Approximating weighted completion time via stronger negative correlation

Citation: Baveja, A., Qu, X. & Srinivasan, A. Approximating weighted completion time via stronger negative correlation. J Sched 27, 319–328 (2024).

As Published: https://doi.org/10.1007/s10951-023-00780-y

Publisher: Springer US

Persistent URL: https://hdl.handle.net/1721.1/156081

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Approximation of the weighted completion time via stronger negative correlation

Journal of Scheduling manuscript No.
(will be inserted by the editor)

Approximation of the weighted completion time via stronger negative correlation

Alok Baveja · Steven Qu · Aravind Srinivasan

the date of receipt and acceptance should be inserted later

Abstract Minimizing the weighted completion time of jobs in the unrelated parallel machines model is a fundamental scheduling problem. The first $(3/2 - c)$–approximation algorithm for this problem, for some constant $c > 0$, was obtained in the work of Bansal, Srinivasan, and Svensson (*SIAM J. Computing*, 2021). A key ingredient in this work was the first dependent-rounding algorithm with a certain guaranteed amount of negative correlation. We improve upon this guaranteed amount from 1/108 to 1/27, thus also improving upon the constant c in the algorithms of Bansal *et al.* and Li (*SIAM J. Computing*, 2020) for weighted completion time. Given the now-ubiquitous role played by dependent rounding in scheduling and combinatorial optimization, our improved dependent rounding is also of independent interest.

Keywords Scheduling · Completion Time · Approximation Algorithms · Dependent Rounding

Mathematics Subject Classification (2020) 68Q25 · 68Q87 · 90B35

Declarations:

Funding: Supported in part by NSF award CCF-1749864, as well as by research awards from Adobe, Amazon, and Google.

Conflicts of interest/Competing interests: Not applicable

Availability of data and material: Not applicable

Code availability: Not applicable

1 Introduction

We consider the well-known problem of scheduling jobs on unrelated machines to minimize the sum of the weighted completion times of the jobs. Herein, we are given a set $J = \{1, 2, \ldots, n\}$ of jobs and a set $M = \{1, 2, \ldots, m\}$ of machines; each job j has a weight $w_j \geq 0$ and requires a given processing time of $p_{ij} \geq 0$ if it gets scheduled on machine $i \in M$. The objective is to find a non-preemptive schedule that minimizes the weighted completion time $\sum_{j \in J} w_j C_j$ of the jobs, where C_j denotes the completion time of job j in the final schedule. The first $(3/2 - c)$–approximation algorithm for this problem, for some constant $c > 0$, was obtained in the work of Bansal et al. (2021). A key ingredient in this work

Supported in part by NSF award CCF-1749864, as well as by research awards from Adobe, Amazon, and Google

Alok Baveja
Department of Supply Chain Management, Rutgers Business School, 100 Rockafeller Road, Room 3151, Piscataway, NJ 08854, USA
E-mail: baveja@business.rutgers.edu

Steven Qu
Department of Electrical Engineering and Computer Science, Massachusetts, Cambridge, MA 02139, USA. Part of this work was done while at Montgomery Blair High School, Silver Spring, MD 20901, USA
E-mail: xsqu@mit.edu

Aravind Srinivasan
Department of Computer Science and UMIACS, University of Maryland, College Park, MD 20742, USA
E-mail: asriniv1@umd.edu

was the first dependent-rounding algorithm with a certain *guaranteed amount* of negative correlation. We improve this "measure of negative correlation" from 1/108 to 1/27, thus also improving upon the constant c for weighted completion time in (Bansal et al., 2021; Li, 2020). Given the growing ubiquity of dependent rounding in scheduling (see, e.g., (Bansal et al., 2021; Im and Shadloo, 2020; Kumar et al., 2009; Li, 2020; Li et al., 2016; Saha and Srinivasan, 2018)) and combinatorial optimization, we believe our improved dependent-rounding approach is of independent interest.

Total completion time is a classical problem that goes at least as far back as the work of Smith (1956). Since many natural versions of the problem are NP-hard, one is interested in approximation algorithms: recall that a ρ-approximation algorithm for a minimization problem, where $\rho \geq 1$, is a polynomial-time (randomized) algorithm whose (expected) objective-function value is at most ρ times optimal. The approximability of completion-time objectives has been studied since the work of Phillips et al. (1998). We now have a nearly-complete understanding of the problem's approximability in simpler models: e.g., for identical and related machines (Afrati et al., 1999; Chekuri and Khanna, 2001; Skutella and Woeginger, 1999) (with some of these holding under other generalizations such as release dates for jobs). For our more-general problem with unrelated parallel machines, (3/2)-approximations were obtained independently by obtained by Skutella (2001) and Sethuraman and Squillante (1999). After a while, the first approximations of the form $(3/2 - \Omega(1))$ were obtained by Bansal et al. (2021); Li (2020); Im and Shadloo (2020) through convex-programming approaches; we contribute to this line of research. On the negative side, the problem is APX-hard (Hoogeveen et al., 2001).

1.1 Dependent rounding and its connection with our problem

A common approach to many discrete-optimization problems is to introduce and solve a continuous relaxation, and to then use (probabilistic) techniques to round the resulting values. For many assignment problems, most notably in scheduling, these techniques are applied to bipartite graphs—with one side corresponding to the jobs, and the other to the machines; see, e.g., Shmoys and Tardos (1993). It is often the case that naive *independent* randomized rounding—where we round the fractional values independently—cannot accommodate hard constraints, or is not powerful enough to prove desired bounds. Thus, *dependent rounding*, where we carefully introduce dependencies among various random variables, has seen much use in scheduling: see applications of this, e.g., in (Kumar et al., 2009; Li et al., 2016; Saha and Srinivasan, 2018) for other scheduling models, and, as mentioned above, (Bansal et al., 2021; Li, 2020; Im and Shadloo, 2020) for our problem of weighted completion time. (More generally, dependent rounding has seen several applications in combinatorial optimization; see, e.g., Bansal (2019) for some exciting progress in the area.) Many of these probabilistic approaches are offshoots of the *deterministic* pipage-rounding technique of Ageev and Sviridenko (2004).

In this work, we will consider a special kind of dependent rounding on bipartite graphs: specifically, one that induces *strong* negative correlation between subsets of vertices without introducing positive correlations at pairs of edges with a common endpoint. This is an approach initiated by Bansal et al. (2021). Formally, consider a bipartite graph $G = (U \cup V, E \subseteq U \times V)$—where U is interpreted as the set of machines and V as the set of jobs—with fractional values $y_e \in [0, 1]$ assigned to each $e \in E$. Let $\delta(v)$ be the set of edges incident to vertex v and denote $y(S) = \sum_{e \in S} y_e$ for $S \subseteq E$. Then, the approach of Bansal et al. (2021) yields a randomized polynomial-time algorithm that rounds each y_e to a random variable $Y_e \in \{0, 1\}$ such that the following properties hold (stated informally here, and defined formally in Theorem 1):

Property 1.1 [Job Assignment]: Let E^* be the set of edges e with $Y_e = 1$. Then $|E^* \cap \delta(v)| = 1$ for all $v \in V$, with probability one.

Property 1.2 [Marginal Preservation]: For every $e \in E$, $\mathbb{E}[Y_e] = \Pr[e \in E^*] = y_e$.

Property 1.3 [Weak and Strong Negative Correlation]: For every $u \in U$ and $e \neq e' \in \delta(u)$, $\Pr[e, e' \in E^*] \leq y_e y_{e'}$ ("weak negative correlation"). Additionally, for some constant $\zeta > 0$, we have that $\Pr[e, e' \in E^*] \leq (1 - \zeta) y_e y_{e'}$ if e and e' belong in "certain given subsets"—to be defined precisely in Theorem 1—of $\delta(u)$; we refer to this as "strong negative correlation".

Properties 1.1 and 1.2, along with the weak-negative-correlation bound of Property 1.3, are present in most dependent-randomized rounding schemes, including the works mentioned above. The work of Bansal et al. (2021) proves the strong-negative-correlation bound of Property 1.3 with the constant of $\zeta = 1/108$, using an algorithm that maintains certain invariants. In this work, we will modify their algorithm in order to improve the invariants, and, importantly, the constant ζ to 1/27.

Strong negative correlation and weighted completion time. The work of Bansal et al. (2021) showed that with a rounding scheme such as the above, we can get a $(3/2 - K\zeta)$–approximation for our problem of minimizing the weighted completion time, for some absolute constant $K > 0$. This result, along with the value $\zeta = 1/108$, led to a $(3/2 - 10^{-7})$–approximation in Bansal et al. (2021), the first $(3/2 - \Omega(1))$–approximation as mentioned above; the constant 10^{-7} can be improved somewhat. Li used the strong-negative-correlation result of Bansal et al. (2021) as a black box with a different, simpler, relaxation and other interesting ideas to obtain a $(3/2 - 1/6000)$–approximation (Li, 2020). Our improved value of ζ helps improve upon both of these algorithms. However, by developing a somewhat-different notion of strong negative correlation with new *iterative fair contention-resolution* techniques, Im & Shadloo have developed the current-best approximation for our problem, which is 1.488 (Im and Shadloo, 2020). Specifically, instead of the "$\Pr[e, e' \in E^*] \le (1 - \zeta) y_e y_{e'}$" in Property 1.3 above, the work of Im and Shadloo (2020) obtains (and uses) the upper bound

$$\left(\frac{e^{y_e} + e^{y'_e}}{1 + e} \right) \cdot y_e y_{e'}$$

as the right-hand side; note that by a slight overload of notation, e here denotes the base of the natural logarithm as well as edges (as in e and e'). Thus, two closely-related notions of strong negative correlation have been key in developing $(3/2 - \Omega(1))$–approximations for our problem. We focus here on improving the value of ζ from 1/108 to 1/27, via a modification of the rounding algorithm of Bansal et al. (2021). This makes progress on a question due to Singh (2016) on what the optimal value of ζ can be. While our algorithm follows from a small added ingredient of randomness to the algorithm of Bansal et al. (2021), the analysis to obtain our improved ζ is significantly more involved. We also remark that the work of Im and Shadloo (2020) does not yield an improvement such as ours on ζ: thus, there could conceivably be future applications where our result is applicable while that of Im and Shadloo (2020) is not.

2 Algorithm

We will first formally state the theorem from Section 1, and then describe the modified version of the algorithm developed in Bansal et al. (2021) which proves the theorem. Along with several steps of the algorithm, we will also make observations key to the analysis. As mentioned above, our main contribution here is that we improve the value of the negative-correlation constant ζ of Bansal et al. (2021) from 1/128 to 1/27. We describe the main changes from the algorithm and analysis of Bansal et al. (2021) in Section 2.1.

Theorem 1 ***(From Bansal et al. (2021), but with our improved*** ζ***.)*** *Let* $\zeta = 1/27$*. Consider a bipartite graph* $G = (U \cup V, E)$ *and let* $y \in [0,1]^E$ *be fractional values on the edges satisfying* $y(\delta(v)) = 1$ *for all* $v \in V$*. For each vertex* $u \in U$*, select any family of disjoint subsets of edges incident to* u*,* $E_u^{(1)}, E_u^{(2)}, \ldots, E_u^{(\kappa_u)} \subseteq \delta(u)$*, such that* $y(E_u^{(\ell)}) \le 1$ *for* $\ell = 1, 2, \ldots, \kappa_u$*. Then there exists a randomized polynomial-time algorithm that outputs a random subset of the edges* $E^* \subseteq E$ *satisfying*

(a) For every $v \in V$*, we have* $|E^* \cap \delta(v)| = 1$ *with probability* 1*;*
(b) For every $e \in E$*,* $\Pr[e \in E^*] = y_e$*;*
(c) For every $w \in U$ *and all* $e \neq e' \in \delta(w)$*,*

$$\Pr[e \in E^* \wedge e' \in E^*] \le \begin{cases} (1 - \zeta) \cdot y_e y_{e'} & \text{if } e, e' \in E_w^{(\ell)} \text{ for some } \ell \in \{1, 2, \ldots, \kappa_w\}, \\ y_e y_{e'} & \text{otherwise.} \end{cases}$$

Remark. We typically reserve symbols such as u to denote elements of U; however, we use "w" in part (c) of Theorem 1 since this part (c) is heavily connected to the situation depicted in Figure 1, wherein "u" has a special meaning.

Notation. A value $z \in [0, 1]$ will be called "floating" if $z \in (0, 1)$.

We divide the algorithm into three phases and present each phase along with some observations that will be useful in Sections 3 and 4.

Phase 1 (Forming the collection of edges R^***).** Let y^* denote the initial fractional assignment, and let s be a fixed integer that we will optimize over later. For each $v \in V$, partition $\delta(v)$ into at most s disjoint groups by letting each group—except possibly for at most one group—be an *inclusion-wise minimal* set of edges with y-value at least $1/s$. (Note that this is possible because $y(\delta(v)) = 1$,

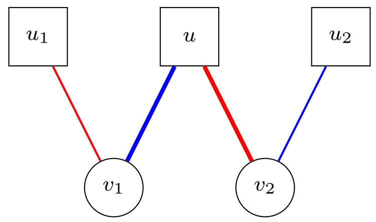

Fig. 1: Illustration of the update step in Phase 2. Either both red edges are increased by α and both blue edges are decreased by α, or vice versa. The thick edges are in R and the thin edges are not. With probability at least 1/2, at least one of these four edges will get an integral y-value after the probabilistic update of Phase 2.

and that the following simple greedy algorithm will construct the partition. Let $S(v)$ be the sequence $(y_e : e \in \delta(v))$ ordered in some fixed way; take the first block of the partition to be the edges in the first prefix of $S(v)$ that adds up to at least $1/s$, and then iterate on the remaining suffix of $S(v)$ to construct the further blocks of the partition.) We then select a random group, uniformly at random and independently for each vertex v, and let R be the set of selected edges.

Observation 2.1. Let $e, e' \in \delta(u)$ for some $u \in U$. Then, $\Pr[e, e' \in R^*] \geq 1/s^2$.

Proof The events that $e \in R^*$ and that $e' \in R^*$ are independent, as they are incident to different vertices in V. The statement then follows as for each $v \in V$ we select a random group out of at most s many. □

We can vary s as we please in order to optimize the value of ζ. The result in Bansal et al. (2021) uses $s = 6$, but we will use $s = 3$ here. The exact details will be elaborated on in Section 3.

Phase 2 (Updating the assignment). Initially let $y = y^*, R = R^*$. Repeat the following steps while there exist edges $\{u, v_1\}, \{u, v_2\} \in R \cap E_u^{(l)}$ for some l and $\{u_1, v_1\} \in \delta(v_1) \setminus R$ and $\{u_2, v_2\} \in \delta(v_2) \setminus R$, with all four of these edges having floating y-values. Here $u, u_1, u_2 \in U, v_1, v_2 \in V$, but are otherwise arbitrary. See Figure 1.

1. Let α be the following minimum of eight quantities:

$$\alpha = \min\{y_{u_1,v_1},\ 1 - y_{u_1,v_1},\ y_{u,v_1},\ 1 - y_{u,v_1},\ y_{u,v_2},\ 1 - y_{u,v_2},\ y_{u_2,v_2},\ 1 - y_{u_2,v_2}\};$$

   note that $\alpha > 0$ due to our "all four of these edges have floating y-values" assumption, and that $\alpha \leq 1/2$ since $\alpha \leq \min\{y_{u_1,v_1},\ 1 - y_{u_1,v_1}\}$.
2. With probability 1/2, update y as follows for each $e \in E$:

$$y_e = \begin{cases} y_e + \alpha & \text{if } e = \{u_1, v_1\} \text{ or } e = \{u, v_2\}, \\ y_e - \alpha & \text{if } e = \{u, v_1\} \text{ or } e = \{u_2, v_2\}, \\ y_e & \text{otherwise.} \end{cases}$$

   Otherwise, with the remaining probability 1/2, update y as follows for each $e \in E$:

$$y_e = \begin{cases} y_e - \alpha & \text{if } e = \{u_1, v_1\} \text{ or } e = \{u, v_2\}, \\ y_e + \alpha & \text{if } e = \{u, v_1\} \text{ or } e = \{u_2, v_2\}, \\ y_e & \text{otherwise.} \end{cases}$$

   (We remark that this step, where we either add or subtract the *same* value α probabilistically, is different from that of Bansal et al. (2021), where different values α and β can be added/subtracted. Our approach can double the expected run-time, but makes the analysis simpler since only one parameter α is involved.)
3. For $v \in \{v_1, v_2\}$, if $y(\delta(v) \cap R) = 1$, then update R as

$$R = (R \setminus \delta(v)) \cup \{e\},$$

   where the single edge $e \in \delta(v)$ is selected with probability y_e. As with the randomized steps above, the random choice here is made independently of all random choices made thus far by the algorithm.

Our **key modification** in Phase 2 is in the third step: in contrast, the algorithm in Bansal et al. (2021) removes all but the edge e with the largest y-value. Our modification, as we will see in Section 4, makes for our improved ζ. Because α depends on the values of all four edges in consideration, there is a probability of at least 1/2 that a single iteration of Phase 2 will have at least one edge with floating y-value reach an integral y-value. Furthermore, these events are independent from each other. We can thus observe the following:

Observation 2.2. With high probability, Phase 2 will terminate within $c \cdot 2|E|$ iterations, for any given constant $c > 1$.

Proof Note that a single iteration of Phase 2 yields a probability of at least 1/2 that at least one edge with floating y-value will reach an integral y-value in that iteration. Thus, since there are $|E|$ edges in total, the random number N of iterations of Phase 2 is upper-bounded (i.e., stochastically dominated) by the time it takes for a sequence of independent tosses of a fair coin to output "Heads" at least $|E|$ times. That is, we have a Negative Binomial distribution with parameter 1/2 from which we ask of $|E|$ "Heads". It is immediate that the expected value of N is at most $2|E|$; for a concentration bound on N, note by the Chernoff bound (Chernoff, 1952) that the probability that $2c|E|$ independent tosses of a fair coin yield less than $|E|$ Heads, is exponentially small in $|E|$ for any given constant $c > 1$. □

Observation 2.3. Phase 2 satisfies the invariants $y(\delta(v)) = 1$ and $y_e \geq 0$ for every $v \in V, e \in E$.

Proof Notice that the selection of α in Phase 2 guarantees $y_e \geq 0$ at all times, and keeps $y(\delta(v))$ constant. Thus the statement follows as $y(\delta(v)) = 1$ at the start. □

Observation 2.4. The set R does not get any new elements during Phase 2.

Proof This follows directly from Step 3 of Phase 2. □

Observation 2.5. When Phase 2 terminates, we have for all $u \in U$ and $l \in \{1, 2, \ldots, \kappa_u\}$ that $|\{e \in E_u^{(l)} \cap R : \ y_e > 0\}| \leq 1$.

Proof Suppose that there exist $e_1, e_2 \in E_u^{(l)} \cap R$ with $y_{e_1}, y_{e_2} > 0$. Then, the fact that any iteration of Phase 2 maintains the value of $y(E_u^{(l)} \cap R)$, along with Observation 2.4, shows that $y(E_u^{(l)} \cap R) \leq y^*(E_u^{(l)} \cap R^*) \leq 1$, implying that y_{e_1}, y_{e_2} are floating. Now let $e_1 = \{v_1, u\}, e_2 = \{v_2, u\}$. Then, Step 3 of Phase 2 guarantees that $y(\delta(v_1) \cap R)$, $y(\delta(v_2) \cap R) < 1$. Thus, there exist edges $\{v_1, u_1\} \in \delta(v_1) \setminus R$ and $\{v_2, u_2\} \in \delta(v_2) \setminus R$ with floating y-values, implying that Phase 2 has not terminated yet, which is a contradiction. □

Phase 3 (Randomized Rounding). (Note that Phase 2 has a "repeat while" loop. Phase 3 is run after Phase 2 terminates.) Construct E^* by, independently for each vertex $v \in V$, selecting a single edge $e \in \delta(v)$ so that an edge e is selected with probability y_e. This is possible because $y(\delta(v)) = 1$ and $y_e \geq 0$ for all $v \in V$ and for all e.

2.1 Main changes from Bansal et al. (2021)

As mentioned above, our key modification to Bansal et al. (2021) is in Step 3 of Phase 2: in contrast to our randomized approach, the algorithm of Bansal et al. (2021) removes all but the edge e with the largest y-value. This modification makes for our improved ζ. A less-important modification—to optimize the value of ζ—is that we use $s = 3$ here, as opposed to the choice $s = 6$ in Bansal et al. (2021).

In terms of the analysis, our main improvement is in Invariant 3 of Section 3: the corresponding invariant in Bansal et al. (2021) has a further factor of 2 in the right-hand side of Invariant 3, which we are able to avoid due to our modified Step 3 of Phase 2.

3 Analysis

We first note that the algorithm terminates in (random) polynomial time. Phase 1 and Phase 3 both clearly run in polynomial time, and Phase 2 does as well by Observation 2.2.

Now, we will prove the properties stated in Theorem 1. First, property (a) holds from Observation 2.3 and the mechanics of Phase 3. To show properties (b) and (c), we will inductively show some invariants.

Let $Y^{(k)} = (y_e^{(k)} : e \in E)$ denote the collection of y-values of edges and let $R^{(k)}$ be the set R at the end of iteration k of Phase 2. For an edge $e = \{u, v\} \in R$ with $u \in U, v \in V$, let $R_{\overline{e}} = \{e' \in \delta(v) \cap R : e' \neq e\}$ be the other edges in R incident to v.

We show that the following invariants hold after each iteration k, where conditioning an event on $Y^{(k)}$ and $R^{(k)}$ means the probability of that event if the iterations in Phase 2 are applied starting from the assignments of $Y^{(k)}$ and $R^{(k)}$.

Invariant 1. $\Pr[e \in E^* \mid Y^{(k)}, R^{(k)}] = y_e^{(k)}$ for all $e \in E$.

Invariant 2. $\Pr[e \in E^* \wedge e' \in E^* \mid Y^{(k)}, R^{(k)}] \leq y_e^{(k)} y_{e'}^{(k)}$ for all $w \in U$ and all distinct $e, e' \in \delta(w)$.

Invariant 3. $\Pr[e \in E^* \wedge e' \in E^* \mid Y^{(k)}, R^{(k)}] \leq \left(y^{(k)}(R_{\overline{e}}^{(k)}) + y^{(k)}(R_{\overline{e'}}^{(k)})\right) y_e^{(k)} y_{e'}^{(k)}$ for all $w \in U, l \in \{1, \ldots, \kappa_w\}$, and all distinct $e, e' \in \left(E_w^{(l)} \cap R^{(k)}\right)$.

Remark. As mentioned in Section 2.1, our *key improvement* is in Invariant 3: the corresponding invariant in Bansal et al. (2021) has a further factor of 2 in the right-hand side of Invariant 3, which we are able to avoid due to our modified Phase 2.

We defer the proof of these invariants to Section 4. Let us first show properties (b) and (c) of Theorem 1, assuming these invariants.

By definition of Phase 2 we have $y^{(0)} = y^*$ and $R^{(0)} = R^*$. Thus, applying Invariant 1 to $k = 0$ yields property (b) and applying Invariant 2 to $k = 0$ yields the weak bound in property (c). For the stronger bound, consider two edges $e \neq e' \in E_w^{(l)}$. Then by Invariant 3, we have that

$$
\begin{aligned}
\Pr[e \in E^* \wedge e' \in E^*] &= \mathbb{E}_{R^*}[\Pr[(e \in E^* \wedge e' \in E^*) \mid Y^*, R^*]] \\
&\leq \Pr[e, e' \in R^*] \cdot \left(y^*(R_{\overline{e}}^*) + y^*(R_{\overline{e'}}^*)\right) y_e^* y_{e'}^* + (1 - \Pr[e, e' \in R^*]) \cdot y_e^* y_{e'}^* \\
&\leq \Pr[e, e' \in R^*] \cdot \frac{2 y_e^* y_{e'}^*}{3} + (1 - \Pr[e, e' \in R^*]) \cdot y_e^* y_{e'}^* \\
&= y_e^* y_{e'}^* \cdot (1 - (1/3) \cdot \Pr[e, e' \in R^*]) \\
&\leq \frac{26}{27} y_e^* y_{e'}^*;
\end{aligned}
$$

the second inequality follows from the fact that $y^*(R_{\overline{e}}^*), y^*(R_{\overline{e'}}^*) \leq 1/3$, which is because in Phase 1 we chose $s = 3$. The third inequality follows from Observation 2.1.

4 Induction

We will prove the three invariants of Section 3 using reverse induction on the iterations of Phase 2, as in Bansal et al. (2021). This will then conclude the proof.

Base Case (when Phase 2 terminates): In this case Phase 2 will not change any of the y-values. Then by the edge selection of Phase 3, we have that $\Pr[e \in E^*] = y_e$, so Invariant 1 holds. Similarly for Invariant 2, we note that for two edges $e \neq e' \in \delta(w)$, it holds that $\Pr[e \in E^* \wedge e' \in E^*] = y_e y_{e'}$, as desired.

Finally, Observation 2.5 says that the number of edges in $E_w^{(l)} \cap R$ with positive value is at most 1, so for $e \neq e' \in E_w^{(l)} \cap R$, we must have that $\Pr[e \in E^* \wedge e' \in E^*] = 0$ and Invariant 3 holds.

For the inductive step, we will assume that our invariants hold at the end of iteration k and prove that they hold at the end of iteration $k - 1$. Let us denote $Y = Y^{(k-1)}$ and $R = R^{(k-1)}$; also let $Y' = Y^k, R' = R^k$ respectively denote the updated y-values and set R at the end of iteration k. Note that for a given Y and R, Y' and R' are random variables.

Inductive step for Invariants 1 and 2: For Invariant 1, notice that $\Pr[e \in E^* \mid Y] = \mathbb{E}_{Y'|Y}[\Pr[e \in E^* \mid Y']] = \mathbb{E}_{Y'|Y}[y'_e]$ by the inductive hypothesis. But if Phase 2 did not update the value of edge e, then $y'_e = y_e$. Otherwise, $\mathbb{E}_{Y'|Y}[y'_e] = \frac{1}{2}(y_e + \alpha) + \frac{1}{2}(y_e - \alpha) = y_e$, as desired.

We proceed similarly for Invariant 2. By the inductive hypothesis, we have

$$
\Pr[e \in E^* \wedge e' \in E^* \mid Y] = \mathbb{E}_{Y'|Y}[\Pr[e \in E^* \wedge e' \in E^* \mid Y']] \leq \mathbb{E}_{Y'|Y}[y'_e y'_{e'}].
$$

If Phase 2 changes the y-value of at most one of e and e', then by independence the right-hand side is at most $\mathbb{E}_{Y'|Y}[y_e]\mathbb{E}_{Y'|Y}[y'_e] = y_e y_{e'}$. On the other hand, if Phase 2 changed both values, then we have

$$
\mathbb{E}_{Y'|Y}[y'_e y'_{e'}] = \frac{1}{2}(y_e + \alpha)(y_{e'} - \alpha) + \frac{1}{2}(y_e - \alpha)(y_{e'} + \alpha) \leq y_e y_{e'}.
$$

This completes the proof for Invariant 2.

Inductive step for Invariant 3: The proof for Invariant 3 is more tricky because the set R may change after an iteration. Fix $w \in U$, $l \in \{1, \ldots, \kappa_w\}$, and $e \neq e' \in \left(E_w^{(l)} \cap R^{(k-1)}\right)$. Call an edge $\{x_0, x_1\}$ *dangerous*, where $x_0 \in U$ and $x_1 \in V$, if it is possible that $y(\delta(x_1) \cap R) = 1$ *just after* the edge values are updated in iteration k (i.e., just after Step 2). Let $e_1 = \{u, v_1\}$ and $e_2 = \{u, v_2\}$ along with $\{u_1, v_1\}$ and $\{u_2, v_2\}$ be the edges being updated in iteration k, as in Figure 1. let $\alpha > 0$ be the update parameter as in our algorithm. We will then split our proof into cases based on how many among e and e' lie in $\{e_1, e_2\}$, and then into sub-cases based, e.g., on the number of dangerous edges among e and e'. *Note in particular that since e and e' lie in $R = R^{(k-1)}$, they cannot be either of the edges (u_1, v_1) and (u_2, v_2) in Figure 1, since these latter edges do not lie in R.*

Case 1: $e, e' \notin \{e_1, e_2\}$. We consider sub-cases here, based on how many of e and e' (if any) have v_1 or v_2 as an end-point.

Case 1(a): Neither e nor e' has v_1 or v_2 as an end-point. In this case, none of y_e, $y_{e'}$, $R_{\overline{e}}$, $R_{\overline{e'}}$, $y(R_{\overline{e}})$, and $y(R_{\overline{e'}})$ changes in iteration k. Therefore,

$$\Pr[e \in E^* \wedge e' \in E^* \mid Y, R] = \Pr[e \in E^* \wedge e' \in E^* \mid Y', R'] \leq (y(R_{\overline{e}}) + y(R_{\overline{e'}}))\, y_e y_{e'},$$

where the inequality follows from the induction hypothesis.

Case 1(b): Exactly one of e and e' has v_1 or v_2 as an end-point. Suppose (without loss of generality) that this edge is e, and suppose e has v_1 as an end-point. We now consider two further sub-cases based on whether e is dangerous or not.

Case 1(b)(i): e is not dangerous. In this case, two outcomes are possible in iteration k: with probability 1/2, $y(R_{\overline{e}})$ is incremented by α, and with the remaining probability of 1/2, $y(R_{\overline{e}})$ is decremented by α. All other parameters relevant to us remain unchanged. Thus,

$$\begin{aligned}
\Pr[e \in E^* \wedge e' \in E^* \mid Y, R] &= \mathbb{E}_{Y',R'|Y,R}[\Pr[e \in E^* \wedge e' \in E^* \mid Y', R']] \\
&\leq \mathbb{E}_{Y',R'|Y,R}\left[(y'(R'_{\overline{e}}) + y'(R'_{\overline{e'}}))\, y'_e y'_{e'}\right] \quad \text{(by the inductive hypothesis)} \\
&= \left(\mathbb{E}_{Y',R'|Y,R}\left[y'(R'_{\overline{e}})\right] + y(R_{\overline{e'}})\right) y_e y_{e'} \\
&= (y(R_{\overline{e}}) + y(R_{\overline{e'}}))\, y_e y_{e'}
\end{aligned}$$

as required.

Case 1(b)(ii): e is dangerous. This case needs more work. Note that since e is dangerous,

$$y(R_{\overline{e}}) = 1 - y_e - \alpha.$$

There are three possible relevant outcomes in iteration k here:

(O1) with probability 1/2, $y(e_1)$ gets decremented by α, in which case the only relevant change for us is that $y(R_{\overline{e}})$ is also decremented by α.

(O2) with probability $(1/2) \cdot y_e$, $y(e_1)$ gets incremented by α, and in the random pruning of R in step 3, e is chosen to be kept in R'. Here, e stays in R', but $y'(R'_{\overline{e}})$ becomes zero since $R'_{\overline{e}}$ is the empty set.

(O3) with probability $(1/2) \cdot (1 - y_e)$, $y(e_1)$ gets incremented by α, and due to the random pruning of R in step 3, e is not present in R'; in this case, we can use Invariant 2.

Thus, by the inductive hypothesis (Invariant 3 for outcomes (O1) and (O2), and Invariant 2 for (O3)), we obtain

$$\begin{aligned}
\Pr[e \in E^* \wedge e' \in E^* \mid Y, R] &\leq \frac{1}{2} \cdot (y(R_{\overline{e}}) - \alpha + y(R_{\overline{e'}})) \cdot y_e y_{e'} + \\
&\quad \frac{y_e}{2} \cdot (0 + y(R_{\overline{e'}})) \cdot y_e y_{e'} + \\
&\quad \frac{(1 - y_e)}{2} \cdot y_e y_{e'} \\
&= \frac{y_e y_{e'}}{2} \cdot (y(R_{\overline{e}}) - \alpha + y(R_{\overline{e'}}) + y(R_{\overline{e'}}) y_e + 1 - y_e) \\
&\leq \frac{y_e y_{e'}}{2} \cdot (y(R_{\overline{e}}) - \alpha + y(R_{\overline{e'}}) + y(R_{\overline{e'}}) + 1 - y_e) \\
&= y_e y_{e'} \cdot (y(R_{\overline{e}}) + y(R_{\overline{e'}}))
\end{aligned}$$

as desired, since $y(R_{\overline{e}}) = 1 - y_e - \alpha$.

Case 1(c): Each of e and e' has an end-point in $\{v_1, v_2\}$. We may assume that e has v_1 as an end-point and that e' has v_2 as an end-point. (Note that since e and e' already share the end-point w, their respective other end-points have to differ.) We consider three sub-cases, based on how many of e and e' are dangerous.

Case 1(c)(i): neither of e and e' is dangerous. Here, the only relevant changes for us are:

- with probability 1/2, $y(R_{\overline{e}})$ gets incremented by α and $y(R_{\overline{e'}})$ gets decremented by α; and
- with the remaining probability of 1/2, $y(R_{\overline{e}})$ gets decremented by α and $y(R_{\overline{e'}})$ gets incremented by α.

Either way, the sum $y(R_{\overline{e}}) + y(R_{\overline{e'}})$, as well as y_e and $y_{e'}$, remain unchanged, so we are done by Invariant 3 of the inductive hypothesis.

Case 1(c)(ii): exactly of e and e' is dangerous. Suppose without loss of generality that e is dangerous; so, $y(R_{\overline{e}}) = 1 - y_e - \alpha$. As in Case 1(b)(ii), there are three possible outcomes in iteration k, which happen with the respective probabilities of 1/2, $(1/2) \cdot y_e$, and $(1/2) \cdot (1 - y_e)$:

- $y(R_{\overline{e}})$ gets decremented by α and $y(R_{\overline{e'}})$ gets incremented by α;
- $y(R_{\overline{e'}})$ gets decremented by α, $e \in R'$, and $R'_{\overline{e}}$ becomes empty (hence $y'(R'_{\overline{e}}) = 0$); and
- $y(R_{\overline{e'}})$ gets decremented by α, and $e \notin R'$.

As in Case 1(b)(ii), we apply Invariant 3 to the first two of these outcomes and Invariant 2 to the third outcome, to obtain inductively that $\Pr[e \in E^* \wedge e' \in E^* \mid Y, R]$ is at most

$$
\begin{aligned}
& y_e y_{e'} \cdot [(1/2) \cdot (y(R_{\overline{e}}) - \alpha + y(R_{\overline{e'}}) + \alpha) + (y_e/2) \cdot (y(R_{\overline{e'}}) - \alpha) + ((1 - y_e)/2) \cdot 1] \\
&= \frac{y_e y_{e'}}{2} \cdot [y(R_{\overline{e}}) + y(R_{\overline{e'}}) + 1 - y_e + y_e \cdot (y(R_{\overline{e'}}) - \alpha)] \\
&\le \frac{y_e y_{e'}}{2} \cdot [y(R_{\overline{e}}) + y(R_{\overline{e'}}) + 1 - y_e + y(R_{\overline{e'}}) - \alpha] \\
&= y_e y_{e'} \cdot [y(R_{\overline{e}}) + y(R_{\overline{e'}})] ,
\end{aligned}
$$

and hence the inductive step for Case 1(c)(ii) is complete.

Case 1(c)(iii): both e and e' are dangerous. There are four possible relevant outcomes now:

(O1') with probability $(1/2) \cdot y_e$, y_{e_1} gets incremented by α, y_{e_2} gets decremented by α, and e is selected to be in R'. Here, $R'_{\overline{e}} = \emptyset$, and $y(R_{\overline{e'}})$ is decremented by α.

(O2') with probability $(1/2) \cdot (1 - y_e)$, y_{e_1} gets incremented by α, y_{e_2} gets decremented by α, and e is *not* selected to be in R'.

(O3') with probability $(1/2) \cdot y_{e'}$, y_{e_2} gets incremented by α, y_{e_1} gets decremented by α, and e' is selected to be in R'. Here, $R'_{\overline{e'}} = \emptyset$, and $y(R_{\overline{e}})$ is decremented by α.

(O4') with probability $(1/2) \cdot (1 - y_{e'})$, y_{e_2} gets incremented by α, y_{e_1} gets decremented by α, and e' is *not* selected to be in R'.

Thus by inductively applying Invariant 3 to outcomes (O1') and (O3'), and Invariant 2 to outcomes (O2') and (O4'), we get that $\Pr[e \in E^* \wedge e' \in E^* \mid Y, R]$ is at most

$$
\begin{aligned}
& (1/2) \cdot y_e \cdot (0 + y(R_{\overline{e'}}) - \alpha) y_e y_{e'} + \\
& (1/2) \cdot (1 - y_e) \cdot y_e y_{e'} + \\
& (1/2) \cdot y_{e'} \cdot (y(R_{\overline{e}}) - \alpha + 0) y_e y_{e'} + \\
& (1/2) \cdot (1 - y_{e'}) \cdot y_e y_{e'} \\
&= \frac{y_e y_{e'}}{2} \cdot [y_e y(R_{\overline{e'}}) + y_{e'} y(R_{\overline{e}}) - \alpha(y_e + y_{e'}) + 2 - (y_e + y_{e'})] . \qquad (1)
\end{aligned}
$$

The equalities $y(R_{\overline{e}}) = 1 - y_e - \alpha$ and $y(R_{\overline{e'}}) = 1 - y_{e'} - \alpha$ help us simplify (1) and to arrive at

$$
\Pr[e \in E^* \wedge e' \in E^* \mid Y, R] \le y_e y_{e'} \cdot [1 - y_e y_{e'} - \alpha(y_e + y_{e'})] . \qquad (2)
$$

This thus leads us to *what we want to show* (WTS), in order to complete the induction:

$$
\begin{aligned}
1 - y_e y_{e'} - \alpha(y_e + y_{e'}) &\le y(R_{\overline{e}}) + y(R_{\overline{e'}}), \quad \text{i.e.,} \\
1 - y_e y_{e'} - \alpha(y_e + y_{e'}) &\le (1 - y_e - \alpha) + (1 - y_{e'} - \alpha), \quad \text{i.e.,} \\
(1 - y_e)(1 - y_{e'}) + \alpha(y_e + y_{e'} - 2) &\ge 0. \qquad (3)
\end{aligned}
$$

We now establish (3). Suppose without loss of generality that

$$y_e \geq y_{e'}. \tag{4}$$

Note from (O3') that $y(R_{\overline{e}}) - \alpha \geq 0$, i.e., $1 - y_e - \alpha - \alpha \geq 0$, which implies that $\alpha \leq \frac{1-y_e}{2}$. Thus, since the multiplier "$(y_e + y_{e'} - 2)$" of α in (3) is non-positive and the rest of (3) does not involve α, it suffices to verify (3) for $\alpha = \frac{1-y_e}{2}$. So, our WTS (3) becomes

$$\begin{aligned}(1-y_e)(1-y_{e'}) + \frac{(1-y_e)}{2}\cdot(y_e + y_{e'} - 2) &\geq 0, \quad \text{i.e.,}\\ 2(1-y_{e'}) + y_e + y_{e'} - 2 &\geq 0, \quad \text{i.e.,}\\ y_e &\geq y_{e'},\end{aligned}$$

which is true by our hypothesis (4). Therefore, the proof of the inductive step for Case 1(c)(iii) is complete.

Case 2: Exactly one of e and e' lies in $\{e_1, e_2\}$. We assume without loss of generality that $e = e_1$, and hence that $e' \neq e_2$. We first argue that we may assume that e' does *not* have an end-point in $\{v_1, v_2\}$: if this was indeed the case, then e' would have been a parallel edge with e_1 or e_2, but we have a simple graph and this is not possible. Thus we will assume that e' does not have an end-point in $\{v_1, v_2\}$. We consider two cases, based on whether e is dangerous or not.

Case 2(a): $e = e_1$ is *not* dangerous, and e' does not have an end-point in $\{v_1, v_2\}$. The two relevant possibilities here are:

- with probability 1/2, y_e gets decremented by α, while $y(R_{\overline{e}})$, $y(R_{\overline{e'}})$, and $y_{e'}$ remain unchanged.
- with the remaining probability of 1/2, y_e gets incremented by α, while $y(R_{\overline{e}})$, $y(R_{\overline{e'}})$, and $y_{e'}$ remain unchanged.

Thus, by Invariant 3 applied inductively, we get

$$\begin{aligned}\Pr[e \in E^* \wedge e' \in E^* \mid Y, R] &\leq y_{e'}\left(y(R_{\overline{e}}) + y(R_{\overline{e'}})\right) \cdot \mathbb{E}_{Y',R'|Y,R}\left[y'_e\right] \\ &= y_e y_{e'}\left(y(R_{\overline{e}}) + y(R_{\overline{e'}})\right)\end{aligned} \tag{5}$$

as needed.

Case 2(b): $e = e_1$ *is* dangerous, and e' does not have an end-point in $\{v_1, v_2\}$. We have $y(R_{\overline{e}}) = 1 - y_e - \alpha$ here again. The three relevant outcomes are:

(C1) [Happens with probability 1/2:] y_e gets decremented by α;
(C2) [Happens with probability $(1/2)\cdot(y_e+\alpha)$):] y_e gets incremented by α, e remains in R', and $R'_{\overline{e}}$ becomes empty.
(C3) [Happens with probability $(1/2)\cdot(1-y_e-\alpha)$):] y_e gets incremented by α, and e does not lie in R'.

Thus, by applying Invariant 3 inductively to cases (C1) and (C2) and Invariant 2 to (C3), we see that $\Pr[e \in E^* \wedge e' \in E^* \mid Y, R]$ is at most the sum of the following three terms, each of which corresponds to the respective outcome above (i.e., term (T1) corresponds to case (C1), etc.):

(T1) $(1/2)\cdot(y(R_{\overline{e}}) + y(R_{\overline{e'}}))\cdot(y_e - \alpha)\cdot y_{e'}$;
(T2) $(1/2)\cdot(y_e+\alpha)\cdot y(R_{\overline{e'}})\cdot(y_e+\alpha)\cdot y_{e'}$; and
(T3) $(1/2)\cdot(1-y_e-\alpha)\cdot(y_e+\alpha)\cdot y_{e'}$.

Note from cases (C2) and (C3) that $y_e + \alpha \leq 1$; we of course also have $y_e \geq 0$ and $\alpha \geq 0$. Thus, using

$$(y_e+\alpha)^2 \leq y_e + \alpha \tag{6}$$

and upper-bounding (T2) by $(1/2)\cdot(y_e+\alpha)\cdot y(R_{\overline{e'}})\cdot y_{e'}$ and adding (T1) and (T3) to this upper-bound, we get that

$$\Pr[e \in E^* \wedge e' \in E^* \mid Y, R] \leq \left(y(R_{\overline{e}}) + y(R_{\overline{e'}})\right) y_e y_{e'}.$$

Case 3: $\{e, e'\} = \{e_1, e_2\}$. We will assume without loss of generality that $e = e_1$ and $e' = e_2$.

Case 3(a): Neither e_1 nor e_2 is dangerous. As in Case 2(a), neither of $y(R_{\overline{e}})$ and $y(R_{\overline{e'}})$ changes here. Thus Invariant 3 yields, analogously to (5),

$$\begin{aligned}\Pr[e \in E^* \wedge e' \in E^* \mid Y, R] &\leq \left(y(R_{\overline{e}}) + y(R_{\overline{e'}})\right)\cdot \mathbb{E}_{Y',R'|Y,R}\left[y'_e y'_{e'}\right] \\ &= \left(y(R_{\overline{e}}) + y(R_{\overline{e'}})\right)\cdot\left((1/2)(y_e+\alpha)(y_{e'}-\alpha) + (1/2)(y_e-\alpha)(y_{e'}+\alpha)\right) \\ &\leq \left(y(R_{\overline{e}}) + y(R_{\overline{e'}})\right)\cdot y_e y_{e'}.\end{aligned}$$

Case 3(b): Exactly one of e_1 and e_2 is dangerous. Without loss of generality suppose $e = e_1$ is dangerous and that $e' = e_2$ is not. Then it must be the case that $y(R_{\overline{e}}) = 1 - y_e - \alpha$.

The three possible outcomes here are:

- with probability 1/2, we have $y'(e) = y(e) - \alpha$ and $y'(e') = y(e') + \alpha$;
- with probability $(1/2) \cdot (y_e + \alpha)$, we have $y'(e) = y(e) + \alpha$ as well as $y'(e') = y(e') - \alpha$, and we choose to keep e in R';
- with probability $(1/2) \cdot (1 - y_e - \alpha)$, we have $y'(e) = y(e) + \alpha$ as well as $y'(e') = y(e') - \alpha$, and e is not in R'.

Thus, inductively applying Invariant 3 to the first two outcomes and Invariant 2 to the third, we obtain

$$\begin{aligned}
\Pr[e \in E^* \wedge e' \in E^* \mid Y, R] \leq\ & \frac{1}{2}(1 - y_e - \alpha + y(R_{\overline{e'}}))(y_e - \alpha)(y_{e'} + \alpha) + \\
& \frac{1}{2}(y_e + \alpha)(y(R_{\overline{e'}}))(y_e + \alpha)(y_{e'} - \alpha) + \\
& \frac{1}{2}(1 - y_e - \alpha)(y_e + \alpha)(y_{e'} - \alpha) \\
\leq\ & \frac{1}{2}(1 - y_e - \alpha + y(R_{\overline{e'}}))(y_e - \alpha)(y_{e'} + \alpha) + \\
& \frac{1}{2}(y(R_{\overline{e'}}))(y_e + \alpha)(y_{e'} - \alpha) + \qquad \text{(we used (6) in this line)} \\
& \frac{1}{2}(1 - y_e - \alpha)(y_e + \alpha)(y_{e'} - \alpha) \\
=\ & \frac{1}{2}(1 - y_e - \alpha + y(R_{\overline{e'}})) \cdot [(y_e - \alpha)(y_{e'} + \alpha) + (y_e + \alpha)(y_{e'} - \alpha)] \\
\leq\ & (1 - y_e - \alpha + y(R_{\overline{e'}})) \cdot y_e y_{e'} \\
=\ & (y(R_{\overline{e}}) + y(R_{\overline{e'}})) \cdot y_e y_{e'}
\end{aligned}$$

as desired.

Case 3(c): Both of e_1 and e_2 are dangerous. This is our most involved case. As before, it must be the case that $y(R_{\overline{e}}) = 1 - y_e - \alpha$ and $y(R_{\overline{e'}}) = 1 - y_{e'} - \alpha$. Transitioning from Y, R to Y', R' yields one of four possible cases now:

- with probability $\frac{1}{2} \cdot (y_e + \alpha)$, we have $y'_e = y_e + \alpha$, $y'_{e'} = y_{e'} - \alpha$, $e \in R'$, and $R'_{\overline{e}} = \emptyset$.
- with probability $\frac{1}{2} \cdot (1 - y_e - \alpha)$, we have $y'_e = y_e + \alpha$, $y'_{e'} = y_{e'} - \alpha$, and $e \notin R'$.
- with probability $\frac{1}{2} \cdot (y_{e'} + \alpha)$, we have $y'_e = y_e - \alpha$, $y'_{e'} = y_{e'} + \alpha$, $e' \in R'$, and $R'_{\overline{e'}} = \emptyset$.
- with probability $\frac{1}{2} \cdot (1 - y_{e'} - \alpha)$, we have $y'_e = y_e - \alpha$, $y'_{e'} = y_{e'} + \alpha$, and $e' \notin R'$.

As before, we apply Invariant 3 of the inductive hypothesis to the first and third of these cases, and Invariant 2 to the second and fourth. Hence we obtain

$$\begin{aligned}
\Pr[e \in E^* \wedge e' \in E^* \mid Y, R] \leq\ & \frac{1}{2}(y_e + \alpha)(y_e + \alpha)(y_{e'} - \alpha)(1 - y_{e'} - \alpha) \\
& + \frac{1}{2}(1 - y_e - \alpha)(y_e + \alpha)(y_{e'} - \alpha) \\
& + \frac{1}{2}(y_{e'} + \alpha)(y_{e'} + \alpha)(y_e - \alpha)(1 - y_e - \alpha) \\
& + \frac{1}{2}(1 - y_{e'} - \alpha)(y_{e'} + \alpha)(y_e - \alpha).
\end{aligned}$$

We must show that the right-hand side above (the sum of four terms) does not exceed our targeted term $(y(R_{\overline{e}}) + y(R_{\overline{e'}}))\, y_e y_{e'} = (1 - y_e - \alpha + 1 - y_{e'} - \alpha) y_e y_{e'}$. Expanding this right-hand side above yields our WTS ("what we want to show", as mentioned above):

$$\alpha^4 + \alpha^3(y_e + y_{e'}) - \alpha^2 - \alpha y_e y_{e'}(y_e + y_{e'}) - y_e^2 y_{e'}^2 + y_e y_{e'} \leq (2 - y_e - y_{e'} - 2\alpha) y_e y_{e'}. \tag{7}$$

For the sake of notation, let $P = y_e,\ Q = y_{e'}$. We are given that $\alpha \leq \min\{P, Q\}$, and we can take $P \leq Q$ without loss of generality. By construction we have that $\alpha \leq 1 - Q$, which also means that $\alpha \leq 0.5$.

We next fix α as well as Q, and take P as our variable. Moving everything in (7) to the left-hand side yields a quadratic in P with leading term $Q(1 - Q - \alpha) \geq 0$, which is maximized at one of its two endpoints that are $P = Q$ and $P = \alpha$. Thus, it suffices to show that inequality (7) holds at these two endpoints.

In the former case where $P = Q$, the WTS (7) simplifies to

$$\begin{aligned} \alpha^4 - Q^4 - \alpha^2 + Q^2 - 2\alpha Q^3 + 2\alpha^3 Q &\le (2 - 2Q - 2\alpha)Q^2, \quad \text{i.e.,} \\ (Q^2 - \alpha^2)(1 - (Q+\alpha)^2) &\le 2(1 - (Q+\alpha))Q^2, \quad \text{i.e.,} \\ (Q^2 - \alpha^2)(1 + Q + \alpha) &\le 2Q^2, \end{aligned}$$

where we get the last line from $1 - (Q + \alpha) \ge 0$; we are done here because $1 + Q + \alpha \le 2$ (this holds since $\alpha \le 1 - Q$).

In the latter case where $P = \alpha$, the WTS (7) simplifies to

$$2\alpha^3 - \alpha - 2\alpha Q^2 \le Q - Q^2 - 3\alpha Q, \tag{8}$$

where we divided both sides by α because $\alpha = 0$ yields the desired result trivially. Moving everything to the left-hand side, this is a quadratic in Q with leading term $1 - 2\alpha \ge 0$, so again we only have to prove that (8) holds for the endpoints, i.e., for $Q = \alpha$ and $Q = 1 - \alpha$. When $Q = \alpha$, (8) becomes

$$\begin{aligned} 2\alpha^3 - \alpha - 2\alpha^3 &\le \alpha - \alpha^2 - 3\alpha^2, \quad \text{i.e.,} \\ 0 &\le 2\alpha - 4\alpha^2 = 2\alpha(1 - 2\alpha), \end{aligned}$$

which is true because $\alpha \le 0.5$. When $Q = 1 - \alpha$, (8) becomes

$$\begin{aligned} 2\alpha^3 - \alpha - 2\alpha(1-\alpha)^2 &\le (1-\alpha) - (1-\alpha)^2 - 3\alpha(1-\alpha), \quad \text{i.e.,} \\ 2\alpha(2\alpha - 1) - \alpha &\le (1-\alpha)\alpha - 3\alpha(1-\alpha), \quad \text{i.e.,} \\ 0 &\le \alpha + (1 - 2\alpha)2\alpha - 2\alpha(1-\alpha) = \alpha(1 - 2\alpha), \end{aligned}$$

which is also true because $\alpha \le 0.5$. Therefore, our WTS inequality is true and the proof is complete.

5 Conclusion

We have shown how to improve the strong-negative-correlation constant ζ of Bansal et al. (2021) via a modified algorithm. A few natural questions remain. The first is, as asked in (Singh, 2016): what is the best ζ possible? The second is to pinpoint the approximability of our fundamental weighted-completion-time problem: e.g., is an approximation such as 4/3 achievable? Third, the *flow time* is a more-challenging objective in the computational context, as compared to the completion time (Becchetti et al., 2016); a general open question is to expand our understanding of this key metric. Fourth, in a somewhat different, but critical dimension, there is an increasing push for energy-efficient computing (Pruhs, 2019); since scheduling over the cloud, for instance, has expanded significantly, it would be very interesting to investigate improved—as well as modern—models and algorithms for problems such as ours. Finally, it would be fruitful to obtain further applications of strong negative correlation and of related notions in scheduling and combinatorial optimization.

Acknowledgment. We thank Yifan Xu for his very timely LaTeX help. We also thank the referees for their valuable comments and suggestions.

References

Foto N. Afrati, Evripidis Bampis, Chandra Chekuri, David R. Karger, Claire Kenyon, Sanjeev Khanna, Ioannis Milis, Maurice Queyranne, Martin Skutella, Clifford Stein, and Maxim Sviridenko. Approximation schemes for minimizing average weighted completion time with release dates. In *Foundations of Computer Science, FOCS*, pages 32–44, 1999.

Alexander A. Ageev and Maxim Sviridenko. Pipage rounding: A new method of constructing algorithms with proven performance guarantee. *J. Comb. Optim.*, 8(3):307–328, 2004.

Nikhil Bansal. On a generalization of iterated and randomized rounding. In Moses Charikar and Edith Cohen, editors, *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 1125–1135. ACM, 2019.

Nikhil Bansal, Aravind Srinivasan, and Ola Svensson. Lift-and-round to improve weighted completion time on unrelated machines. *SIAM Journal on Computing*, 50(3):STOC16–138–STOC16–159, 2021. doi: 10.1137/16M1099583. Conference version: Proc. ACM Symposium on Theory of Computing, pages 156-167, 2016.

Luca Becchetti, Stefano Leonardi, Alberto Marchetti-Spaccamela, and Kirk Pruhs. Flow time minimization. In *Encyclopedia of Algorithms*, pages 766–768. 2016.

Chandra Chekuri and Sanjeev Khanna. A PTAS for minimizing weighted completion time on uniformly related machines. In Fernando Orejas, Paul G. Spirakis, and Jan van Leeuwen, editors, *Automata, Languages and Programming, 28th International Colloquium, ICALP 2001, Crete, Greece, July 8-12, 2001, Proceedings*, volume 2076 of *Lecture Notes in Computer Science*, pages 848–861. Springer, 2001.

Herman Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathematical Statistics*, 23:493–509, 1952.

Han Hoogeveen, Petra Schuurman, and Gerhard J. Woeginger. Non-approximability results for scheduling problems with minsum criteria. *INFORMS J. Comput.*, 13(2):157–168, 2001.

Sungjin Im and Maryam Shadloo. Weighted completion time minimization for unrelated machines via iterative fair contention resolution. In Shuchi Chawla, editor, *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SODA 2020, Salt Lake City, UT, USA, January 5-8, 2020*, pages 2790–2809. SIAM, 2020.

V. S. Anil Kumar, Madhav V. Marathe, Srinivasan Parthasarathy, and Aravind Srinivasan. A unified approach to scheduling on unrelated parallel machines. *J. ACM*, 56(5):28:1–28:31, 2009.

Shi Li. Scheduling to minimize total weighted completion time via time-indexed linear programming relaxations. *SIAM Journal on Computing*, 49(4), 2020. Published online April 2020.

Yong Li, Depeng Jin, Pan Hui, and Zhu Han. Optimal base station scheduling for device-to-device communication underlaying cellular networks. *IEEE J. Sel. Areas Commun.*, 34(1):27–40, 2016.

Cynthia A. Phillips, Clifford Stein, and Joel Wein. Minimizing average completion time in the presence of release dates. *Math. Program.*, 82:199–223, 1998.

Kirk Pruhs. Green computing algorithmics. In Bernhard Steffen and Gerhard J. Woeginger, editors, *Computing and Software Science - State of the Art and Perspectives*, volume 10000 of *Lecture Notes in Computer Science*, pages 161–183. Springer, 2019.

Barna Saha and Aravind Srinivasan. A new approximation technique for resource-allocation problems. *Random Struct. Algorithms*, 52(4):680–715, 2018.

Jay Sethuraman and Mark S. Squillante. Optimal scheduling of multiclass parallel machines. In *ACM-SIAM Symposium on Discrete Algorithms, SODA*, pages 963–964, 1999.

David B. Shmoys and Éva Tardos. An approximation algorithm for the generalized assignment problem. *Math. Program.*, 62:461–474, 1993.

Mohit Singh. Personal communication through Nikhil Bansal. 2016.

Martin Skutella. Convex quadratic and semidefinite programming relaxations in scheduling. *J. ACM*, 48(2):206–242, 2001.

Martin Skutella and Gerhard J. Woeginger. A PTAS for minimizing the weighted sum of job completion times on parallel machines. In *Symposium on Theory of Computing, STOC*, pages 400–407, 1999.

Wayne E. Smith. Various optimizers for single-stage production. *Naval Research Logistics*, 3:59–66, 1956.