AD}, t) = k \cdot h \leq 2 \cdot h \cdot \lceil \frac{avg(t)}{h} \rceil$ for $t \in \mathcal{I}_{>h}$. On the other hand, if $k < 3$, $\ell(\mathcal{AD}, t) \leq 2h < 2 \cdot avg(t) \leq 2 \cdot h \cdot \lceil \frac{avg(t)}{h} \rceil$.

For $t \in \mathcal{I}_{\leq h}$, $avg(t) \leq h$ by definition. That is, the sum of densities of all available jobs at t is no more than h. By the InsertQueue procedure all jobs will be in at most two adjacent queues. Hence, $\ell(\mathcal{AD}, t) \leq 2h$ for $t \in \mathcal{I}_{\leq h}$. $\qquad \square$

By Lemma 34 and Lemma 30, we have Theorem 35 by setting $c = 2$ and $c' = 2$.

Theorem 35. *For jobs with uniform height, arbitrary width and agreeable deadlines, $\mathcal{AD}$ is $(\frac{(8\alpha)^\alpha}{2} + 2^\alpha)$-competitive.*

6 Exact Algorithms

In this section, we propose exact algorithms and derive lower bounds on the running time of exact algorithms. Table 2 gives a summary of our results.

6.1 Fixed parameter algorithms

In parameterized complexity theory, the complexity of a problem is not only measured in terms of the input size, but also in terms of parameters. The theory focuses on situations where the parameters can be assumed to be small, and the time complexity depends mainly on these small parameters. The problems having such small parameters are captured by the

---

[1]Let $r = \frac{avg(t)}{h}$, since $t \in \mathcal{I}_{>h}$, $r > 1$. Since there are k queues at t, the total density at t is at least $\lceil \frac{k-1}{2} \rceil$. Hence, $\lceil \frac{k-1}{2} \rceil < r \leq \lceil r \rceil$. Since $\lceil \frac{k-1}{2} \rceil$ is a natural number, $\lceil \frac{k-1}{2} \rceil \leq \lceil r \rceil - 1$. It follows that $\lceil r \rceil - 1 > \frac{k-1}{2} - 1$. Therefore, $k \leq 2\lceil r \rceil = 2\lceil \frac{avg(t)}{h} \rceil$ since k is a natural number.

concept "fixed-parameter tractability". An algorithm with parameters $p_1, p_2, \ldots$ is said to be an *fixed parameter algorithm* if it runs in $f(p_1, p_2, \ldots) \cdot O(g(N))$ time for any function f and any polynomial function g, where N is the size of input. A parameterized problem is *fixed-parameter tractable* if it can be solved by a fixed parameter algorithm. In this section, we show that the general case of Grid problem, jobs with arbitrary release times, deadlines, width and height, is fixed-parameter tractable with respect to a few small parameters.

6.1.1 Key notions

We design two fixed parameter algorithms that are based on a dynamic programming fashion. Roughly speaking, we divide the timeline into k contiguous windows in a specific way, where each window W_i represents a time interval $[b_i, b_{i+1})$ for $1 \leq i \leq k$. The algorithm visits all windows accordingly from the left to the right and maintains a candidate set of schedules for the visited windows that no optimal solution is deleted from the set. In the first fixed parameter algorithm, the parameters of the algorithm are the maximum width of jobs, the maximum number of overlapped feasible intervals and the maximum size of windows, where the latter two can be parameterized if we interpret the input job set as an "interval graph". We will drop out the last parameter in the second algorithm. All parameters do not increase necessarily as the number of jobs grows, and can be assumed to be small in practice. For example, a width of a job is a requested amount of time to run an appliance, and the running time is usually a few hours, which is small when we make a timeslot to be an hour. And the number of overlapped feasible intervals is at most the number of appliances.

Interval graph. A graph $G = (V, E)$ is an *interval graph* if it captures the intersection relation for some set of intervals on the real line. Formally, for each $v \in V$, we can associate v to an interval I_v such that (u, v) is in E if and only if $I_u \cap I_v \neq \emptyset$. It has been shown in [20, 21] that an interval graph has a "consecutive clique arrangement", i.e., its maximal cliques can be linearly ordered in a way that for every vertex v in the graph, the maximal cliques containing v occur consecutively in the linear order. For any instance of the Grid problem, we can transform it into an interval graph $G = (V, E)$: For each job J with interval $I(J)$, we create a vertex $v(J) \in V$ and an edge is added between $v(J)$ and $v(J')$ if and only if $I(J)$ intersects $I(J')$. We can then obtain a set of maximal cliques in linear order, $C_1, C_2, \cdots, C_k$, by sweeping a vertical line from the left to the right, where k denotes the number of maximal cliques thus obtained. The parameter of our algorithm, the maximum number of overlapped feasible intervals, is just the maximum size of these maximal cliques.

Boundaries and windows. Based on the maximal cliques described above, we define some "windows" $W_1, W_2, \cdots, W_k$ with "boundaries" $b_1, b_2, \cdots, b_{k+1}$ as follows. We first give the definition of boundaries for the first algorithm. This definition will be generalized in section 6.1.4 for the second algorithm. For $1 \leq i \leq k$, the i-th *boundary* b_i is defined as the earliest release time of jobs in clique C_i but not in cliques before C_i, precisely, $b_i = \min\{t \mid t = r(J) \text{ and } J \in C_i \setminus (\cup_{s=1}^{i-1} C_s)\}$. The rightmost boundary b_{k+1} is defined as the latest deadline among all jobs. With the boundaries, we partition the timeslots into contiguous intervals called *windows*. The i-th window W_i is defined as $[b_i, b_{i+1})$.

Example. Figure 1 is an example of a set of jobs, its corresponding interval graph and the corresponding maximal cliques. The cliques are put in such a way that any vertex appears consecutively if there is two or more of it. The boundaries of windows are determined by the leftmost vertex of the maximal cliques.

6.1.2 Framework of the algorithms

We propose two exact algorithms, both of which runs in k stages corresponding to each of the k windows. We maintain a table T_{left} that stores all "valid" configurations of jobs in all the windows that have been considered so far. A configuration of a job corresponds to an execution
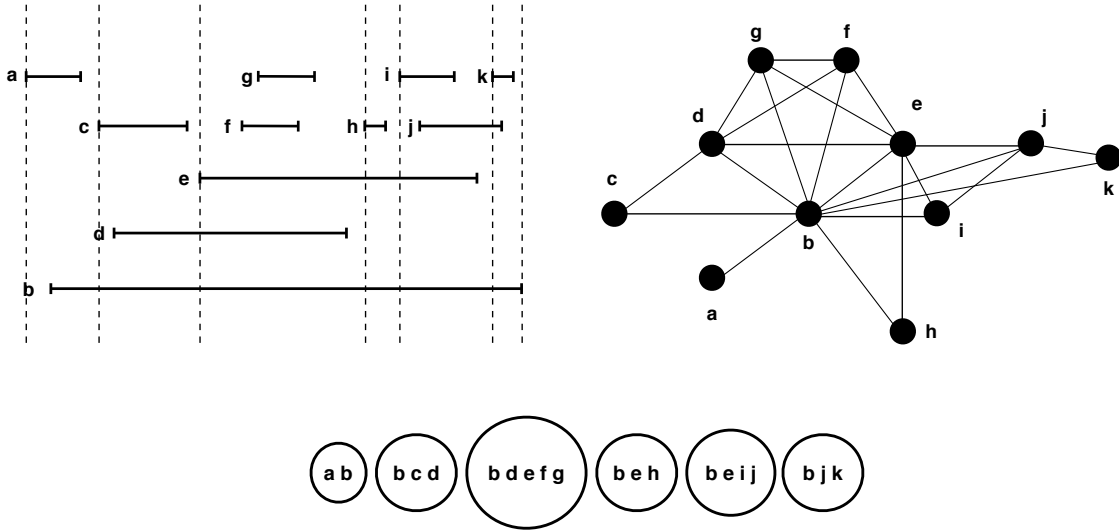
Figure 1: The figure on the left is a set of jobs, where the horizontal line segments are the feasible time intervals of jobs and the vertical lines are boundaries of windows. The figure on the right is an interval graph of the corresponding job set. And the figure at the bottom is a set of all the maximal cliques in the interval graph.

segment. And a row in the table consists of the configurations of all the jobs. In addition, for each window W_i, we compute a table T_{right_i} to store all possible configurations of start and end time of jobs available in W_i. The configurations in T_{right_i} would then be "concatenated" to some configurations in T_{left} that are "compatible" with each other. These merged configurations will be filtered to remove those non-optimal ones. The remaining configurations will become the new T_{left} for the next window. To describe the details of the algorithm, we explain several notions below. We denote by W_{left} the union of the windows corresponding to T_{left}. More formally, in the i-th stage, $W_{\text{left}} = \cup_{j<i} W_j$. And we use T_{right} to denote T_{right_i} when the context is clear.

Configurations. A *configuration* $F_i(J)$ of job J in window W_i is an execution segment, denoted by $[st_i(J), et_i(J))$ contained completely by W_i. That is, $st_i(J) \in \{b_i, b_i + 1, \cdots, b_{i+1} - 1\} \cup \{b_i - 1, b_{i+1}\}$ and $et_i(J) \in \{b_i + 1, b_i + 2, \cdots, b_{i+1}\} \cup \{b_i, b_{i+1} + 1\}$. Setting $st_i(J) = b_i - 1$ and $et_i(J) = b_i$ means J is executed completely before W_i. Similarly, setting $st_i(J) = b_{i+1}$ and $et_i(J) = b_{i+1} + 1$ means J starts execution after W_i. Also, setting $st_i(J) = b_i - 1$ and $et_i(J) = b_{i+1} + 1$ means J starts execution before W_i, crosses the whole window W_i, and ends execution after W_i. We say that $st_i(J) \in W_i$ or $et_i(J) \in W_i$ if $st_i(J) \in [b_i, b_{i+1})$ or $et_i(J) \in (b_i, b_{i+1}]$ respectively. And J is executed in W_i if both $st_i(J) \in W_i$ and $et_i(J) \in W_i$ hold. For a collection C of jobs, we use $F_i(C)$ to denote the set of configurations of all jobs in C, and $F_{\text{left}}(J)$ and $F_{\text{left}}(C)$ for the counterparts corresponding to T_{left}. The cost of $F_i(C)$ is the cost corresponding to the execution segments in $F_i(C)$. That is, $\text{cost}(F_i(C)) = \sum_{t \in W_i} (\sum_{J \in C: t \in F_i(J)} h(J))^\alpha$.

Validity. A configuration $F_i(J)$ is *invalid* if one of the following conditions hold: (i) $st_i(J) \geq et_i(J)$; (ii) $et_i(J) > st_i(J) + w(J)$ meaning that the length of execution segment of J is larger than the width of J; (iii) $(et_i(J) < st_i(J) + w(J)) \wedge (st_i(J) \geq b_i) \wedge (et_i(J) \leq b_{i+1})$ meaning that the length of execution segment of J is smaller than the width of J; (iv) $(st_i(J) < r(J)) \wedge (st_i(J) < b_{i+1})$ meaning that the start time of J is earlier than the release time of J; (v) $(et_i(J) > d(J)) \wedge (et_i(J) > b_i)$ meaning that the end time of J exceeds the deadline of J. Note that for $F_{\text{left}}(J)$, the validity is defined on the boundaries b_1 (instead of b_i) and b_{i+1}. And for T_{left}, $F_{\text{left}}(J)$ is also invalid if $st_{\text{left}}(J) = b_1 - 1$ since there is no window on the left of W_{left}. Similarly, $F_k(J)$ is invalid if $et_k(J) = b_{k+1} + 1$. A configuration $F_i(C)$ is *invalid* if there exists $J \in C$ such that $F_i(J)$ is invalid.

19

Compatibility. For job J, the two configurations $F_{\text{left}}(J)$ and $F_i(J)$ are *compatible* if (i) J is executed in W_{left} for $F_{\text{left}}(J)$, and J is executed before W_i for $F_i(J)$; (ii) J starts execution in W_{left} and ends execution after W_{left} for $F_{\text{left}}(J)$, and J starts execution before W_i and ends execution either in W_i or after W_i for $F_i(J)$; (iii) J is executed completely after W_{left} for $F_{\text{left}}(J)$, and J does not start before W_i for $F_i(J)$.

Concatenating configurations. To concatenate two configurations $F_{\text{left}}(J)$ and $F_i(J)$, we create a new $F_{\text{left}}(J)$ by the following setting based on the three types of compatible configurations described in the previous paragraph: for type (i), $st_{\text{left}}(J)$ and $et_{\text{left}}(J)$ leave unchanged; for type (ii), $st_{\text{left}}(J)$ leaves unchanged and set $et_{\text{left}}(J) \leftarrow et_i(J)$; and for type (iii), set $st_{\text{left}}(J) \leftarrow st_i(J)$ and $et_{\text{left}}(J) \leftarrow et_i(J)$. *Concatenating $F_{\text{left}}(C)$ and $F_i(C)$* is to concatenate the configurations of each job in C. The corresponding cost is simply adding the cost of the two configurations.

Uncertainty and identity. A configuration $F_i(J)$ is *uncertain* if $et_i(J) = b_{i+1} + 1$ meaning that the end time of J is not determined yet, and we are not sure at the i-th stage whether $F_i(J)$ will be valid after concatenating $F_i(J)$ and $F_{i+1}(J)$. Two configurations $F_i(C)$ and $F_i'(C)$ are *identical* if (i) $F_i(J)$ is uncertain if and only if $F_i'(J)$ is uncertain for all job $J \in C$; and (ii) the start time of $F_i(J)$ is equal to the start time of $F_i'(J)$ for all uncertain configuration $F_i(J)$ and $J \in C$. That is, we only consider the differences among the start times of those jobs with uncertain configurations when we distinguish two configurations of a set of jobs.

6.1.3   An algorithm with three parameters

Algorithm $\mathcal{E}$. The algorithm consists of three components: ListConfigurations, ConcatenateTables and FilterTable. In the algorithm, we first transform the input job set $\mathcal{J}$ to an interval graph, and obtain the maximal cliques C_i for $1 \leq i \leq k$ and the corresponding windows W_i. We start with T_{left} containing the only configuration, which sets $st_0(J) = b_1$ and $et_0(J) = b_1 + 1$ for all jobs J. That is, the configuration treats all the jobs to be not yet executed. Then we visit the windows from the left to the right.

ListConfigurations: For window W_i and jobs in C_i, we construct T_{right} storing all configurations of $J \in C_i$. We enumerate all $st_i(J) \in W_i$ and $et_i(J) \in W_i$ for each job $J \in C_i$, list all the combinations of all the jobs J with all of its start times and end times, and store the results in T_{right} in the way that one row is for one configuration $F_i(C_i)$. In another words, T_{right} stores all the combinations of execution segments in W_i for all jobs $J \in C_i$. Note that the jobs with release time later than W_i are considered to execute after W_i and the jobs with deadline earlier than W_i are considered to execute before W_i. For each configuration $F_i(C_i)$, we also store its cost contribution $cost(F_i(C_i))$ together. We also check each of the configurations and delete those invalid ones.

ConcatenateTables: We then concatenate compatible configurations in T_{left} and T_{right}. The resulting table is the new T_{left}. More specifically, for each configuration $F_{\text{left}}(C)$ in T_{left} and each configuration $F_{\text{right}}(C)$ in T_{right}, we concatenate $F_{\text{left}}(C)$ and $F_{\text{right}}(C)$ if they are compatible, and store the result to a new row in T_{left}. We also check each of the configurations in the new T_{left} and delete those invalid ones.

FilterTable: After concatenation, we filter non-optimal configurations. We classify all the configurations in T_{left} into groups such that the configurations in a group are identical and no two configurations from different groups are identical. For each group, we only leave the configuration with the lowest cost (choosing anyone to break tie if any) and remove the others in the group. In the current T_{left}, no two configurations are identical.

After processing all windows, the only configuration in the final T_{left} is returned as the solution. Algorithm 1 is the pseudocode of this algorithm.

Lemma 36. *Algorithm $\mathcal{E}$ outputs an optimal solution.*

Proof. In each stage, we list all possible configurations. A configuration is deleted only when it is invalid or it is identical to another configuration with lower cost. Hence, an invalid configuration

---

Algorithm 1 The fixed parameter algorithm $\mathcal{E}$

  Input: a set of job $\mathcal{J}$
  Output: an optimal configuration of $\mathcal{J}$
  $\{(W_i, C_i)\}_{i=1}^k \leftarrow$ the windows and their corresponding cliques of $\mathcal{J}$
  $T_{\text{left}} \leftarrow$ a configuration that sets all jobs $J \in \mathcal{J}$ to be not yet executed
  for i from 1 to k **do**
      $T_{\text{right}} \leftarrow \text{ListConfigurations}(W_i, C_i)$
      $T_{\text{left}} \leftarrow \text{ConcatenateTables}(T_{\text{left}}, T_{\text{right}})$
      $T_{\text{left}} \leftarrow \text{FilterTable}(T_{\text{left}})$
  return any configuration in T_{left}

---

cannot be optimal. So we focus on the other case. Given two identical configurations $F_{\text{left}}(C)$ and $F'_{\text{left}}(C)$ with $\text{cost}(F_{\text{left}}(C)) < \text{cost}(F'_{\text{left}}(C))$, we show that $F'_{\text{left}}(C)$ cannot be optimal. Suppose there is an optimal solution F^* containing $F'_{\text{left}}(C)$, which means each execution segment $F'_{\text{left}}(J)$ in $F'_{\text{left}}(C)$ is completely contained by the corresponding execution interval of J in F^*. Since $F_{\text{left}}(C)$ and $F'_{\text{left}}(C)$ are identical, the start times of J are the same in the two configurations for all uncertain jobs J. In W_{left}, this means the uncertain jobs do not make the costs of the two configurations to be different, and the jobs $\mathcal{J}_c$ that are not uncertain do. Note that $\mathcal{J}_c$ is consisted of the jobs with their end times being determined. This means we can replace the configurations of $\mathcal{J}_c$ in $F'_{\text{left}}(C)$ by the configurations of $\mathcal{J}_c$ in $F_{\text{left}}(C)$ and this action will not affect the procedures in the algorithm thereafter. However, this also results in a solution of lower cost and contradicts the assumption that F^* is optimal. Thus $F'_{\text{left}}(C)$ cannot be optimal. Therefore, none of the deleted configuration can be part of an optimal schedule. That is, no optimal schedule would be removed through out the whole process. $\qquad\square$

Theorem 37. *Algorithm $\mathcal{E}$ computes an optimal solution in $O(k \cdot w_{\max}{}^{2m} \cdot (W_{\max} + 1)^{4m} \cdot n)$ time, where n is the number of jobs, $w_{\max}$ is the maximum width of jobs, m is the maximum size of cliques, $W_{\max}$ is the maximum length of windows, and k is the number of windows.*

Proof. We first compute the time complexities for the three components of the algorithm, and then compute the total time complexity. For ListConfigurations, there are at most $(W_{\max} + 1)^{2m}$ configurations in the outputted table T_{right}, since there are at most $W_{\max} + 1$ possible start times and end times respectively and at most m jobs that should be considered in the current window. For each configuration, it takes $O(n)$ time for construction and validity checking. It also takes $O(nW_{\max})$ to compute the cost of a configuration. So, the time complexity for ListConfigurations is

$$O((W_{\max} + 1)^{2m} \cdot nW_{\max}) = O((W_{\max} + 1)^{2m+1} \cdot n) \ .$$

Before computing the time complexities of the other components, we focus on the number of configurations of T_{left} at the end of each iteration in the algorithm. Since T_{left} is filtered to have no identical configurations, the number of configurations can be upper bounded. This number depends on the number of different start times of uncertain jobs. There are at most m uncertain jobs, and for each such job, the number of start times is at most $w_{\max}$. Note that the end times of these jobs are all set to be later than the current window and will not affect the number of configurations. So the number of configurations of T_{left} at the end of each iteration is at most $w_{\max}{}^m$.

For ConcatenateTables, there are at most $w_{\max}{}^m \cdot (W_{\max} + 1)^{2m}$ configurations in the outputted table T_{left}. This is because for each configuration in the input T_{left}, we need to compare it with all the configurations in T_{right} for compatibility checking. For each configuration, it takes $O(n)$ time for compatibility checking, concatenation and validity checking. Thus the time complexity for ConcatenateTables is $O(w_{\max}{}^m \cdot (W_{\max} + 1)^{2m} \cdot n)$.

21

For FilterTable, the number of configurations in the outputted table T_{left} is at most the number of configurations outputted by ConcatenateTables. Also, the number of groups is at most its number of configurations. Thus it takes

$$O([w_{\max}{}^m \cdot (W_{\max} + 1)^{2m}]^2 \cdot n) = O(w_{\max}{}^{2m} \cdot (W_{\max} + 1)^{4m} \cdot n)$$

time for classification. And it takes $O(w_{\max}{}^m \cdot (W_{\max} + 1)^{2m})$ time for deletion. So the time complexity for FilterTable is $O(w_{\max}{}^{2m} \cdot (W_{\max} + 1)^{4m} \cdot n)$. Since there are k iterations, the total time complexity is $O(k \cdot w_{\max}{}^{2m} \cdot (W_{\max} + 1)^{4m} \cdot n)$. $\qquad\square$

In the worst case, there are at most $O(n)$ windows. So algorithm $\mathcal{E}$ also runs in $f(w_{\max}, m, W_{\max}) \cdot O(n^2)$ time where $f(w_{\max}, m, W_{\max}) = w_{\max}{}^{2m} \cdot (W_{\max} + 1)^{4m}$.

Corollary 38. *Grid problem is fixed parameter tractable with respect to the maximum width of jobs, the maximum number of overlapped feasible intervals, and the maximum length of windows.*

6.1.4 An algorithm with two parameters

This section describes how to drop out the parameter $W_{\max}$ in the previous algorithm by generalizing the definitions of windows and boundaries.

At the beginning of Algorithm $\mathcal{E}$, we transform a set of jobs to its corresponding interval graph and obtain a sequence of windows by the set of maximal cliques in the interval graph. We require in the algorithm that all the cliques should be maximal. However, the algorithm is still optimal and has parameterized bound of time complexity if we divide a maximal clique into multiple non-maximal cliques in a specific way. Given a maximal clique C_i and its corresponding window W_i, we divide W_i into a set of contiguous windows $W_{i_1}, W_{i_2}, \ldots$ such that $W_i = \cup_j W_{i_j}$. Note that the set of jobs C_{i_j} corresponding to W_{i_j} is a clique in the interval graph since C_i is a clique and $C_{i_j} \subseteq C_i$. In this way, the number of jobs in the window W_{i_j} is still at most m. Furthermore, since this window division does not affect the proof of lemma 36, the algorithm is still optimal. Thus we have the following observation.

Observation 39. *Algorithm $\mathcal{E}$ outputs an optimal solution if it receives a set of contiguous windows containing all the jobs such that each window represents a clique (not necessarily maximal) in the interval graph of the input jobs. And we have the number of jobs in each window is at most the maximum number of overlapped feasible intervals.*

To drop out the parameter $W_{\max}$ in the previous algorithm, we divide windows into smaller ones such that the number of configurations in a window can be bounded by $w_{\max}$ and m. In the new algorithm, we set the locations of boundaries at the release times and deadlines of all the jobs and construct the windows bases on these boundaries. In this way, there is no job being released or attaining its deadline in the middle of a window, and all the jobs in the window can be put anywhere in the window. Thus the number of used timeslots is at most $m \cdot w_{\max} + 2(w_{\max} - 1)$. This is because in the worst case, all jobs in a window are scheduled such that no job overlaps to another and these jobs consume at most $m \cdot w_{\max}$ timeslots. In addition, we need to consider the cases that a job's start time is earlier than the window or its deadline is later than the window. Both cases consume at most $w_{\max} - 1$ timeslots respectively. Note that this window division results in a set of windows that their sizes are smaller than their original counterparts, and thus observation 39 can be applied. Based on this new window division, we have the following algorithm.

Algorithm $\mathcal{E}^+$. This algorithm is similar to algorithm $\mathcal{E}$ except the definitions of boundaries and the component ListConfigurations. Given a set of jobs $\mathcal{J}$, the algorithm uses the set of boundaries $\{r(J) \mid J \in \mathcal{J}\} \cup \{d(J) \mid J \in \mathcal{J}\}$ to construct the windows and obtain the corresponding cliques. Let k denotes by the number of windows. There are k stages for the algorithm. At the i-th stage, the algorithm runs ListConfigurations, ConcatenateTables and

FilterTable accordingly as algorithm 1 does. It finally outputs the only configuration in T_{left}. For the component ListConfigurations, we only consider to schedule jobs on the timeslots used instead of enumerating all possibilities of start times and end times. The algorithm tries all $m \cdot w_{\text{max}}$ timeslots (the worst case described in the previous paragraph) as the start time of a job, and also the $2(w_{\text{max}} - 1)$ schedules that a job is partially executed in the window. In addition, the component shall includes the cases that either a job is completely executed before the window, it is completely executed after the window, or it crosses the window.

Theorem 40. *Algorithm* $\mathcal{E}^+$ *computes an optimal solution in* $f(w_{\text{max}}, m) \cdot O(n^2)$ *time, where n is the number of jobs, w_{max} is the maximum width of jobs, m is the maximum size of cliques, and* $f(w_{\text{max}}, m) = (4m \cdot w_{\text{max}}^2)^{2m}$.

Proof. As in the proof of theorem 37, we compute the running time of the three components and then the total time complexity. For the component ListConfigurations, there are at most $(m \cdot w_{\text{max}} + 2(w_{\text{max}} - 1) + 3)^m$ outputted configurations, since there are at most $m \cdot w_{\text{max}} + 2(w_{\text{max}} - 1) + 3$ schedules for a job (see the description in the previous paragraph) and at most m jobs in a window. It takes $O(n(m \cdot w_{\text{max}} + 2(w_{\text{max}} - 1))) \leq O(n \cdot m \cdot w_{\text{max}})$ time to compute the cost for each configuration. Thus the time complexity for ListConfigurations is at most

$$O((m \cdot w_{\text{max}} + 2(w_{\text{max}} - 1) + 3)^m \cdot (n \cdot m \cdot w_{\text{max}})) \leq O((4m \cdot w_{\text{max}})^{m+1} \cdot n) \ .$$

The time complexities of ConcatenateTables and FilterTable are similar to that in the proof of theorem 37 except the number of outputted configurations. For ConcatenateTables and FilterTable, both the number of outputted configurations are at most $w_{\text{max}}^m \cdot (4m \cdot w_{\text{max}})^m$. Thus their running time are at most $O(w_{\text{max}}^{2m} \cdot (4m \cdot w_{\text{max}})^{2m} \cdot n)$. Since there are $k = O(n)$ iterations, the total time complexity of the algorithm is at most

$$O((4m \cdot w_{\text{max}}^2)^{2m} \cdot n^2) = f(w_{\text{max}}, m) \cdot O(n^2) \ . \qquad \square$$

Corollary 41. *Grid problem is fixed parameter tractable with respect to the maximum width of jobs, and the maximum number of overlapped feasible intervals.*

6.2 An exact algorithm without parameter

For the case with unit width and arbitrary height of Grid problem, we can use algorithm $\mathcal{E}$ to design an exact algorithm that its time complexity is only measured in the size of the input.

In Algorithm $\mathcal{E}$, we maintain two tables T_{left} and T_{right} for each stage. At each stage, the core operations are to construct T_{right}, merge T_{left} and T_{right}, and filter the resulting table. In the case with unit width and arbitrary height, one may observe that the functionalities of these core operations are not affected by the length of the windows representing T_{left} and T_{right}. For example, we can restrict the window length to be a constant but not be related to the cliques in the interval graph, and the algorithm still works correctly. By fixing the lengths of all windows, a new exact algorithm is obtained. Without loss of generality, we assume that the number of timeslots τ is even. We enforce all windows to have length 2, i.e. we have $\tau/2$ windows in total. By this setting, the new algorithm runs in $O((\tau/2) \cdot 4^{2n} \cdot n)$ time where n is the number of jobs. This is because the numbers of configurations for the three components in the algorithm are at most 4^n. Note that the input size N of the problem is $3n \log \tau + n \log h_{\text{max}}$ where h_{max} is the maximum height over all jobs. Since $\log \tau = O(N)$, the running time becomes $2^{O(N)}$. Thus we have the following theorem.

Theorem 42. *There is an exact algorithm running in* $2^{O(N)}$ *time for the Grid problem with unit width and arbitrary height where N is the length of the input.*

Our algorithm is highly more efficient than a brute force search. Such naive method would enumerate all possible schedules and check if they are feasible and optimal, which requires $O(\tau^n n)$ time. The running time can be rewritten as $2^{O(Nn)}$ or more clearly, $(2^{O(N)})^n$. The exact algorithm modified from our fixed parameter algorithm indeed crosses out an 'n' in the exponent.

6.3 Lower bound

This section provides two lower bounds on the running time of the Grid problem under a certain condition.

Jansen et al. [24] derived several lower bounds for scheduling and packing problems which can be used to develop lower bounds for our problem. Their lower bounds assume *Exponential Time Hypothesis* (ETH) holds, which conjectures that there is a positive real ϵ such that 3-Sat cannot be decided in time $2^{\epsilon n} N^{O(1)}$ where n is the number of variables in the formula and N is the length of the input. A lower bound for other problems can be shown by making use of strong reductions, i.e. reductions that increase the parameter at most linearly. Through a sequence of strong reductions, they obtain two lower bounds for Partition, $2^{o(n)} N^{O(1)}$ and $2^{o(\sqrt{N})}$ where n is the cardinality of the given set and N is the length of the input.

Reduction. We design a strong reduction from Partition to the decision version of Grid problem with unit width and arbitrary height. Here is a sketch of the reduction. Recall that Partition is a decision problem that decides if a given set S of integers can be partitioned into two disjoint subsets such that the two subsets have equal sum. For each integer $s \in S$, we convert it to a job J with $r(J) = 0$, $d(J) = 2$, $w(J) = 1$ and $h(J) = 2s$. We claim that S is a partition if and only if the set of jobs J can be scheduled with cost at most $2(\sum_{s \in S} s)^\alpha$. Note that the specified cost appears when jobs can be put into two timeslots with equal loads. By setting the length of the input as the parameter, we observe that the parameter increases at most linearly from Partition to our problem. (Note that a strong reduction from Partition to the case with unit height and arbitrary width can be done similarly, and the results also apply on that case.) Furthermore, we can choose the number of jobs as a parameter of the problem. Note that the reduction above does not increase this parameter with respect to the parameter of Partition, which is the number of integers.

Theorem 43. *There is a lower bound of $2^{o(\sqrt{N})}$ and a lower bound of $2^{o(n)} N^{O(1)}$ on the running time for the Grid problem unless ETH fails, where n is the number of jobs and N is the length of the input.*

7 Minimizing peak and non-preemptive machine minimization

In this section, we investigate extension of our solutions to other objectives and other problems. In particular, we consider the objective of minimizing peak electricity cost in the GRID model and we name it the GRID$_{\text{peak}}$ problem. We also consider the classical non-preemptive machine minimization problem denoted as MACHINE.

The GRID$_{\text{peak}}$ problem. The input is the same as the GRID problem. The goal is to find a feasible non-preemptive schedule such that the maximum load over the time horizon is minimized. GRID$_{\text{peak}}$ has been proven to be NP-hard [45] and approximation algorithms are known for requests having common feasible interval with approximation ratio 4 and for requests having agreeable deadlines with approximation ratio $O(\log \frac{w_{\max}}{w_{\min}})$.

The MACHINE problem. The input is a set of jobs each with a processing time $p(J)$, release time $r(J)$, and deadline $d(J)$. Each job has to be scheduled non-preemptively on one of the (infinite number of) machines. For each machine, at most one job can be executed at any time. The goal is to minimize the number of machines used.

The MACHINE problem can be seen as a special case of the $\mathsf{GRID}_{\mathsf{peak}}$ problem where jobs have unit height. A lower bound of $\log_3 \frac{w_{\max}}{w_{\min}}$ on the competitive ratio of any online algorithm has been shown in [42]. As a result, this lower bound also applies to $\mathsf{GRID}_{\mathsf{peak}}$. The author also provided an $O(\log \frac{p_{\max}}{p_{\min}})$-competitive algorithm for the MACHINE problem. The algorithm classifies jobs by processing times and applies the constant approximate algorithm for each class.[2] The employed algorithm can be the 6-competitive algorithm proposed by Yu and Zhang [52] or our algorithm for uniform widths jobs (to be analyzed in Section 7.1.2).

In this section, we show that our online algorithm solves the $\mathsf{GRID}_{\mathsf{peak}}$ problem optimally in an asymptotical sense and provide an alternative asymptotically optimal competitive algorithm for the MACHINE problem.

7.1 Online algorithms

In this section, we prove that the online algorithm $\mathcal{G}$ proposed in Section 4 is asymptotically optimal for the $\mathsf{GRID}_{\mathsf{peak}}$ problem. We first state two properties, one for $\mathcal{BKP}'$ that $\mathcal{G}$ is based on and the other for the optimal algorithm w.r.t. the peak objective. Let function $\mathrm{peak}(S)$ denote the maximum load (cf. speed) of any schedule S, i.e., $\mathrm{peak}(S) = \max_t \ell(S, t)$. Combining Lemma 8 and the fact that $\mathcal{BKP}$ is e-competitive w.r.t. maximum speed [4], we have the following observation.

Observation 44. *The $\mathcal{BKP}'$ algorithm is $e(1 + e)$-competitive with respect to maximum speed.*

Proof. By Lemma 8, for any integral t and $0 < \Delta < 1$, $\ell(\mathcal{BKP}, t + \Delta) \leq (1 + e) \cdot \ell(\mathcal{BKP}, t)$. Also, by $\mathcal{BKP}'$, $\ell(\mathcal{BKP}', t) \geq \ell(\mathcal{BKP}, t + \Delta)$ for any $0 < \Delta < 1$. Hence, $\mathrm{peak}(\mathcal{BKP}') \leq (1 + e) \cdot \mathrm{peak}(\mathcal{BKP})$. $\square$

On the other hand, $\mathcal{YDS}$ guarantees that the maximum speed is minimized [4]. Similar to Observation 1, $\mathcal{YDS}$ gives a lower bound for the $\mathsf{GRID}_{\mathsf{peak}}$ problem.

Observation 45. *Let $\mathcal{O}_D$ and $\mathcal{O}_G$ be the optimal schedule for the DVS and $\mathsf{GRID}_{\mathsf{peak}}$ problem, respectively. Given a job set $\mathcal{J}_G$ for the $\mathsf{GRID}_{\mathsf{peak}}$ problem. Let $\mathcal{J}_D$ denote the job set after converting $\mathcal{J}_G$ into a job set for the DVS problem. Then, $\mathrm{peak}(\mathcal{O}_D(\mathcal{J}_D)) \leq \mathrm{peak}(\mathcal{O}_G(\mathcal{J}_G))$.*

Recall that in Sections 3 and 4, we have presented three algorithms $\mathcal{V}$, $\mathcal{UV}$, and $\mathcal{G}$ for unit-width jobs, uniform-width jobs, and arbitrary width jobs, respectively. Here, we analyze their performance w.r.t. $\mathsf{GRID}_{\mathsf{peak}}$. In summary, we show that for $\mathsf{GRID}_{\mathsf{peak}}$, we have

- $\mathcal{V}$ is $2(e + e^2)$-competitive for unit-width job sets (Theorem 46);

- $\mathcal{UV}$ is $(6(e + e^2) + 3)$-competitive for unit-width job sets (Theorem 49); and

- $\mathcal{G}$ is $(18(e + e^2) + 9) \cdot \lceil \log \frac{w_{\max}}{w_{\min}} \rceil$-competitive for arbitrary-width job sets (Theorem 50).

7.1.1 Unit-width jobs.

Recall that for each timeslot t, $\mathcal{V}$ schedules jobs to start at t such that $\ell(\mathcal{V}, t)$ is at least $\ell(\mathcal{BKP}', t) = (1 + e) \cdot \ell(\mathcal{BKP}, t)$ or until all available jobs have been scheduled. We prove that although $\ell(\mathcal{V}, t)$ might be higher than $\ell(\mathcal{BKP}', t)$, the peak of $\mathcal{V}$ is no more than $2(e + e^2)$ times of the peak of the optimal.

Theorem 46. *For any job set $\mathcal{J}$ where for each job has unit width, $\mathrm{peak}(\mathcal{V}(\mathcal{J})) \leq 2(e + e^2) \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}))$.*

---

[2]The paper did not state explicitly which algorithm to use and a result [11] cited in the paper has later been retreated by the same authors [10].

Proof. Let $h_{\max}(\mathcal{V}, t)$ be the maximum height of jobs scheduled at t by $\mathcal{V}$; we set $h_{\max}(\mathcal{V}, t) = 0$ if $\mathcal{V}$ assigns no job at t. We classify each timeslot t into two types: (i) $h_{\max}(\mathcal{V}, t) < \ell(\mathcal{BKP}', t)$, and (ii) $h_{\max}(\mathcal{V}, t) \geq \ell(\mathcal{BKP}', t)$. We denote by $\mathcal{I}_1$ and $\mathcal{I}_2$ the union of all timeslots of Type (i) and (ii), respectively. (Notice that $\mathcal{I}_1$ and $\mathcal{I}_2$ can be empty and the union of $\mathcal{I}_1$ and $\mathcal{I}_2$ covers the entire time line.)

We first prove that for any job set $\mathcal{J}$ where for each job $J \in \mathcal{J}$, $w(J) = 1$, $\mathrm{peak}(\mathcal{V}(\mathcal{J}), \mathcal{I}_1) \leq 2(e + e^2) \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}))$; and $\mathrm{peak}(\mathcal{V}(\mathcal{J}), \mathcal{I}_2) \leq 2 \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}))$. For every timeslot $t \in \mathcal{I}_1$, $\ell(\mathcal{V}, t) < \ell(\mathcal{BKP}', t) + h_{\max}(\mathcal{V}, t) \leq 2 \cdot \ell(\mathcal{BKP}', t) \leq 2(1 + e) \cdot \ell(\mathcal{BKP}, t)$. By definition, let t bet the timeslot where $\ell(\mathcal{V}, t) \geq \ell(\mathcal{V}, t')$ for any other t', $\mathrm{peak}(\mathcal{V}, \mathcal{I}_1) = \ell(\mathcal{V}, t) \leq 2(1 + e) \cdot \ell(\mathcal{BKP}, t)$. Therefore, $\mathrm{peak}(\mathcal{V}, \mathcal{I}_1) \leq 2e(1 + e) \cdot \mathrm{peak}(\mathcal{O})$ by Observation 45 and Lemma 44. For every timeslot $t \in \mathcal{I}_2$, $\ell(\mathcal{V}, t) < \ell(\mathcal{BKP}', t) + h_{\max}(\mathcal{V}, t) \leq 2 \cdot h_{\max}(\mathcal{V}, t)$. In the optimal schedule, the job with height $h_{\max}(\mathcal{V}, t)$ has to be scheduled somewhere or the schedule is not feasible, so $\mathrm{peak}(\mathcal{O}) \geq h_{\max}(\mathcal{V}, t)$. Hence, $\mathrm{peak}(\mathcal{V}, \mathcal{I}_2) = \ell(\mathcal{V}, t) \leq 2h_{\max}(\mathcal{V}, t) \leq 2 \cdot \mathrm{peak}(\mathcal{O})$ where t is the timeslot with peak power.

Since $\mathcal{I}_1$ and $\mathcal{I}_2$ are disjoiont, $\mathrm{peak}(\mathcal{V}) = \max\{\mathrm{peak}(\mathcal{V}, \mathcal{I}_1), \mathrm{peak}(\mathcal{V}, \mathcal{I}_2)\}$. Therefore, $\mathrm{peak}(\mathcal{V}) = \max\{2(e + e^2) \cdot \mathrm{peak}(\mathcal{O}), 2 \cdot \mathrm{peak}(\mathcal{O})\} = 2(e + e^2) \cdot \mathrm{peak}(\mathcal{O})$. $\qquad\square$

7.1.2 Uniform-width jobs.

Recall that in handling uniform-width jobs, we classify jobs into tight and loose jobs. Let $\mathcal{J}^*$ denote the input set with uniform width jobs, $\mathcal{J}_{\mathrm{T}}^*$ and $\mathcal{J}_{\mathrm{L}}^*$ denote the set of tight jobs and loose jobs in $\mathcal{J}^*$, respectively. We first prove that any feasible schedule for tight jobs is 3-competitive due to the "inflexibility" (Lemma 47). Then, we prove that $\mathcal{UV}$ is $O(1)$-competitive for loose jobs (Lemma 48).

Lemma 47. *For any feasible schedule S, $peak(S(\mathcal{J}_T^*)) \leq 3 \cdot peak(\mathcal{O}(\mathcal{J}^*))$.*

Proof. We prove it by showing that even if the execution intervals of jobs are considered as the whole feasible interval, the peak is not too much larger than the peak in the optimal schedule.

We first *extend* jobs $J \in \mathcal{J}_{\mathrm{T}}^*$ to J^+ as follows: $r(J^*) = r(J)$, $d(J^*) = d(J)$, $w(J^*) = d(J) - r(J)$, and $h(J^*) = h(J)$. That is, every job has its width as the length of its feasible interval. We denote the resulting job set by $\mathcal{J}^+$. Since each job in $\mathcal{J}^+$ are not shiftable, there is only one feasible schedule for $\mathcal{J}^+$ and it is optimal. It is clear that $\mathrm{peak}(S(\mathcal{J}_{\mathrm{T}}^*)) \leq \mathrm{peak}(\mathcal{O}(\mathcal{J}^+))$.

Similar to Lemma 16 (i), we bound the load at time t of $\mathcal{O}(\mathcal{J}^+)$ by the loads of constant number of timeslots in $S(\mathcal{J}_{\mathrm{T}}^*)$. Consider the job J corresponding to J^+, the execution interval of J in any feasible schedule must contain either timeslot $t - (w - 1)$, $t + (w - 1)$, or t. Hence, we can upper bound the load at time t in $\mathcal{O}(\mathcal{J}^+)$ as follows: $\ell(\mathcal{O}(\mathcal{J}^+), t) \leq \ell(\mathcal{O}(\mathcal{J}_{\mathrm{T}}^*), t - (w - 1)) + \ell(\mathcal{O}(\mathcal{J}_{\mathrm{T}}^*), t + (w - 1)) + \ell(\mathcal{O}(\mathcal{J}_{\mathrm{T}}^*), t)$. Hence, $\mathrm{peak}(S(\mathcal{J}_{\mathrm{T}}^*)) \leq \mathrm{peak}(\mathcal{O}(\mathcal{J}^+)) \leq 3 \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}_{\mathrm{L}}^*))$. $\qquad\square$

Lemma 48. *For loose jobs set $\mathcal{J}_L^*$ where jobs have uniform width, $peak(\mathcal{UV}(\mathcal{J}_L^*)) \leq 6(e + e^2) \cdot peak(\mathcal{O}(\mathcal{J}^*))$.*

Proof. According to $\mathcal{UV}$, the job set $\mathcal{J}_{\mathrm{L}}^*$ is transformed into a job set $\mathcal{J}'$ by ALIGNFI and $\mathcal{V}$ is run on $\mathcal{J}'$. Then, the schedule $\mathcal{V}(\mathcal{J}')$ is transformed to a schedule of $\mathcal{J}_{\mathrm{L}}^*$ by Transformation FREESCH. Hence, by Theorem 46, $\mathrm{peak}(\mathcal{UV}(\mathcal{J}_{\mathrm{L}}^*)) \leq \mathrm{peak}(\mathcal{V}(\mathcal{J}')) \leq 2(e + e^2) \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}'))$.

By Observation 11, given the optimal schedule of $\mathcal{J}_{\mathrm{L}}^*$, we have shown the load at any time in the schedule S' generated by ALIGNSCH is no more than the sum of the load of the optimal schedule at three timeslots and hence no more than three times the peak of the optimal schedule. Therefore, $\mathrm{peak}(\mathcal{O}(\mathcal{J}')) \leq \mathrm{peak}(S') \leq 3 \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}_{\mathrm{L}}^*))$

In summary, $\mathrm{peak}(\mathcal{UV}(\mathcal{J}_{\mathrm{L}}^*)) \leq 2(e + e^2) \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}')) \leq 6(e + e^2) \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}_{\mathrm{L}}^*))$. $\qquad\square$

Theorem 49. *For any jobs set $\mathcal{J}^*$ where jobs have uniform width, $peak(\mathcal{UV}(\mathcal{J}^*)) \leq (6(e + e^2) + 3) \cdot peak(\mathcal{O}(\mathcal{J}^*))$.*

Proof. By definition, $\mathrm{peak}(\mathcal{UV}(\mathcal{J}^*)) \leq \mathrm{peak}(\mathcal{UV}(\mathcal{J}_T^*)) + \mathrm{peak}(\mathcal{UV}(\mathcal{J}_L^*))$. By Lemma 47 and 48, $\mathrm{peak}(\mathcal{UV}(\mathcal{J}^*)) \leq 3 \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}^*)) + 6(e+e^2) \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}^*)) = (6(e+e^2)+3) \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}^*))$. $\square$

7.1.3   Arbitrary width jobs.

Finally, we bound the competitive ratio of $\mathcal{G}$.

Theorem 50. *For any job set $\mathcal{J}$, $peak(\mathcal{G}(\mathcal{J})) \leq (18(e+e^2)+9) \cdot \lceil \log \frac{w_{\max}}{w_{\min}} \rceil \cdot peak(\mathcal{O}(\mathcal{J}))$.*

Proof. Recall that $\mathcal{G}(\mathcal{J})$ partition jobs into subsets $\mathcal{J}_p$ such that in each $\mathcal{J}_p$ jobs have bounded widths. For each $\mathcal{J}_p$, it is transformed into $\mathcal{J}_p^*$ and $\mathcal{UV}$ is applied independently on each class. Then, $\mathcal{UV}(\mathcal{J}_p^*)$ is transformed into a schedule for $\mathcal{J}_p$ by Transformation SHRINKSCH. By Observation 22, $\ell(\mathcal{G}(\mathcal{J}_p), t) \leq \ell(\mathcal{UV}(\mathcal{J}_p^*), t)$. Hence, $\ell(\mathcal{G}(\mathcal{J}), t) = \sum_p \ell(\mathcal{G}(\mathcal{J}_p), t) \leq \sum_p \ell(\mathcal{UV}(\mathcal{J}_p^*), t)$ for all t.

Consider the timeslot t with peak load in $\mathcal{G}(\mathcal{J})$, $\mathrm{peak}(\mathcal{G}(\mathcal{J})) = \ell(\mathcal{G}(\mathcal{J}), t) \leq \sum_p \ell(\mathcal{UV}(\mathcal{J}_p^*), t) \leq \sum_p \mathrm{peak}(\mathcal{UV}(\mathcal{J}_p^*))$. By Theorem 49, $\mathrm{peak}(\mathcal{G}(\mathcal{J})) \leq \sum_p (6(e+e^2)+3) \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}_p^*))$.

Now we prove that $\mathrm{peak}(\mathcal{O}(\mathcal{J}_p^*)) \leq 3 \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}))$. Consider the optimal schedule $\mathcal{O}(\mathcal{J}_p)$, there exists schedule $S(\mathcal{J}_p^*)$ generated by Transformation RELAXSCH where $\mathcal{J}_p^*$ is the job set corresponding to $\mathcal{J}_p$ generated by CONVERT. By Lemma 20, the load of any timeslot in $S(\mathcal{J}_p^*)$ is no more than the sum of loads of three timeslots in $\mathcal{O}(\mathcal{J}_p)$, and hence no more than three times of the peak in $\mathcal{O}(\mathcal{J}_p)$. Therefore, $\mathrm{peak}(\mathcal{O}(\mathcal{J}_p^*)) \leq \mathrm{peak}(S(\mathcal{J}_p^*)) \leq 3 \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}_p)) \leq 3 \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}))$.

In summary, $\mathrm{peak}(\mathcal{G}(\mathcal{J})) \leq \sum_p (6(e+e^2)+3) \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}_p^*)) \leq \sum_p (6(e+e^2)+3) \cdot 3 \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J})) \leq (18(e+e^2)+9) \cdot \lceil \log \frac{w_{\max}}{w_{\min}} \rceil \cdot \mathrm{peak}(\mathcal{O}(\mathcal{J}))$. $\square$

Since the MACHINE problem is a special case of the $\mathsf{GRID_{peak}}$ problem where jobs have uniform height, we have the following corollary:

Corollary 51. *The algorithm $\mathcal{G}$ is $(18(e+e^2)+9)$-competitive for the MACHINE problem.*

7.2   The interval graph approach on the $\mathsf{GRID_{peak}}$ problem

In Section 6.1, we introduced an exact algorithm $\mathcal{E}$ using the *linear clique arrangement* property of the interval graphs. The linear property of the consecutive clique arrangement of interval graphs gives a direction to design a dynamic programming algorithm, which breaks down a problem into overlapped subproblems until the subproblems are simple enough to be solved. In the following of this section, we show that $\mathcal{E}$ can be used to solve the $\mathsf{GRID_{peak}}$ problem by changing the objective function.

Algorithm $\mathcal{E}_{peak}$ for the $\mathsf{GRID_{peak}}$ problem (also see Section 6.1). The jobs are considered as time intervals and the time horizon is chopped into "windows". The algorithm visits all windows accordingly from the left to right. In Stage i, the i-th window is visited and the algorithm maintains a candidate set of schedules for the visited windows that no optimal solution is deleted from the set. In each Stage i, the algorithm consists of three procedures: ListConfigurations, ConcatenateTables and FilterTable.

The ListConfigurations procedure lists all possible configurations (i.e., execution segments) of the jobs in C_i within W_i. The invalid configurations will be deleted. The valid configurations together with *the peak load within W_i* will be stored in a table.

The ConcatenateTables procedure concatenates the configurations in the current window W_i and the configurations in the windows which have been seen so far. If the execution interval after concatenation is not valid, it is deleted from the table. *The peak load of the new configuration is simply the maximum among the peaks of the two concatenated configurations.*

The FilterTable procedure filters non-optimal configurations. The idea is, given a configuration of the jobs in C_i, there must be a best decision of the jobs in $\bigcup_{k=1}^{i-1} C_k \setminus C_i$ which has minimum *peak load* within the intervals $[0, b_{i+1})$, where b_{i+1} is the right boundary of the window W_i. For each configuration, we only keep the (partial) schedule with the minimum *peak load*.

After processing all the windows, the schedule with minimum *peak load* can be found in the final table.

It can be seen that we list all possible configurations. A configuration is deleted only when it is invalid or it is identical to another configuration with lower peak peak. Hence in the end we get an optimal schedule. It also shows that the $\mathsf{GRID}_{\mathsf{peak}}$ problem is fixed parameter tractable with respect to the maximum width of jobs and the maximum number of overlapped feasible intervals, and the maximum length of windows.

Corollary 52. *The $\mathsf{GRID}_{\mathsf{peak}}$ problem is fixed parameter tractable with respect to the maximum width of jobs, and the maximum number of overlapped feasible intervals.*

8 Conclusion

We develop the first online algorithm with polylog-competitive ratio and the first FPT algorithms for non-preemptive smart grid scheduling problem in general case. We also derive matching lower bound for the competitive ratio. Constant competitive online algorithms are presented for several special input instances.

There are quite a few directions in extending the problem setting: different cost functions perhaps to capture varying electricity cost over time of the day; jobs with varying power requests during its execution (it is a constant value in this paper); other objectives like response time. A preliminary result is that we can extend our online algorithm to the case where a job may have varying power requests during its execution, in other words, a job can be viewed as having rectilinear shape instead of being rectangular. In such case, the competitive ratio is increased by a factor depending on the maximum height to the minimum height ratio of a job. Precisely, the competitve ratio becomes $(36H\lceil \log K \rceil)^{\alpha} \cdot (8e^{\alpha} + 1)$, where $H = \max_J \frac{h_H(J)}{h_L(J)}$, $h_H(J)$ and $h_L(J)$ are respectively the highest and lowest power request of job J.

Acknowledgement

References

[1] S. Alamdari, T. Biedl, T. M. Chan, E. Grant, K. R. Jampani, S. Keshav, A. Lubiw, and V. Pathak. Smart-grid electricity allocation via strip packing with slicing. In *WADS*, pages 25–36. Springer, 2013.

[2] S. Albers. Energy-efficient algorithms. *Commun. ACM*, 53(5):86–96, 2010.

[3] Y. Azar. On-line load balancing. In *Online Algorithms*, pages 178–195. Springer, 1998.

[4] N. Bansal, T. Kimbrel, and K. Pruhs. Speed scaling to manage energy and temperature. *J. ACM*, 54(1), 2007.

[5] P. C. Bell and P. W. H. Wong. Multiprocessor speed scaling for jobs with arbitrary sizes and deadlines. *J. Comb. Optim.*, 29(4):739–749, 2015.

[6] M. Burcea, W. Hon, H. Liu, P. W. H. Wong, and D. K. Y. Yau. Scheduling for electricity cost in a smart grid. *J. Scheduling*, 19(6):687–699, 2016.

[7] S. Caron and G. Kesidis. Incentive-based energy consumption scheduling algorithms for the smart grid. In *SmartGridComm*, pages 391–396. IEEE, 2010.

[8] C. Chen, K. G. Nagananda, G. Xiong, S. Kishore, and L. V. Snyder. A communication-based appliance scheduling scheme for consumer-premise energy management systems. *IEEE Trans. Smart Grid*, 4(1):56–65, 2013.

[9] L. Chen, N. Megow, and K. Schewior. An $O(\log m)$-competitive algorithm for online machine minimization. In *SODA*, pages 155–163. ACM, 2016.

[10] J. Chuzhoy and P. Codenotti. Erratum: Resource minimization job scheduling. In *APPROX-RANDOM*, page E1. Springer, 2009.

[11] J. Chuzhoy and P. Codenotti. Resource minimization job scheduling. In *APPROX-RANDOM*, pages 70–83. Springer, 2009.

[12] J. Chuzhoy, S. Guha, S. Khanna, and J. S. Naor. Machine minimization for scheduling jobs with interval constraints. In *FOCS*, pages 81–90. IEEE, 2004.

[13] M. Cieliebak, T. Erlebach, F. Hennecke, B. Weber, and P. Widmayer. Scheduling with release times and deadlines on a minimum number of machines. In *IFIP TCS*, pages 209–222. Springer, 2004.

[14] M. Ž. Djurović, A. Milačić, and M. Kršulja. A simplified model of quadratic cost function for thermal generators. In *DAAAM*, pages 25–28. DA, 2012.

[15] European Commission. Europen smartgrids technology platform. `ftp://ftp.cordis.europa.eu/pub/fp7/energy/docs/smartgrids_en.pdf`, 2006.

[16] K. Fang, N. A. Uhan, F. Zhao, and J. W. Sutherland. Scheduling on a single machine under time-of-use electricity tariffs. *Annals OR*, 238(1-2):199–227, 2016.

[17] X. Fang, S. Misra, G. Xue, and D. Yang. Smart grid – the new and improved power grid: A survey. *IEEE Communications Surveys and Tutorials*, 14(4):944–980, 2012.

[18] H. Farhangi. The path of the smart grid. *IEEE Power and Energy Mag.*, 8(1):18–28, 2010.

[19] X. Feng, Y. Xu, and F. Zheng. Online scheduling for electricity cost in smart grid. In *COCOA*, pages 783–793. Springer, 2015.

[20] D. Fulkerson and O. Gross. Incidence matrices and interval graphs. *Pacific Journal of Mathematics*, 15(3):835–855, 1965.

[21] R. Halin. Some remarks on interval graphs. *Combinatorica*, 2(3):297–304, 1982.

[22] K. Hamilton and N. Gulhar. Taking demand response to the next level. *IEEE Power and Energy Mag.*, 8(3):60–65, 2010.

[23] A. Ipakchi and F. Albuyeh. Grid of the future. *IEEE Power & Energy Mag.*, 7(2):52–62, 2009.

[24] K. Jansen, F. Land, and K. Land. Bounding the running time of algorithms for scheduling and packing problems. *SIAM J. Discrete Math.*, 30(1):343–366, 2016.

[25] L. Kannberg, D. Chassin, J. DeSteese, S. Hauser, M. Kintner-Meyer, P. N. N. L. (U.S.), and U. S. D. of Energy. *GridWise: The Benefits of a Transformed Energy System*. Pacific Northwest National Laboratory, 2003.

[26] M. M. Karbasioun, G. Shaikhet, E. Kranakis, and I. Lambadaris. Power strip packing of malleable demands in smart grid. In *ICC*, pages 4261–4265. IEEE, 2013.

[27] I. Koutsopoulos and L. Tassiulas. Control and optimization meet the smart power grid: Scheduling of power demands for optimal energy management. In *e-Energy*, pages 41–50. ACM, 2011.

[28] R. Krishnan. Meters of tomorrow [in my view]. *IEEE Power and Energy Mag.*, 6(2):96–94, 2008.

[29] H. Li and R. C. Qiu. Need-based communication for smart grid: When to inquire power price? In *GLOBECOM*, pages 1–5. IEEE, 2010.

[30] Z. Li and Q. Liang. Performance analysis of multiuser selection scheme in dynamic home area networks for smart grid communications. *IEEE Trans. Smart Grid*, 4(1):13–20, 2013.

[31] F. Liu, H. Liu, and P. W. H. Wong. Optimal nonpreemptive scheduling in a smart grid model. In *ISAAC*, pages 53:1–53:13. LIPIcs, 2016.

[32] A. Llaria, J. Jiménez, and O. Curea. Study on communication technologies for the optimal operation of smart grids. *Trans. Emerging Telecommunications Technologies*, 25(10):1009–1019, 2014.

[33] Lockheed Martin. SEELoad™ Solution. `http://www.lockheedmartin.co.uk/us/products/energy-solutions/seesuite/seeload.html`.

[34] T. Logenthiran, D. Srinivasan, and T. Z. Shun. Demand side management in smart grid using heuristic optimization. *IEEE Trans. Smart Grid*, 3(3):1244–1252, 2012.

[35] T. Lui, W. Stirling, and H. Marcy. Get smart. *IEEE Power & Energy Mag.*, 8(3):66–78, 2010.

[36] C. Y. T. Ma, D. K. Y. Yau, and N. S. V. Rao. Scalable solutions of markov games for smart-grid infrastructure protection. *IEEE Trans. Smart Grid*, 4(1):47–55, 2013.

[37] S. Maharjan, Q. Zhu, Y. Zhang, S. Gjessing, and T. Basar. Dependable demand response management in the smart grid: A stackelberg game approach. *IEEE Trans. Smart Grid*, 4(1):120–132, 2013.

[38] G. M. Masters. *Renewable and efficient electric power systems*. John Wiley & Sons, 2013.

[39] A.-H. Mohsenian-Rad, V. Wong, J. Jatskevich, and R. Schober. Optimal and autonomous incentive-based energy consumption scheduling algorithm for smart grid. In *ISGT*, pages 1–6. IEEE, 2010.

[40] A.-H. Mohsenian-Rad, V. W. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia. Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid. *IEEE Trans. Smart Grid*, 1(3):320–331, 2010.

[41] REGEN Energy Inc. ENVIROGRID™ SMART GRID BUNDLE. `http://www.regenenergy.com/press/announcing-the-envirogrid-smart-grid-bundle/`.

[42] B. Saha. Renting a cloud. In *FSTTCS*, pages 437–448. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2013.

[43] S. Salinas, M. Li, and P. Li. Multi-objective optimal energy consumption scheduling in smart grids. *IEEE Trans. Smart Grid*, 4(1):341–348, 2013.

[44] P. Samadi, A.-H. Mohsenian-Rad, R. Schober, V. W. Wong, and J. Jatskevich. Optimal real-time pricing algorithm based on utility maximization for smart grid. In *SmartGridComm*, pages 415–420. IEEE, 2010.

[45] S. Tang, Q. Huang, X.-Y. Li, and D. Wu. Smoothing the energy consumption: Peak demand reduction in smart grid. In *INFOCOM*, pages 1133–1141. IEEE, 2013.

[46] Toronto Hydro Corporation. Peaksaver Program. `http://www.peaksaver.com/peaksaver_THESL.html`.

[47] UK Department of Energy & Climate Change. Smart grid: A more energy-efficient electricity supply for the UK. `https://www.gov.uk/smart-grid-a-more-energy-efficient-electricity-supply-for-the-uk`, 2013.

[48] US Department of Energy. The Smart Grid: An Introduction. `http://www.oe.energy.gov/SmartGridIntroduction.htm`, 2009.

[49] F. Yao, A. Demers, and S. Shenker. A scheduling model for reduced CPU energy. In *FOCS*, pages 374–382. IEEE, 1995.

[50] S. Yaw and B. Mumey. An exact algorithm for non-preemptive peak demand job scheduling. In *COCOA*, pages 3–12. Springer, 2014.

[51] S. Yaw, B. Mumey, E. McDonald, and J. Lemke. Peak demand scheduling in the smart grid. In *SmartGridComm*, pages 770–775. IEEE, 2014.

[52] G. Yu and G. Zhang. Scheduling with a minimum number of machines. *Oper. Res. Lett.*, 37(2):97–101, 2009.

[53] Zpryme Research & Consulting. Power systems of the future: The case for energy storage, distributed generation, and microgrids. `http://smartgrid.ieee.org/images/features/smart_grid_survey.pdf`, 2012.