Improved Approximations for Min Sum Vertex Cover and Generalized Min Sum Set Cover

Nikhil Bansal[*]   Jatin Batra[†]   Majid Farhadi[‡]   Prasad Tetali[§]

Abstract

We study the *generalized min sum set cover* (GMSSC) problem, wherein given a collection of hyperedges E with arbitrary covering requirements $\{k_e \in Z^+ : e \in E\}$, the goal is to find an ordering of the vertices to minimize the total cover time of the hyperedges; a hyperedge e is considered covered by the first time when k_e many of its vertices appear in the ordering.

We give a 4.642 approximation algorithm for GMSSC, coming close to the best possible bound of 4, already for the classical special case (with all $k_e = 1$) of *min sum set cover* (MSSC) studied by Feige, Lovász and Tetali [11], and improving upon the previous best known bound of 12.4 due to Im, Sviridenko and van der Zwaan [20]. Our algorithm is based on transforming the LP solution by a suitable kernel and applying randomized rounding. This also gives an LP-based 4 approximation for MSSC. As part of the analysis of our algorithm, we also derive an inequality on the lower tail of a sum of independent Bernoulli random variables, which might be of independent interest and broader utility.

Another well-known special case is the *min sum vertex cover* (MSVC) problem, in which the input hypergraph is a graph (i.e., $|e| = 2$) and $k_e = 1$, for every edge $e \in E$. We give a $16/9 \simeq 1.778$ approximation for MSVC, and show a matching integrality gap for the natural LP relaxation. This improves upon the previous best 1.999946 approximation of Barenholz, Feige and Peleg [6]. (The claimed 1.79 approximation result of Iwata, Tetali and Tripathi [21] for the MSVC turned out have an unfortunate, seemingly unfixable, mistake in it.)

Finally, we revisit MSSC and consider the ℓ_p norm of cover-time of the hyperedges. Using a dual fitting argument, we show that the natural greedy algorithm simultaneously achieves approximation guarantees of $(p+1)^{1+1/p}$, for all

$p \geq 1$, giving another proof of the result of Golovin, Gupta, Kumar and Tangwongsan [13], and showing its tightness up to NP-hardness. For $p = 1$, this gives yet another proof of the 4 approximation for MSSC.

1 Introduction

In the *min sum set cover problem (MSSC)*, formally introduced by Feige, Lovász and Tetali [11], given a collection of sets whose union is V, one seeks an ordering of the elements of V so as to minimize the sum of the cover times of the sets. Feige et al. showed that a natural greedy algorithm provides a factor 4 approximation (a result implicit in the prior work of Bar-Noy et al. [5], albeit with a more complicated proof), and also established that it is hard to approximate it to within a ratio $4 - \epsilon$, for every $\epsilon > 0$.

Azar, Gamzu and Yin [1] introduced a significant generalization of MSSC as the multiple intents re-ranking or the **generalized min sum set cover (GMSSC)** problem. In GMSSC, the input consists of a hypergraph $H = (V, E)$ where every hyperedge has a covering requirement k_e where $k_e \leq |e|$. The output is an ordering of the vertices, i.e., an assignment of V to time slots $1, \ldots, n$. We say that e is covered at time t, if t is the earliest time by which k_e of the vertices contained in e have appeared in the ordering. The objective is to minimize the total cover-time of all the hyperedges. Note that the case of $k_e = 1$, for all $e \in E$, corresponds to MSSC with the hyperedges corresponding to the sets. GMSSC has applications in query results diversification and broadcast scheduling among others [1, 23, 24, 30].

Azar et al. provided an $O(\log r)$ approximation for the problem, where $r = \max_e k_e$. Bansal, Gupta, and Krishnaswamy [3] were the first to provide a constant-factor (of 485) approximation for GMSSC by introducing a strong LP formulation using the so-called Knapsack Cover (KC) inequalities. Their approximation bound was improved to 28, by Skutella and Williamson [29], using the idea of α-point rounding and other technical enhancements. This has since been improved to 12.4 by Im, Sviridenko and van der Zwaan [20]. They introduce a different configuration LP, and show how to obtain a *preemptive* schedule from this LP losing a factor 2, and then round it to obtain the final solution at another factor 6.2 loss. In the quest for the best possible 4 approximation, they conjecture that any

preemptive schedule can be rounded to an integral solution at factor 2 loss. Another interesting version of MSSC is where the objective is to minimize the ℓ_p norm of the cover times for $p \geq 1$. For this setting, [13] showed that the greedy algorithm of [11] (which is oblivious to p) gives a $\delta_p := (p+1)^{1+1/p} \leq p + \ln p + 3$ approximation (simultaneously for every p), together with a hardness of approximation of $\Omega(p)$ assuming $P \neq NP$. Their analysis of the greedy algorithm uses a proof technique similar to that of [11].

Min Sum Vertex Cover (MSVC) is a well-known special case of MSSC, in which the hypergraph is a graph, i.e. $|e| = 2$ and $k_e = 1$, $\forall e \in E$. MSVC first arose in optimizing matrix computations involved in the analysis of a heuristic to speed up semidefinite program solvers [7]. Feige et al. [11], provided a randomized 2-approximation algorithm for MSVC, which was later improved to 1.999946 by Barenholz, Feige, and Peleg [6]. It should be noted that a 1.79-approximation result for MSVC, reported in [21], turned out to have a crucial error in the proof of Lemma 1 of [21] according to its authors, thus invalidating the claimed 1.79 bound.

1.1 Our results and techniques Despite the only lower bound on the approximability of GMSSC being 4 for the special case of MSSC by Feige et al., the case of arbitrary covering requirements has remained difficult to approximate within a factor close to 4. We obtain the following result for GMSSC.

THEOREM 1.1. *There is a polynomial time* 4.642-*approximation algorithm for the generalized min sum set cover (GMSSC) problem.*

Numerical evidence suggests that the approximation ratio of our algorithm is no more than 4.5232. We also note that the approach of [20] cannot give an approximation better than 8. In particular, we give a counter-example to their conjecture on the gap between a preemptive schedule and a non-preemptive one.

All known approaches for GMSSC start with obtaining a fractional solution to some strengthening of the natural LP (which has unbounded integrality gap [3]). Applying a naive randomized rounding however, fails even for $k_e = 1$ as the LP may cover a hyperedge e at time t by scheduling r of its vertices to extent $1/r$, while randomized rounding will leave it uncovered with probability $(1 - 1/r)^r \approx 1/e$ for times much greater than t. Therefore, all approaches use some additional idea to increase the probability of a hyperedge being covered early relative to its LP cost.

Our approach. The key to our results is to use a careful linear transformation to perform such a modification. Given

a fractional assignment $x_{v,t}$ of vertices to time-slots, we obtain a new solution $z_{v,t}$ by applying the linear transformation K, referred to henceforth as the *kernel*,

$$(1.1) \qquad z_{v,t} = \sum_{t'} K(t,t') x_{v,t'}.$$

We then apply a standard randomized rounding (α-point rounding) to z. To prove Theorem 1.1, we use the kernel $K(t,t') = \beta/t$ (for $t \geq t'$), for a suitable constant β. To see the idea, suppose for some v, $x_{v,1} > 0$ and $x_{v,t} = 0$ for $t > 1$. Then, $z_{v,t} \approx \beta x_{v,1}/t$. Roughly speaking, the solution x is "spread" to the right at a rate β/t, and eventually the cumulative amount of v scheduled by z will be arbitrarily large.

Choosing $\beta = 2$ gives a new LP rounding based 4-approximation for MSSC, refining the result in [11], as the guarantee there was with respect to the integral optimum solution.

THEOREM 1.2. *There is an LP-rounding based* 4-*approximation algorithm for MSSC.*

Min Sum Vertex Cover. For MSVC, despite significant effort to improve the factor 2 approximation of [11], no approximation guarantee substantially better than 2 is known [6]. We show the following.

THEOREM 1.3. *There is a polynomial time* 16/9 *approximation for MSVC. Further, this is tight with respect to the natural LP relaxation.*

A key idea is to use a different kernel for MSVC based on the following observation: for any hyperedge e, the LP must schedule some vertex $v \in e$ to an extent of at least $1/2$ (in contrast to $1/r$ for MSSC above). As it suffices for randomized rounding that v is scheduled by z to extent 1, we can spread x much less aggressively than for MSSC. We show that using the kernel $K(t,t') \approx \beta(t')^2/t^3$ (for $t \geq t'$) gives us Theorem 1.3.

ℓ_p **norm Set Cover.** Finally, we consider the ℓ_p norms of cover times for MSSC for $p \geq 1$. We give two different proofs, giving other perspectives, for the $\delta_p = (p+1)^{1+1/p}$ approximation of [13]. In particular, we give a *dual fitting* argument for the greedy algorithm, and a randomized rounding based proof using the kernel $K(t,t') = (p+1)/t$ (for $t \geq t'$). We also show an (exact) matching hardness of approximation result of δ_p for every $p \geq 1$.

THEOREM 1.4. *For any $p \in [1, \infty)$ the greedy algorithm guarantees a $\delta_p = (p+1)^{1+1/p}$ approximate solution for the min ℓ_p norm set cover problem. Further, it is NP-hard to approximate the min ℓ_p norm set cover problem better than $\delta_p - \epsilon$ for any $\epsilon > 0$ and any $p \geq 1$.*

In particular, as $\delta_p = 4$ for $p = 1$, the dual-fitting argument in the proof of Theorem 1.4 provides yet another LP based 4-approximation for MSSC.

Analysis. While our algorithms are quite simple, the analysis is more subtle. A key idea which allows us to get tight or close to tight bounds in Theorems 1.1-1.3, is the setting up of a non-linear convex optimization problem to find the worst-case fractional solution which maximizes our approximation ratio.

For GMSSC, the solution to the optimization problem above, is an expression in terms of the following function $P(\gamma)$, involving the sum of Bernoulli random variables: Given independent Bernoulli random variables Y_v with $\mathbb{E}\left[\sum_v Y_v\right] = \gamma k$, where k is a positive integer and $\gamma \geq 1$. What is the best upper bound $P(\gamma)$ on $\Pr[\sum_v Y_v \leq k-1]$, as a function of γ?

Standard Chernoff-Hoeffding or other tail bounds are inadequate for our purposes as they lose relatively large constants. We use the following refined bound which may be of independent interest.

THEOREM 1.5. *Let S be the sum of n independent, not necessarily identical, Bernoulli random variables. For any integer $k \leq \mathbb{E}[S]$, letting $\gamma = \mathbb{E}[S]/k \geq 1$, we have*

$$\Pr[S < k] \leq e^{-\gamma}(e/2)^{e^{1-\gamma}}.$$

In particular, for $\gamma = 1$, $P(\gamma) = 1/2$ which is the best possible as $\Pr[S \leq k - 1]$ approaches $1/2$ for large k and $\mathbb{E}[S] = k$. On the other extreme, for large γ, $P(\gamma)$ approaches $e^{-\gamma}$. This again is the best possible as for $k = 1$, $\Pr[S \leq k - 1] = \Pr[S = 0] \approx \exp(-\gamma)$.

The bound above can probably be strengthened. Numerical evidence suggests that the stronger bound $Pr[S < k] \leq e^{-\gamma}(e/2)^{e^{c(1-\gamma)}}$ with $c = 13.3$ holds for all k. If true, this would imply an improved bound of 4.5232 for GMSSC in Theorem 1.1 as mentioned previously. However, we remark that this is not the only reason that our bound for GMSSC is not 4. Even if we assume the best possible function $P(\gamma) = e^{-\gamma}$, the best achievable bound in Theorem 1.1 using our approach is 4.509. This is because unlike in the argument for MSSC, the use of Knapsack Cover inequalities in the argument for GMSSC leads to an additional loss.

1.2 Other Related Works A well-studied special case of GMSSC is when all $k_e = |e|$, a problem known as the *min-latency set cover*, for which Hassin and Levin [17] gave an e-approximation. However, this is equivalent to the classical problem of minimizing weighted completion time with precedence constraints, $1|\text{prec}|\sum_j w_j C_j$, [31], for which several 2-approximations are known [8, 14, 25]. A matching hardness of $2 - \epsilon$ (for every $\epsilon > 0$) is also known, assuming a variant of the Unique Games Conjecture [4]. Close to

this end also lies the *all-but-one MSSC*, another special case of GMSSC, with $k_e = \max\{1, |e| - 1\} \forall e \in E$, for which Happach and Schulz [16] provided a 4-approximation.

MSSC is closely related to the min-sum graph coloring problem [5] and has found interesting applications in peer to peer networks [9], data streams [2], and data-base query processing optimization [27]. It has inspired more general models and their analyses, e.g., pipelined set cover problem [27] and versions with precedence-constraints [16, 26]. Finally variants of MSSC under submodular and supermodular cost functions have also been studied [12, 15, 19, 21, 28].

1.3 Organization The rest of the paper is organized as follows.

In Section 2, we describe the LP formulation for GMSSC based on the Knapsack Cover inequalities, and the template based on kernels and α-point rounding that we use in all our algorithms.

In Section 3, we consider the GMSSC problem. We begin with the simpler MSSC and min latency problems, describing a 4 and an e approximation for them respectively using our analysis framework. Building on these ideas we then present our main result for GMSSC in §3.3. The lower tail for sums of Bernoulli r.v.s is proved in Section 6 and Appendix B.

In Section 4, we consider the MSVC problem and prove the 16/9 approximation, and also show the matching integrality gap for the LP.

In Section 7, we consider the problem of minimization of the ℓ_p norms of cover times for MSSC. Finally, in Section 8, we describe the counter-example for the conjecture of [20].

2 The Basic Framework

We assume that time is slotted, and time t, for $t = 1, 2, \ldots$, refers to the time interval $(t - 1, t]$. We refer to edges and hyperedges interchangeably.

LP relaxation for MSSC and GMSSC. Consider the following natural formulation for MSSC. For each vertex v and time t we have a variable $x_{v,t}$, which is intended to be 1 if v is assigned to time t. For each edge e and time t, there is a variable $u_{e,t}$ which is intended to be 1 if e is not covered at the start of t. Subsequently, we have the following LP relaxation.

Minimize $\quad \displaystyle\sum_{e,t} u_{e,t} \quad$ s.t. $\quad u_{e,t}, x_{v,t} \geq 0, \quad \forall\, e, v, t$

$$\text{(2.2)} \qquad \sum_v x_{v,t} \leq 1, \quad \forall\, t$$

$$\text{(2.3)} \qquad u_{e,t} + \sum_{v \in e} \sum_{t' < t} x_{v,t'} \geq 1, \quad \forall\, e, t$$

Constraint (2.2) ensures that at most one vertex is assigned

to any time, and constraint (2.3) ensures that $u_{e,t}$ is 0 only if some $v \in e$ is scheduled strictly before t. Note that the objective can also be written equivalently as $\sum_t t(u_{e,t} - u_{e,t+1})$ as $u_{e,t} - u_{e,t+1}$ can be viewed as the amount of e fractionally completed at time t. Later on, we will use both these perspectives interchangeably.

For GMSSC, where the demands k_e are arbitrary, the natural extension of this LP becomes extremely weak, but as shown in [3], it can be strengthened by replacing (2.3) by the following Knapsack Cover (KC) inequalities:

$$(2.4) \quad (k_e - |S|)u_{e,t} + \sum_{v \in e \setminus S} \sum_{t' < t} x_{v,t'} \geq (k_e - |S|),$$
$$\forall\, e, t, S \subseteq e, |S| < k_e.$$

The constraints (2.4) require that for an edge e to be considered covered at time t, no matter which subset S of vertices in e we ignore, at least $(k_e - |S|)$ many must be scheduled from vertices in $e \setminus S$ before t. Henceforth, we will treat the LP with constraints (2.2) and (2.4) as the LP for GMSSC, unless otherwise specified. Even though it has exponentially many constraints, it can be solved efficiently using the Ellipsoid method to any desired accuracy, see [3] for details.

The algorithmic approach. All our algorithms follow the same basic template consisting of three steps:

First, compute some optimum solution x for the corresponding LP relaxation. For a vertex v, let $x_v = (x_{v,t})_{\{t\}}$ denote the vector corresponding to its fractional assignment.

Second, for each vertex v, apply a suitable (problem dependent) linear transformation K to x_v to obtain z_v. That is, for all v and all t,

$$(2.5) \quad z_{v,t} = \sum_{t'} K(t,t')x_{v,t'}.$$

We will refer to the matrix K as a kernel. The solution z may not satisfy (2.4) and (2.2) anymore, but K will be chosen so that z satisfies other useful properties.

Third, create a random tentative schedule τ (where multiple vertices may be assigned to the same time slot), by applying α-point rounding independently for each vertex v, using the solution z_v (see Algorithm 1 below). Next, convert the schedule τ to a proper schedule σ by scheduling one vertex at a time in the order given by τ, breaking ties at random (this randomness will be crucial for some of the problems we consider).

Analysis. The analyses of our algorithms will also follow the same template. We will do a per-edge analysis. We first introduce some notation that will be used throughout. For an edge e, let $c_x(e) = \sum_t u_{e,t}$ denote the cost for e in the LP solution x.

Algorithm 1 The algorithmic template

1: **procedure** KERNEL α-POINT ROUNDING($H = (V, E), k : E \to \mathbb{N}$)
2:     $x \leftarrow$ An optimum fractional schedule for H, k. ▷ i: Solving a Relaxation.
3:     $z_v \leftarrow Kx_v \quad \forall v \in V.$ ▷ ii: Applying the Kernel
4:     **for** $v \in V$ **do**   ▷ iii: α-point Rounding
5:         $\alpha_v \sim uniform[0,1].$
6:         $\tau_{v,t} \leftarrow \mathbb{1}[t =$ the earliest time for which $\sum_{t' \leq t} z_{v,t'} \geq \alpha_v] \quad \forall t.$ ▷ Tentative Schedule
7:     **return** an ordering σ of V : Scheduling vertices according to τ, breaking ties at random.

Given z, and the random choices α_v, the cover time $\mathrm{Cov}_\tau(e)$ of e in the (random) tentative schedule τ is the earliest time t, by which exactly $k(e)$ elements of e are scheduled in τ. Note that the only randomness in $\mathrm{Cov}_\tau(e)$ is due to the choice of α_v, for all $v \in e$.

Similarly, let $\mathrm{Cov}_\sigma(e)$ be the (random) cover time of e in the proper schedule σ, obtained from τ. The randomness in $\mathrm{Cov}_\sigma(e)$ is due to choices of α_v for all v (including vertices not in e, as these can *delay* vertices in e), and the randomness of the tie-breaking rule at each slot.

To implement the aforementioned, first we define a cost $c_z(e)$ that upper bounds $\mathbf{E}[\mathrm{Cov}_\tau(e)]$. Next, we bound the ratio of $c_z(e)$ to $c_x(e)$ and finally, we bound the expected cover time $\mathbf{E}[\mathrm{Cov}_\sigma(e)]$ under σ by a multiple of $c_z(e)$. Relating $c_z(e)$ to $c_x(e)$ will be the hardest part. For each of the problems we consider, we will bound this ratio by considering another optimization problem to solve for the worst-case setting of the variables $x_{v,t}$ that maximizes this ratio and bounding its optimum.

3 Generalized Min Sum Set Cover

In this section we develop the algorithm for GMSSC. To motivate the eventual analysis of the algorithm, we first start with the simpler MSSC problem and describe a tight factor 4 approximation. Then, we describe a factor e approximation for the min-latency version. Finally, we describe our algorithm for GMSSC in §3.3.

For all these problems we apply Algorithm 1, and we use a kernel of the form $K(t,t') = \frac{\beta}{t}$ for $t' \leq t$, and $K(t,t') = 0$ otherwise. Here β is a constant whose choice will be optimized separately for each problem.

Given an optimum solution x to the underlying LP, we obtain z by applying K, so for each v, t, we have

$$z_{v,t} = \frac{\beta}{t} \sum_{t' \leq t} x_{v,t'}.$$

This kernel has the interesting property that even if v is scheduled in the LP solution x to extent at most ϵ, the amount

of v in z will be arbitrarily large. E.g. suppose $x_{v,1} = \epsilon$ and 0 otherwise, then $z_{v,t} = \beta\epsilon/t$ and hence $\sum_t z_{v,t}$ diverges. Yet, crucially, the total *load* at any time step in z is bounded. In particular, observe:

CLAIM 1. *For any t, $\sum_v z_{v,t} \leq \beta$.*

Proof. As (2.2) ensures that $\sum_v x_{v,t'} \leq 1$ for any t', we have $\sum_v z_v = \sum_v K x_v \leq K\mathbf{1} = \beta\mathbf{1}$, where the desired inequality follows as K is a linear map with non-negative coefficients. $\quad\square$

Notation. Before proceeding, it will be useful to define a partial order $\succeq$ on non-negative vectors. For any vector $b = (b_1, b_2, \ldots)$, let $b_{<t} = b_1 + \ldots + b_{t-1}$ and $b_{\leq t} = b_1 + \ldots + b_t$. We call a vector non-negative if all its entries are non-negative. For two non-negative vectors a, a', we denote $a \succeq a'$ if $a_{\leq t} \geq a'_{\leq t}$ for all t. Note on the other hand that $a \geq b$ denotes $a_t \geq b_t$ for all t. Also, for a real number c, $\max(c, 0)$ is denoted by $(c)_+$.

We will use the following repeatedly.

CLAIM 2. *Let a be a non-negative vector and let $b = Ka$. Then for any time t,*

$$b_{<t} = \sum_{t'<t} b_{t'} \geq \beta \sum_{t'<t} a_{t'} \ln \frac{t}{t'} = \beta \sum_{t'\leq t} a_{t'} \ln \frac{t}{t'}.$$

Moreover, if $a \succeq a'$, then $Ka \succeq Ka'$.

Proof. By the definition of K, and using that $\sum_{i=c}^{d-1} \frac{1}{i} \geq \ln(\frac{d}{c})$, for positive integers c, d, we have that

$$b_{<t} = \sum_{t''<t} b_{t''} = \sum_{t''<t} \sum_{t'\leq t''} K(t'', t') a_{t'}$$
$$= \sum_{t'<t} a_{t'} \sum_{t'\leq t''<t} K(t'', t') \geq \beta \sum_{t'<t} a_{t'} \ln \frac{t}{t'}.$$

We now show $(Ka)_{\leq t} \geq (Ka')_{\leq t}$ for every t, whenever $a \succeq a'$. Fix some t.

Let $\gamma_{t'} = \sum_{t' \leq t'' \leq t} K(t'', t') = \sum_{t' \leq t'' \leq t} \frac{\beta}{t''}$, and note that $\gamma_{t'}$ is non-increasing in t'. Then,

$$(Ka)_{\leq t} = \sum_{t'\leq t} a_{t'} \sum_{t'\leq t''\leq t} K(t'', t') = \sum_{t'\leq t} a_{t'} \gamma_{t'}.$$

Recall the Abel summation formula that for any reals $c_1, c_2, \ldots, c_n$ and $d_1, d_2, \ldots, d_n$,

$$\sum_{k=1}^{n} c_k d_k = \sum_{k=1}^{n-1} c_{\leq k}(d_k - d_{k+1}) + c_{\leq n} d_n.$$

Applying this with $c = a$ and $d = \gamma$ gives

$$\sum_{t'=1}^{t} a_{t'} \gamma_{t'} = \sum_{t'=1}^{t-1} a_{\leq t'}(\gamma_{t'} - \gamma_{t'+1}) + a_{\leq t} \gamma_t.$$

As $\gamma_{t'}$ in non-increasing and $\gamma_t \geq 0$, each term is non-negative we have that $(Ka)_{\leq t} \geq (Ka')_{\leq t}$ for $a \succeq a'$. $\quad\square$

Let us state another simple fact that we will use several times later on.

FACT 3.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function that is non-negative on some compact convex region H, and let $g : \mathbb{R}^n \to \mathbb{R}$ be a linear function that is strictly positive on H, then the maximum of $\max_{x \in H} f(x)/g(x)$ is attained at an extreme point of H.*

Proof. Suppose that the maxima is attained at some z in the interior of H, and let us write $z = \lambda x + (1-\lambda)y$ where x, y are either extreme points (or both lie on a lower dimensional face) of H with $\lambda \in (0, 1)$. As f is non-negative and g is strictly positive on H,

$$\max\left(\frac{f(x)}{g(x)}, \frac{f(y)}{g(y)}\right) \geq \frac{\lambda f(x) + (1-\lambda)f(y)}{\lambda g(x) + (1-\lambda)g(y)}$$
$$\geq \frac{f(\lambda x + (1-\lambda)y)}{g(\lambda x + (1-\lambda)y)}$$
$$= \frac{f(z)}{g(z)},$$

where the first inequality uses that

$$\max\left(\frac{a}{b}, \frac{c}{d}\right) \geq \frac{a+c}{b+d}$$

for any non-negative a, b and positive c, d, and the second inequality uses the convexity of f and the linearity of g.

Thus, the value is at least as large as $f(z)/g(z)$ at one of x or y. Iterating the process eventually gives an extreme point with value at least as large as $f(z)/g(z)$. $\quad\square$

3.1 Min Sum Set Cover Here we show the following result for MSSC, refining the previous approximation of [11], which was with respect to the integral optimum solution.

THEOREM 3.1. *For any $\beta \geq 1$, the algorithm described above is a $\beta^2/(\beta-1)$-approximation. In particular, setting $\beta = 2$ gives a 4-approximation.*

Before proving this theorem, we first give some notation. Recall that $\mathrm{Cov}_\tau(e)$ is the cover time of e in the tentative schedule τ, and $\mathrm{Cov}_\sigma(e)$ in the proper schedule σ. Moreover, $c_x(e)$ is the LP cost for e.

By the definition of the schedule τ, note that $\mathbf{E}[\mathrm{Cov}_\tau(e)] = \sum_t p_t(e)$, where $p_t(e)$ is the probability that edge e is uncovered at the beginning of time-slot t in τ. As each v is rounded independently using α-point rounding,

$$(3.6) \qquad p_t(e) = \prod_{v\in e}(1 - z_{v,<t})_+ \leq \exp(-\sum_{v\in e} z_{v,<t}).$$

Let us define

$$c_z(e) = \sum_t \exp(-\sum_{v \in e} z_{v,<t}),$$

and note that it upper bounds $\mathbb{E}\left[\mathrm{Cov}_\tau(e)\right]$. Theorem 3.1 is a direct consequence of the next two lemmas.

LEMMA 3.1. *For each edge* e, $c_z(e) \leq \frac{\beta}{\beta-1} c_x(e)$.

LEMMA 3.2. *For each edge* e, $\mathbb{E}\left[Cov_\sigma(e)\right] \leq \beta c_z(e)$.

Intuitively, Lemma 3.2 follows from what we observed as Claim 1: As τ has at most β vertices in expectation at any time t, the expected cover time of e in σ should be at most β times that in τ, so can be upper bounded by $\beta c_z(e)$. However, to make this intuition precise, one needs to *condition* on when some vertex of e was first scheduled in τ.

We address this in §5, and focus here on Lemma 3.1 instead.

Proof. [Proof of Lemma 3.1] As $z_v = Kx_v$, by Claim 2, we have that $z_{v,<t} \geq \beta \sum_{t' \leq t} x_{v,t'} \ln(t/t')$, and hence

(3.7)
$$\begin{aligned}
c_z(e) &= \sum_t \exp(-z_{v,<t}) \\
&\leq \sum_t \exp\Big(-\beta \sum_{v \in e} \sum_{t' \leq t} x_{v,t'} \ln(t/t')\Big).
\end{aligned}$$

Recalling that

$$c_x(e) = \sum_t u_{e,t} = \sum_t (1 - \sum_{v \in e} x_{v,<t})_+,$$

to prove the lemma, it suffices to show that

(3.8)
$$\begin{aligned}
&\sum_t \exp\Big(-\beta \sum_{t' \leq t} \sum_{v \in e} x_{v,t'} \ln(t/t')\Big) \\
&\leq \frac{\beta}{(\beta-1)} \sum_t (1 - \sum_{v \in e} x_{v,<t})_+.
\end{aligned}$$

holds for any setting of $x_{v,t}$ variables.

Let us define $x_{e,t} = u_{e,t} - u_{e,t+1}$ (the amount of e satisfied at time t in solution x). Noting that $x_{e,t} \leq \sum_{v \in e} x_{v,t}$, replacing $\sum_{v \in e} x_{v,t'}$ by $x_{e,t'}$ can only increase the left hand side of (3.8).

Moreover, $x_{e,<T} = 1$ for large enough T, as e will eventually be satisfied by the LP. So, the cost $c_x(e)$ can be written as $\sum_t t x_{e,t}$ and our goal is to show that for any $x_{e,t} \geq 0$ satisfying $\sum_t x_{e,t} = 1$, the following inequality holds:

$$\sum_t \exp\Big(-\beta \sum_{t' \leq t} x_{e,t'} \ln(t/t')\Big) \leq \frac{\beta}{(\beta-1)} \sum_t t x_{e,t}.$$

To this end, we consider the following optimization problem in the variables a_t.

$(\mathcal{F})$ Maximize $\dfrac{\sum_t \exp(-\beta \sum_{t' \leq t} a_{t'} \ln \frac{t}{t'})}{\sum_t t a_t}$,

          Subject to $\|a\|_1 = 1$, $a_t \geq 0$.

CLAIM 3. *For any* $\beta > 1$, *the optimum value of the problem* $\mathcal{F}$ *is at most* $\beta/(\beta-1)$.

Proof. Consider the numerator of the objective in $\mathcal{F}$. Each summand $\exp(-\beta \sum_{t' \leq t} a_{t'} \ln(t/t'))$ is of the form $f = \exp(-\sum_{t'} c_{t'} a_{t'})$ (note $c_{t'} = \beta \ln(t/t')$ are constants), and hence is convex in a. As the sum of convex functions is convex, the numerator is convex. Moreover, the domain H of a is the non-negative simplex and the denominator of $\mathcal{F}$ is linear and non-negative on H. By Fact 3.1, the optimum of $\mathcal{F}$ is at some extreme point of the unit simplex, which must be of the form $a_u = 1$ for some u and $a_{u'} = 0, u' \neq u$.

For such an extreme point, the objective of $\mathcal{F}$ is easy to compute. In particular, the denominator is exactly u and the numerator is

$$\begin{aligned}
\sum_{t=1}^u 1 + \sum_{t>u} \exp(-\beta \ln t/u) &= u + \sum_{t>u} (u/t)^\beta \\
&\leq u + \int_{x=u}^\infty (u/x)^\beta dx \\
&= u\Big(1 + \frac{1}{\beta-1}\Big) \\
&= u\Big(\frac{\beta}{\beta-1}\Big).
\end{aligned}$$

$\square$

This proves that $c_z(e) \leq (\beta/(\beta-1)) c_x(e)$ as desired. $\square$

3.2 Min Latency Set Cover We now consider the min latency setting where $k_e = |e|$, and show the following.

THEOREM 3.2. *For any* $\beta \geq 1$, *the algorithm is a* $\beta e^{1/\beta}$-*approximation. In particular for* $\beta = 1$, *this gives an* e-*approximation.*[1]

For min latency, the constraints (2.4) reduce to simply that for each e, t and $v \in e$, that $u_{e,t} + x_{v,<t} \geq 1$. In particular, we can assume $u_{e,t} = \max_{v \in e}(1 - x_{v,<t})$. Let us define $x_{e,t} = u_{e,t} - u_{e,t+1}$ (the amount of e satisfied by the LP at time t), and note that $x_{e,<t} = 1 - u_{e,t}$. So we have that $x_{v,<t} \geq x_{e,<t}$ for all $v \in e$ and all t, i.e., $x_v \succeq x_e$.

For an edge e, let t_e be the earliest time such that $z_{v,\leq t_e} \geq 1$ for all $v \in e$. We define $c_z(e) = t_e$. Note

---

[1]The use of e both for an edge and the constant $e = 2.718\ldots$ is a bit unfortunate, but hopefully it is not confusing.

that the α-point rounding will surely schedule each $v \in e$ in τ by time t_e, and hence $\mathrm{Cov}_\tau(e) \leq c_z(e)$ with probability 1.

Theorem 3.2 is a direct consequence of the following Lemmas 3.3 and 3.4.

LEMMA 3.3. *For each edge e, $c_z(e) \leq e^{1/\beta} c_x(e)$.*

Proof. Fix an edge e. Let t^* (not necessarily an integer) be such that

$$(3.9) \qquad \beta \sum_{t' \leq \lfloor t^* \rfloor} x_{e,t'} \ln \frac{t^*}{t'} = 1 \,.$$

We claim that $t_e \leq \lfloor t^* \rfloor$. This follows as for any $v \in e$, by Claim 2 with $t = \lfloor t^* \rfloor + 1$,

$$\begin{aligned} z_{v,\leq \lfloor t^* \rfloor} &= (Kx_v)_{\leq \lfloor t^* \rfloor} \\ &\geq (Kx_e)_{\leq \lfloor t^* \rfloor} \\ &\geq \beta \sum_{t' \leq \lfloor t^* \rfloor} x_{e,t'} \ln \frac{\lfloor t^* \rfloor + 1}{t'} \\ &\geq \beta \sum_{t' \leq \lfloor t^* \rfloor} x_{e,t'} \ln \frac{t^*}{t'} \\ &= 1 \end{aligned}$$

where the first and second inequalities use Claim 2 and $x_v \succeq x_e$, and the last equality uses (3.9).

As $c_z(e) = t_e \leq \lfloor t^* \rfloor \leq t^*$, to prove the lemma it suffices to show that $c_x(e) \geq e^{-1/\beta} t^*$. As $c_x(e) = \sum_t t x_{e,t}$, this is equivalent to showing that the optimum solution of the following optimization problem with variables a_t (corresponding to $x_{e,t}$), is at least $e^{-1/\beta} t^*$.

$$(3.10)$$
$$(\mathcal{F}) \qquad \text{Minimize} \quad \sum_t t a_t$$
$$\text{Subject to} \quad \|a\|_1 = 1 \text{ and } \beta \sum_{t \leq \lfloor t^* \rfloor} a_t \ln \frac{t^*}{t} = 1 \,.$$

The first constraint follows as $x_{e, \leq T} = 1$ for T large enough (since there are n vertices, we can assume that $x_{v, \leq n} = 1$ for all v), and the second constraint is by the relation (3.9) defining t^*.

CLAIM 4. *The optimum of $\mathcal{F}$ is at least $t^* e^{-1/\beta}$.*

Proof. As $\mathcal{F}$ is an LP with 2 non-trivial constraints, there is some optimum solution with at most 2 non-zero a_t. Let u and v be those indices with $u \leq v$ and let $a_u = s$ and $a_v = 1 - s$ so that $\|a\|_1 = 1$.

We consider two cases, depending on whether $v \leq \lfloor t^* \rfloor$ or not. Suppose first that $v \leq \lfloor t^* \rfloor$. Then $(\mathcal{F})$ reduces to the

following problem on three variables u, v, s.

$$\begin{aligned} &\text{Minimize } su + (1-s)v \\ &\text{Subject to } s \ln(t^*/u) + (1-s)\ln(t^*/v) = 1/\beta, \\ &\qquad\qquad s \in [0,1], u \leq v \leq \lfloor t^* \rfloor \,. \end{aligned}$$

Setting $u = t^* e^{-c}$ and $v = t^* e^{-d}$ where $c, d \geq 0$ as $u, v \leq t^*$, this becomes

$$\begin{aligned} &\text{Minimize} \quad t^*(se^{-c} + (1-s)e^{-d}) \\ &\text{Subject to} \quad (sc + (1-s)d) = 1/\beta, \\ &\qquad\qquad s \in [0,1], c \geq d \geq 0. \end{aligned}$$

But as e^{-x} is convex, the objective is always at least $t^*(e^{-(sc+(1-s)d)}) = t^*(e^{-1/\beta})$, as claimed.

Next, suppose that $v > \lfloor t^* \rfloor$. The value a_v does not affect t^* here, and $(\mathcal{F})$ now becomes

$$\begin{aligned} &\text{Minimize} \quad su + (1-s)v \\ &\text{Subject to} \quad s\ln(t^*/u) = 1/\beta, \\ &\qquad\qquad s \in [0,1], u \leq \lfloor t^* \rfloor, v > \lfloor t^* \rfloor \,. \end{aligned}$$

The constraint implies $u = t^* e^{-1/s\beta}$, and let us further relax the constraints to $u \leq t^*, v \geq t^*$ as this can only reduce the objective. As the only constraint on v is $v \geq t^*$, $v = t^*$ in any optimum solution.

So the problem reduces to minimizing

$$t^*(se^{-1/s\beta} + (1-s))$$

for $s \in [0,1]$. Setting $f(s) = (se^{-1/s\beta} + (1-s))$ and checking that $f'(s) = e^{-1/s\beta}(1 + 1/s\beta) - 1 \leq 0$ for $s \geq 0$ (as $1 + x \leq e^x$ for all x), we get that it is minimized at $s = 1$ with value $t^* e^{-1/\beta}$. This completes the proof of the claim. □

Putting all this together, $c_z(e) \leq t^*$ and $c_x(e) \geq t^* e^{-1/\beta}$, which completes the proof of Lemma 3.3. □

Finally we relate the expected cost under σ to $c_z(e)$. The argument is in fact quite simple for min-latency.

LEMMA 3.4. *For each edge e, $\mathbb{E}[Cov_\sigma(e)] \leq \beta t_e = \beta c_z(e)$.*

Proof. As all the vertices of e are scheduled by t_e in τ, the cover time $\mathrm{Cov}_\sigma(e)$ of e in σ is trivially upper bounded by the number of vertices not in e that appear in τ by time t_e, plus k_e, the number of vertices in e needed to cover e.

As $\sum_v z_{v, \leq t_e} \leq \beta t_e$ by Claim 1, and as $z_{v, \leq t_e} \geq 1$ for each $v \in e$, the expected number of vertices not in e that appear by time t_e is $\sum_{v \notin e} z_{v, \leq t_e} \leq \beta t_e - k_e$. So $\mathbb{E}[\mathrm{Cov}_\sigma(e)] \leq \beta t_e - k_e + k_e = \beta t_e$. □

3.3 Generalized Min Sum Set Cover We now consider the setting with general k_e. The algorithm is exactly the same as before: we use the kernel $K(t, t') = \beta/t$ for $t' \leq t$, and follow the algorithmic template. Our analysis will (necessarily) combine aspects from both MSSC and the min-latency setting, and will be more technically involved. We will also need to use the KC inequalities in a careful way.

Briefly recalling earlier notation, let x denote some optimum LP solution, and let z be obtained by applying the kernel K to x, and let τ be the tentative schedule obtained by applying the α-point rounding to z for each v.

Fix a time t. For each vertex v, let Y_v be a Bernoulli random variable which is 1 if v is picked before t in schedule τ and 0 otherwise. Then, by the property of α-point rounding $\mathbb{E}[Y_v] = \min(1, z_{v,<t})$, the random variables Y_v for different vertices v are independent, and the probability that an edge e is not yet covered before t in τ can be written as

$$(3.11) \qquad p_t(e) = \Pr[\sum_{v \in e} Y_v < k_e].$$

The $P(\cdot)$-function. As $\mathbb{E}[\mathrm{Cov}_\tau(e)] = \sum_t p_t(e)$, this motivates understanding the following question:

Given Bernoulli random variables Y_v with $\mathbb{E}[\sum_v Y_v] = \gamma k$, where k is a positive integer and $\gamma \geq 1$, what is the best upper bound $P(\gamma)$ on $\Pr[\sum_v Y_v \leq k-1]$, as a function of γ?

Remark. Note that we require $P(\gamma)$ to only depend on γ (as k and $\mathbb{E}[Y_v]$, will be completely arbitrary for us). Moreover, we assume that $\gamma \geq 1$, since if $\gamma < 1$, we simply assume the worst case that $p_t(e) = 1$ (indeed if $\gamma < 1$, then as k_e get large, the probability that e is covered by time t will tend to 0).

Let us first note that for $k = 1$, this question is precisely $\Pr[\sum_v Y_v = 0]$ if $\mathbb{E}[\sum_v Y_v] = \gamma$, which tends to $\exp(-\gamma)$ (e.g. in the Poisson regime). So the best $P(\gamma)$ we can hope for is $\exp(-\gamma)$. However, in general $P(\gamma)$ must be strictly worse: in particular, if k is large and $\gamma = 1$, then $\Pr[\sum_v Y_v \leq k-1]$ is arbitrarily close to $1/2$, see [22].

One of our main technical contributions is coming up with refined functions $P(\gamma)$ that are essentially optimum[2]. In particular Theorem 1.5 provides us with

$$(3.12) \qquad P(\gamma) \leq e^{-\gamma}(e/2)^{\exp(1-\gamma)}.$$

Note that here, $P(1) = 1/2$ (the best we can hope for). Moreover, and as γ increases, the factor $(e/2)^{\exp(1-\gamma)}$ approaches rapidly to 1 (as $\exp(1-\gamma)$ approaches 0), and hence

---

[2]One could use standard Chernoff-Hoeffding bounds or other tail bounds for sums of Bernoulli random variables, but these are quite crude for our purposes as we lose relatively large constants, leading to substantially worse approximation factors.

$P(\gamma)$ is essentially $e^{-\gamma}$ for larger γ. So the bound is close to the best possible for the entire range of γ. The derivation of this bound is rather technical and is deferred to the Appendix. In §6, we show a weaker (but still quite non-trivial) bound of $P(\gamma) = e^{1-\gamma}/2$.

Our result. Our main result is the following, which we state in terms of a general function $P(\cdot)$.

THEOREM 3.3. *For any $\beta \geq 1$, and any tail bound function $P(\cdot)$ that is convex and non-increasing on $[1, \infty)$ and satisfies $P(1) \leq 1$, the approximation ratio of the algorithm is at most $\beta r(\beta)$ where*

$$(3.13) \qquad r(\beta) := e^{1/\beta}\left(1 + \frac{1}{\beta}\int_1^\infty P(x)e^{\frac{x-1}{\beta}}dx\right).$$

In particular, for the bound $P(\gamma) = e^{1-\gamma}/2$, this gives the approximation $\beta(2\beta - 1)e^{1/\beta}/(2\beta - 2)$ which has value at most 4.9102, for $\beta = 2.191$. Using the stronger bound (3.12) for $P(\gamma)$, the approximation factor is at most 4.642, for $\beta = 2.0715$.

Remark. Based on numerical computations, it seems that $P(\gamma)$ satisfies the stronger bound

$$P(\gamma) \leq e^{-\gamma}(e/2)^{\exp(c(1-\gamma))},$$

for $c = 13.3$ (note that the bound in (3.12) corresponds to $c = 1$). However, we are unable to prove this analytically. Plugging this better bound for P gives an approximation ratio of at most 4.5232.

We do not pursue further optimizations of the function $P(\gamma)$ as 4.509 seems a natural bottleneck for our approach, as even assuming the best possible function $P(\gamma) = \exp(-\gamma)$, Theorem 3.3 only gives a 4.509 approximation Roughly, the reason for this loss is that for $k_e = 1$ (where we got the 4 approximation), we could use that an edge e had a decent chance of being covered even for $\gamma < 1$ (i.e., $P(\gamma) = e^{-\gamma}$ even for $0 \leq \gamma \leq 1$). However for arbitrary k_e, we are unable to say anything better than $P(\gamma) = 1$ for $\gamma \leq 1$. In particular, it seems that getting a bound of 4 (if possible at all) will require further new ideas.

Analysis. Fix an edge e. and as previously, define $x_{e,t} = u_{e,t} - u_{e,t+1}$ (the fraction of e satisfied by the LP at t). Then $c_x(e) = \sum_t u_{e,t} = \sum_t t x_{e,t}$. Let $z_e = Kx_e$ be obtained by applying the kernel K to x_e. Let t_e be the earliest time such that $z_{e,\leq t_e} \geq 1$. Let us define

$$(3.14) \qquad c_z(e) = t_e + \sum_{t > t_e} P(z_{e,<t}).$$

We will show that $c_z(e)$ is an upper bound on the expected cost $\mathbb{E}[\mathrm{Cov}_\tau(e)]$ of e in the tentative schedule τ. Theorem 3.3 will follow directly from the following two lemmas.

LEMMA 3.5. *For any edge e, $c_z(e) \leq r(\beta)c_x(e)$.*

LEMMA 3.6. *For each edge e, $\mathbb{E}[Cov_\sigma(e)] \leq \beta c_z(e)$.*

Proof. (Lemma 3.5) We first show that $\mathbb{E}[Cov_\tau(e)] \leq c_z(e)$.

Recall by (3.11) that $p_t(e)$ denotes the probability that e is not covered before t in τ. Then $\mathbb{E}[Cov_\tau(e)] = \sum_t p_t(e)$. To upper bound $p_t(e)$ we assume that $p_t(e) = 1$ for $t \leq t_e$. Now, fix some time $t > t_e$. We claim that $p_t(e) \leq P(z_{e,<t})$. This directly gives that $\mathbb{E}[Cov_\tau(e)] \leq c_z(e)$, by the definition of $c_z(e)$ in (3.14).

Let us define $A_e(t) := \{v \in e : z_{v,<t} \geq 1\}$ to be the set of vertices in e that are surely picked in schedule τ before time t, and let $B_e(t) = e \setminus A_e(t)$ be the remaining vertices in e. Let $k_e(t) = \max(k_e - |A_e(t)|, 0)$ be the number of vertices that must still be picked from $B_e(t)$ to satisfy e.

By the KC inequality for e with $S = A_e(t)$, the LP satisfies

$$u_{e,t} + \sum_{v \in B_e(t)} x_{v,<t}/(k_e(t)) \geq 1,$$

or equivalently, $\sum_{v \in B_e(t)} x_{v,<t} \geq k_e(t)x_{e,<t}$. Applying the kernel K on both sides then gives (via Claim 2)

$$(3.15) \qquad \sum_{v \in B_e(t)} z_{v,<t} \geq k_e(t)z_{e,<t}.$$

Now e is not covered in τ before t, if and only if strictly fewer than $k_e(t)$ vertices from $B_e(t)$ are picked in τ. Using the random variables Y_v described earlier, this is exactly if $\sum_{v \in B_e(t)} Y_v < k_e(t)$, and hence from the definition of $P(\cdot)$ and (3.15), it follows that $p_t(e) \leq P(z_{e,<t})$, as claimed.

We now upper bound $c_z(e)$ in terms of $c_x(e)$. Let t^* (not necessarily integer) be such that

$$\beta \sum_{t' \leq \lfloor t^* \rfloor} x_{e,t'} \ln(t^*/t') = 1.$$

By Claim 2, as $z_e = Kx_e$, we have $z_{e,\leq t} \geq \beta \sum_{t' \leq t} x_{e,t'} \ln(t+1)/t'$. Hence, $z_{e,\leq \lfloor t^* \rfloor} \geq 1$ implying that $\lfloor t^* \rfloor \geq t_e$. Moreover as $P(\gamma)$ is non-increasing for $\gamma \geq 1$ (by our hypothesis), we have

$$c_z(e) = t_e + \sum_{t > t_e} P(z_{e,<t})$$
$$\leq \lfloor t^* \rfloor + \sum_{t > \lfloor t^* \rfloor} P(z_{e,<t})$$
$$\leq \lfloor t^* \rfloor + \sum_{t > \lfloor t^* \rfloor} P\left(\beta \sum_{t' \leq t} x_{e,t'} \ln \frac{t}{t'}\right).$$

Our goal henceforth is to upper bound the maximum possible ratio of the right hand side above to $c_x(e)$.

Optimization problem. As previously, we set up the following optimization problem $\mathcal{F}$ with variables a_t (corresponding to $x_{e,t}$), and some fixed real t^*.

$$(\mathcal{F}) \quad \text{Maximize} \quad \frac{\lfloor t^* \rfloor + \sum_{t > \lfloor t^* \rfloor} P(\beta(\sum_{t' \leq t} a_{t'} \ln t/t'))}{\sum_t t a_t}$$
$$\text{Subject to} \quad \|a\|_1 = 1, \quad \beta \sum_{t' \leq \lfloor t^* \rfloor} a_{t'} \ln(t^*/t') = 1.$$

We bound this in Lemma 3.7 which will imply the result. □

LEMMA 3.7. *If P is non-increasing and convex in $[1, \infty)$, and satisfies $P(1) \leq 1$. Then, the optimal value of problem $\mathcal{F}$ is at most $r(\beta)$, defined in* (3.13).

Proof. Let us denote $\gamma_t = \beta \sum_{t' \leq t} a_t \ln t/t'$. The constraint in $\mathcal{F}$ implies that, $\gamma_t \geq 1$ for $t \geq t^*$. Also, as $\gamma_{t^*} = 1$, we have $P(\gamma_{t^*}) \leq 1$ by the assumption on P.

As the constraints of $\mathcal{F}$ are linear in a, the feasible region is a polytope. Moreover as γ_t is linear in a and P is convex, $P(\gamma_t)$ is convex in a. So the objective of $\mathcal{F}$ is the ratio of a non-negative convex function and a strictly positive linear function, and by Fact 3.1, attains its optimum at an extreme point. As there are only 2 non-trivial constraints, any extreme point has at most 2 non-zero variables, say a_u and a_v for $u \leq v$.

We consider two cases depending on whether $v \leq t^*$.

Case 1. $v \leq t^*$. Let us denote $a_u = s$, then $a_v = 1 - s$ (as $\|a\|_1 = 1$). Let us substitute $u = t^*e^{-a}$ and $v = t^*e^{-b}$, where $a, b \geq 0$. Then the second constraint in $\mathcal{F}$ is $\beta(sa + (1-s)b) = 1$. Moreover,

$$\gamma_t = \beta(s\ln(t/u) + (1-s)\ln(t/v))$$
$$= \beta \ln t/t^* + \beta(sa + (1-s)b)$$
$$= 1 + \beta \ln(t/t^*).$$

So, the numerator of the objective of $\mathcal{F}$ is simply $\lfloor t^* \rfloor + \sum_{t > \lfloor t^* \rfloor} P(1 + \beta \ln(t/t^*))$, which surprisingly is completely independent of the parameters a_u, a_v, s.

So optimizing $\mathcal{F}$ simply reduces to minimizing the denominator. In particular,

$$\text{Minimize} \quad t^*(se^{-a} + (1-s)e^{-b})$$
$$\text{Subject to} \quad sa + (1-s)b = 1/\beta, \quad a \geq b \geq 0, s \in [0, 1].$$

As $\exp(-x)$ is convex, by Jensen's inequality $se^{-a} + (1-s)e^{-b} \geq e^{-sa-(1-s)b} = e^{-1/\beta}$, and hence the denominator is lower bounded by $t^*e^{-1/\beta}$.

So any optimum solution of $\mathcal{F}$ is at most

$$e^{1/\beta}\Big(\frac{\lfloor t^*\rfloor}{t^*} + \frac{1}{t^*}\sum_{t>\lfloor t^*\rfloor} P\left(1+\beta\ln(t/t^*)\right)\Big)$$

$$\leq e^{1/\beta}\left(1 + \frac{1}{t^*}\int_{t=t^*}^{\infty} P\left(1+\beta\ln(t/t^*)\right)dt\right)$$

$$= e^{1/\beta}\left(1 + \frac{1}{\beta}\int_{x=1}^{\infty} P(x)e^{\frac{x-1}{\beta}}dx\right) = r(\beta)\,,$$

where the first inequality follows as P is non-increasing and as $P(\gamma_{t^*}) \leq 1$. The second equality follows by the change of variables $x = 1 + \beta\ln(t/t^*)$.

Case 2. $v > t^*$. Again, let $a_u = s$ and $a_v = 1 - s$. The second constraint of $\mathcal{F}$ now becomes $\beta s\ln(t^*/u) = 1$ and hence $u = t^* = t^*\exp(-1/s\beta)$. Moreover, the numerator of the objective of $\mathcal{F}$ is

$$\lfloor t^*\rfloor + \sum_{t=\lfloor t^*\rfloor+1}^{v} P\left(\beta s\ln(t/u)\right)$$

$$+ \sum_{t=v+1}^{\infty} P\left(s\beta\ln(t/u) + (1-s)\beta\ln(t/v)\right)\,.$$

Substituting the value of u gives $\beta s\ln(t/u) = 1 + \beta s\ln(t/t^*)$. Further, setting $v = t^*a$, the numerator simplifies to

$$\lfloor t^*\rfloor + \sum_{t=\lfloor t^*\rfloor+1}^{v} P\left(1+\beta s\ln(t/t^*)\right)$$

$$+ \sum_{t=v+1}^{\infty} P\left(1+\beta\ln(t/t^*) - (1-s)\beta\ln a\right)\,.$$

Substituting the variable $x = t/t^*$, upper bounding sums by integrals using that P is non-increasing and that $P(\gamma_{t^*}) \leq 1$, we finally get the following upper bound on the numerator

$$(3.16) \quad t^*\Big(1 + \int_1^a P(1+s\beta\ln x)dx$$
$$+ \int_a^{\infty} P(1+\beta\ln x - \beta(1-s)\ln a)dx\Big)\,.$$

Again, by the substituting the value of u and $v = t^*a$, the denominator now becomes

$$(3.17) \qquad su + (1-s)v = t^*(se^{-1/(s\beta)} + a(1-s))\,.$$

Fix s, β, and let us define $f(a)$ to be expression in (3.16) divided by t^* and $g(a)$ to be expression in (3.17) divided by t^*. Our goal is to show that $\max_{a\geq 1} f(a)/g(a) \leq r(\beta)$.

The next key claim, the proof of which is in Section A in the Appendix, shows that it suffices to consider the value of f and g at $a = 1$.

CLAIM 5. *For any* $s \in [0,1)$ *and* $\beta \geq 1$, *the ratio* $f(a)/g(a)$, *for* $a \geq 1$, *is maximized at* $a = 1$.

Now for $a = 1$, we have $g(1) = se^{-1/(s\beta)} + 1 - s$ which by simple calculus is minimized at $s = 1$ (for $s \in [0,1]$) and hence is at least $\geq e^{-1/\beta}$. Next,

$$f(1) = 1 + \int_1^{\infty} P(1+\beta\ln x)dx$$

$$= 1 + \frac{1}{\beta}\int_1^{\infty} P(y)e^{\frac{y-1}{\beta}}dy$$

$$= r(\beta)e^{-1/\beta}\,.$$

So, $f(a)/g(a) \leq r(\beta)$, which gives the required bound for Lemma 3.7. $\quad\square$

Relating the final schedule σ **to** $c_z(e)$. We now prove Lemma 3.6, that $\mathbb{E}\left[\mathrm{Cov}_\sigma(e)\right] \leq \beta c_z(e)$.

Remark: Even though GMSSC is a generalization of MSSC, the proof of Lemma 3.5 below is easier than that of Lemma 3.2 for MSSC (deferred to Section 5) as the definition of $c_z(e)$ for GMSSC has some slack, as it is not optimized to show a tight approximation ratio of 4.

Proof. (Lemma 3.6) Fix a time t, and condition on the event E_t that e is covered at time t in τ, that is, $\mathrm{Cov}_\tau(e) = t$. Recall that for all t, $\Pr[\mathrm{Cov}_\tau(e) \geq t] \leq P(z_{e<t})$ (with $P(\gamma) = 1$ for $\gamma < 1$). So to upper bound $\mathbb{E}\left[\mathrm{Cov}_\sigma(e)\right]$, we can assume that $\Pr[E_t] = P(z_{e<t}) - P(z_{e\leq t})$, and that $t \geq t_e$, as otherwise $P(z_{e\leq t}) = 1$ and hence $\Pr[E_t] = 0$.

Now conditioned on the event E_t, the cover time $\mathrm{Cov}_\sigma(e)$ of e is at most $k_e + \mathbb{E}\left[N(t)\right]$ where $N(t)$ are the vertices not in e that appear at time $\leq t$ in τ, and the delay of k_e is due to the vertices of e itself. As the rounding for vertices $u \notin e$ is independent of E_t, we have that

$$\mathbb{E}\left[N(t)\right] = \sum_{u\notin e} z_{u,\leq t} \leq \beta t - \sum_{v\in e} z_{v,\leq t}\,.$$

Let $A_e(t)$ be the set of vertices v in e with $z_{v,\leq t} \geq 1$, and $B_e(t) = e \setminus A_e(t)$. Then,

$$\sum_{v\in e} z_{v,\leq t} = \sum_{v\in A_e(t)} z_{v,\leq t} + \sum_{v\in B_e(t)} z_{v,\leq t}$$

$$\geq |A_e(t)| + k_e - |A_e(t)|$$

$$= k_e\,,$$

where we use (3.15) and that $z_{e,\leq t} \geq 1$ for $t \geq t_e$. This gives that,

$$\mathbb{E}\left[\mathrm{Cov}_\sigma(e)\right] \leq \sum_t \Pr[E_t](k_e + \mathbb{E}\left[N(t)\right])$$

$$\leq \Pr[E_t](k_e + (\beta t - k_e))$$

$$= \beta\sum_t \Pr[E_t]\,t\,.$$

As $\sum_t \Pr[E_t]\, t = \sum_t t(P(z_{e<t}) - P(z_{e\leq t})) = \sum_t P(z_{e<t}) = c_z(e)$, the result follows. □

4 Min Sum Vertex Cover

Now we consider MSVC, a special case of GMSSC when $|e| = 2$ and $k(e) = 1$. We will apply the same framework as before, but with a different kernel K and use it to obtain a $16/9 = 1.777\ldots$ approximation.

We use the same LP as that for MSSC, where now constraint (2.3) can be written more explicitly as

$$u_{e,t} \geq 1 - (x_{v,<t} + x_{w,<t}) \qquad \forall e = (v,w).$$

We will also show a matching $16/9$ integrality gap for this LP.

4.1 Approximation Result We apply our generic Algorithm 1 with the following kernel,

$$K(t,t') = \frac{4t'(t'+1)}{t(t+1)(t+2)} \cdot \mathbb{1}[t \geq t'].$$

and show the following

THEOREM 4.1. *The algorithm, with the above kernel, is a $16/9$-approximation for MSVC.*

This kernel should be viewed intuitively as $K(t,t') = 4t'^2/t^3$, but as time is discrete in our formulations, we use the version above. This kernel is qualitatively very different from the one we used previously for MSSC. The idea is that in the vertex cover problem, for any edge $e = (v,w)$ at least one of v, w (say v) will eventually be picked to an extent of at least $1/2$. So we only need a kernel that ensures that v has a mass of 1 in the solution z as $x_{v,\leq t}$ approaches $1/2$ for t large enough. This allows us to pick a kernel where $K(t,t')$ decays more rapidly with t (in particular, as $1/t^3$ here as opposed to $1/t$ in MSSC). In particular, if $x_{v,t'} = 1/2$ for some t' (and 0 elsewhere), then using (4.18) below,

$$\sum_t z_{v,t} = \sum_{t \geq t'} K(t,t')\frac{1}{2} = \sum_{t \geq t'} \frac{2t'(t'+1)}{t(t+1)(t+2)} = 1.$$

Of course, there are several candidate kernel functions that would satisfy this property, and our particular choice was based on considering how an optimum rounding should look like on the worst case integrality gap example described in Section 4.2. We now give the details.

We first state a couple of standard identities that we will use repeatedly below.

$$\sum_{t=a}^{b} t(t+1) = \frac{1}{3}\Big(b(b+1)(b+2) - a(a-1)(a+1)\Big).$$

Similarly, writing $1/t(t+1) = 1/t - 1/(t+1)$ and $2/t(t+1)(t+2) = 1/t - 2/(t+1) - 1/(t+2)$,

$$(4.18)\quad \sum_{t=a}^{b} \frac{2}{t(t+1)(t+2)} = \frac{1}{a(a+1)} - \frac{1}{(b+1)(b+2)}.$$

As usual, to prove Theorem 4.1, it suffices to show the following two lemmas.

LEMMA 4.1. *For any edge e, $\mathbb{E}[Cov_\tau(e)] \leq (4/3)c_x(e)$.*

LEMMA 4.2. *For any edge e, $\mathbb{E}[Cov_\sigma(e)] \leq (4/3)\mathbb{E}[Cov_\tau(e)]$.*

The proof of Lemma 4.2 is given in Section 5, and relies on the fact below that the total z *load* of all the vertices at any time t is at most $4/3$.

LEMMA 4.3. $\sum_v z_{v,t} \leq 4/3$, *for any t.*

Proof. By constraint (2.2), $\sum_v x_v \leq \mathbf{1}$, and hence $\sum_v z_v = \sum_v Kx_v \leq K\mathbf{1}$. Now,

$$(K\mathbf{1})_t = \sum_{t'} K(t,t') = \sum_{t' \leq t} \frac{4t'(t'+1)}{t(t+1)(t+2)}$$
$$= \frac{4}{t(t+1)(t+2)} \cdot \frac{t(t+1)(t+2)}{3} = 4/3.$$

□

We now focus here on proving Lemma 4.1.

Proof. [Proof of Lemma 4.1] Fix an edge $e = (v,w)$. Let $p_{v,t}$ denote the probability that v is not covered before time t in the tentative schedule τ. Then, the expected cover time of e in τ is $\mathbb{E}[Cov_e(\tau)] = \sum_{t \geq 1} p_{v,t} \cdot p_{w,t}$. To compute $p_{v,t}$, note that the v is assigned to a slot $< t$ by α-point rounding with probability $\min\{1, z_{v,<t}\}$, and

$$z_{v,<t} = \sum_{t'<t} z_{v,t'} = \sum_{t'<t}\sum_{t''\leq t'} K(t',t'')x_{v,t''}$$
$$= \sum_{t''<t} x_{v,t''} \sum_{t'=t''}^{t-1} K(t',t'')$$
$$= \sum_{t''<t} 2\left(1 - \frac{t''(t''+1)}{t(t+1)}\right) x_{v,t''}$$

where we use that by (4.18),

$$\sum_{q=t''}^{t-1} K(q,t'') = 4t''(t''+1)\sum_{q=t''}^{t-1} \frac{1}{q(q+1)(q+2)}$$
$$= 2\left(1 - \frac{t''(t''+1)}{t(t+1)}\right).$$

This gives that,

$$p_{v,t} = 1 - \min\{1, z_{v,<t}\}$$

(4.19)
$$= \left(1 - \sum_{t'<t} 2\left(1 - \frac{t'(t'+1)}{t(t+1)}\right)x_{v,t'}\right)_+ .$$

Next, the LP cost $c_x(e)$ for e is exactly $\sum_t \left(1 - \sum_{t'<t}(x_{v,t'} + x_{w,t'})\right)_+$. So our goal is to show that for any setting of $x_{v,t}$ and $x_{w,t}$,

(4.20) $$\sum_t p_{v,t} \cdot p_{w,t} \le \frac{4}{3}\sum_t \left(1 - \sum_{t'<t}(x_{v,t'} + x_{w,t'})\right)_+$$

where $p_{v,t}$ and $p_{w,t}$ are given by (4.19).

To this end, we first make some simple observations. First, we can assume that $x_{v,<t} + x_{w,<t} \le 1$ for all t as this does not affect the right side of (4.20), and can only increase the left hand side. Next, replacing each $x_{v,t}$ and $x_{w,t}$ by their average value $(x_{v,t} + x_{w,t})/2$ does not affect the right side, and can only increase the left side as by AM-GM inequality, $(1-a)(1-b) \le (1-(a+b)/2)^2$, whenever $a, b \in [0,1]$. Thus proving (4.20) reduces to showing the following.

Given a vector a with non-negative entries (where a_t corresponds to $x_{v,t} + x_{w,t}$) and $\|a\|_1 = 1$,

$$\sum_t \left(1 - \sum_{t'<t}\left(1 - \frac{t'(t'+1)}{t(t+1)}\right)a_{t'}\right)^2$$
$$\le \frac{4}{3}\sum_t \left(1 - \sum_{t'<t} a_{t'}\right).$$

Let $f(a)$ denote the left hand side above, and $g(a)$ denote the right side, then we get the following problem.

$(\mathcal{F})$ Maximize $f(a)/g(a)$
            Subject to $\|a\|_1 = 1$, $a_t \ge 0$.

Again, $g(a)$ is linear in a and strictly positive over the positive simplex. For a fixed t, each summand in $f(a)$ is a term of the form $(1 - c \cdot a)^2$, where c is a vector with $\|c\|_\infty \le 1$, implying that $|c \cdot x| \le 1$ and that the term is convex. Thus $f(a)$ is convex. So by Fact 3.1, the optimum is attained at some extreme point of the simplex of the form $a_u = 1$ and $a_{u'} = 0$ for $u' \ne u$.

In this case, $g(a) = 4u/3$ and

$$f(a) = u + \sum_{t>u}\left(\frac{u(u+1)}{t(t+1)}\right)^2$$
$$\le u + \sum_{t>u}\frac{u^2(u+1)^2}{(t-1)t(t+1)(t+2)}$$
$$= u + \frac{u^2(u+1)^2}{3u(u+1)(u+2)}$$
$$\le \frac{4}{3} \cdot u,$$

which proves the desired inequality.   □

4.2 Tight Integrality Gap We now show that the integrality gap of the LP for MSVC is arbitrarily close to 16/9.

Consider an instance consisting of disjoint copies of the complete graphs K_i, for $i = 1, \ldots, k$, where K_i has n_i vertices. We will set $n_i = Ni^{-\alpha}$ where $\alpha = \frac{2}{3} + \epsilon$, with ϵ approaching 0, and for a suitable $k = \omega(1)$. Also choose N large enough (so the issues of floor and ceiling in n_i will not affect us). Note that $n_1 > n_2 \ldots > n_k$. To show the integrality gap, we will upper bound the LP cost suitably, and show that any integral solution has cost at least $16/9 - O_\epsilon(1)$ times the upper bound on the LP cost.

LEMMA 4.4. *The LP cost is at most $N^3/(4\epsilon)$, up to lower order terms.*

Proof. Consider the fractional solution that first schedules each vertex of K_1 to extent $1/n_1$ in the first $n_1/2$ slots, then schedules the vertices of K_2 in the next $n_2/2$ slots and so on. Then, each edge in K_i is completely covered by time $(n_1 + \ldots + n_i)/2$, and the cost of this solution is at most

$$\sum_{i=1}^k \binom{n_i}{2}\frac{n_1 + \ldots n_i}{2} \le \sum_{i=1}^k \frac{N^3}{4}i^{-2\alpha}\sum_{j=1}^i j^{-\alpha}$$
$$\le \sum_{i=1}^k \frac{N^3}{4}i^{-2\alpha}\frac{i^{1-\alpha}}{1-\alpha}$$
$$= \frac{N^3}{4(1-\alpha)}\sum_{i=1}^k i^{1-3\alpha},$$

where we used that $\sum_{j=1}^i j^{-\alpha} \le \int_{j=0}^i x^{-\alpha}dx = i^{1-\alpha}/(1-\alpha)$, as $\alpha < 1$. Moreover, for $\alpha = \frac{2}{3} + \epsilon < 1$, as ϵ gets arbitrarily small

$$\sum_{i=1}^k i^{1-3\alpha} = \sum_{i=1}^k i^{-1-3\epsilon} \le 1 + \int_1^\infty x^{-1-3\epsilon}dx = 1 + \frac{1}{3\epsilon}.$$

For ϵ small and k large enough, the LP cost is bounded by $N^3(1+3\epsilon)/4\epsilon(1-3\epsilon) \approx N^3/4\epsilon$.   □

LEMMA 4.5. *Any integral solution has cost at least $4N^3/(9\epsilon)$, up to lower order terms.*

Proof. We will measure the cost of an integral solution as the sum over time t of the number of uncovered edges at t. In any clique K, if x vertices remain then there are exactly $x(x-1)/2 \approx x^2/2$ uncovered edges. Consider the greedy algorithm, that at any time picks a vertex that covers the most number of edges. For our instance, it is easily verified that this in fact minimizes the total number

of remaining uncovered edges at every time t, and hence is clearly the optimum solution. The algorithm thus proceeds by first covering vertices in K_1, until n_2 vertices are left in K_1, then alternately picking vertices from K_1 and K_2 until n_3 vertices are left in both, then alternating between K_1, K_2, K_3 until n_4 vertices are left and so on.

For $i \in [k]$, let t_i denote the first time when n_i vertices are left in each of $K_1, \ldots, K_i$. Then $t_1 = 0$ and $t_{i+1} - t_i = i(n_i - n_{i+1})$ as the sizes of $K_1, \ldots, K_i$ must shrink from n_i to n_{i+1} during this time.

During t_i to t_{i+1}, the number of uncovered edges is at least $i \cdot \binom{n_{i+1}}{2} + \sum_{j=i+1}^{k} \binom{n_j}{2}$, where the first contribution is due to all the edges between the n_{i+1} unpicked vertices in each of the first i cliques $K_1, \ldots, K_i$, and the latter due to all the edges in the remaining cliques $K_{i+1}, \ldots, K_k$. So the objective is at least

(4.21)
$$\sum_{i=1}^{k-1} (t_{i+1} - t_i) \left(i \cdot \binom{n_{i+1}}{2} + \sum_{j=i+1}^{k} \binom{n_j}{2} \right)$$
$$= \sum_{i=1}^{k-1} i(n_i - n_{i+1}) \left(i \cdot \binom{n_{i+1}}{2} + \sum_{j=i+1}^{k} \binom{n_j}{2} \right).$$

Note that we are even ignoring the contribution due to edges covered after time t_k.

Replacing $\binom{n_i}{2} \approx \frac{1}{2} n_i^2$ we have

$$i \cdot \binom{n_{i+1}}{2} + \sum_{j=i+1}^{k} \binom{n_j}{2} \approx \frac{N^2}{2} \left(i^{1-2\alpha} + \int_i^k j^{-2\alpha} dj \right)$$
$$= \frac{N^2}{2} \left(i^{1-2\alpha} + \frac{1}{2\alpha-1} \left(i^{1-2\alpha} - k^{1-2\alpha} \right) \right)$$

Substituting this in (4.21) and writing $n_i - n_{i+1} \approx N\alpha i^{-1-\alpha}$ gives

$$\frac{N^3}{2} \sum_{i=1}^{k} i \cdot \alpha i^{-1-\alpha} \left(\frac{2\alpha}{2\alpha-1} i^{1-2\alpha} - \frac{1}{2\alpha-1} k^{1-2\alpha} \right)$$
$$= \frac{N^3}{2} \left(\sum_{i=1}^{k} \frac{2\alpha^2}{2\alpha-1} i^{1-3\alpha} - \frac{\alpha k^{1-2\alpha}}{2\alpha-1} \sum_{i=1}^{k} i^{-\alpha} \right)$$

As $\alpha > 2/3$, as k gets large, the first term converges to $N^3\alpha^2/(2\alpha-1)(3\alpha-2) \approx 4N^3/(9\epsilon)$. Next, as $\alpha < 1$, the second term is about $\frac{\alpha}{2\alpha-1} \frac{k^{1-\alpha}}{1-\alpha} k^{1-2\alpha} \cdot (N^3/2) = O(k^{2-3\alpha}) \cdot (N^3/2) = o(N^3)$ for $k = \omega(1)$. So any integral solution has cost at least $4N^3/(9\epsilon)$, up to lower order terms. □

Lemmas 4.4 and 4.5 give an integrality gap of $(4/9)/(1/4) = 16/9$.

5 Conditioning for MSSC and MSVC

We now prove Lemmas 3.2 and 4.2, that relate the expected cover time for an edge e in σ to $c_z(e)$ for MSSC and MSVC. The conditioning is more subtle here, and handling it needs more care, and crucially uses the randomness in the tie-breaking rule (i.e., line 7, Algorithm 1).

Fix some time t, and an edge e. Let E_t denote the event that $\text{Cov}_\tau(e) = t$. Recall that, given a solution z, we defined $c_z(e) = \sum_t p_t(e)$, where $p_t(e)$ was an upper bound on the probability that e is uncovered just before t in τ, i.e. $\Pr[\text{Cov}_\tau(e) \geq t] \leq p_t(e)$.

Let us condition on the event E_t. Then the expected cover time of e in σ is 1 plus the delay due to vertices not in e. Among the vertices not in e, let $A_{<t}$ be those scheduled before t in τ, and A_t be those scheduled at t. If k vertices of e are scheduled at t, then as the ties at time t are broken at random, the expected delay of e due to vertices in A_t is exactly $|A_t|/(k+1) \leq |A_t|/2$ (as $k \geq 1$ conditioned on E_t).

So the expected cover time of e in σ conditioned on E_t is at most $1 + \mathbb{E}[|A_{<t}|] + \mathbb{E}[|A_t|/2]$. Now, any $v \notin e$ is scheduled before t in τ with probability $z_{v,<t}$, and this is independent of E_t. So,

$$\mathbb{E}[|A_{<t}|] = \sum_{v \notin e} z_{v,<t} = \beta(t-1) - \sum_{v \in e} z_{v,<t} \leq \beta(t-1) - z_{e,<t}.$$

Here we used $z_{e,t} \leq \sum_{v \in e} z_{v,t}$ for MSSC and MSVC. Similarly, $\mathbb{E}[|A_t|] \leq \beta - z_{e,t}$. Hence, $\mathbb{E}[\text{Cov}_\sigma(e)]$ is at most

$$\sum_t \Pr[E_t] \left(1 + \beta(t-1) - z_{e,<t} + (\beta - z_{e,t})/2 \right)$$
$$\leq \sum_t (p_t(e) - p_{t+1}(e))(1 + \beta(t-1) - z_{e,<t}$$
(5.22) $$+ (\beta - z_{e,t})/2),$$

where we use that as $(\beta(t-1) - z_{e,<t} + (\beta - z_{e,t})/2)$ is non-decreasing in t and $\Pr[\text{Cov}_\tau(e) \geq t] \leq p_t(e)$, replacing $\Pr[E_t]$ by $p_t(e) - p_{t+1}(e)$ can only increase the right hand side.

As $c_z(e) = \sum_t p_t(e) = \sum_t t(p_t(e) - p_{t+1}(e))$, to prove $\mathbb{E}[\text{Cov}_\sigma(e)] \leq \beta c_z(e)$, it suffices to show that the expression in (5.22) is at most $\beta \sum_t t(p_t(e) - p_{t+1}(e))$.

Using $\sum_t (p_t(e) - p_{t+1}(e)) = 1$ and simplifying, this reduces to showing that

$$1 - \beta/2 \leq \sum_t (p_t(e) - p_{t+1}(e))(z_{e,<t} + z_{e,t}/2).$$

As $\beta \geq 1$, the left side is at most $1/2$, and the right hand side simplifies to $\sum_t z_{e,t}(p_t(e) + p_{t+1}(e))/2$. So our goal is to show that

$$\sum_t z_{e,t}(p_t(e) + p_{t+1}(e)) \geq 1.$$

We now do this separately for MSSC and MSVC.

MSSC. In the definition of $c_z(e)$, we defined $p_t(e) = \exp(-z_{e,<t})$, so our goal is to show that

$$(5.23) \qquad \sum_t \big(\exp(-z_{e,<t}) + \exp(-z_{e,<t+1})\big) z_{e,t} \geq 1.$$

Now, for any convex function f and any two points a and b it holds that $(f(a) + f(b))(b - a) \geq 2 \int_a^b f(u)du$. Setting $f(u) = \exp(-u)$ and $a = z_{e,<t}$ and $b = z_{e,<t+1}$ for each t, the left hand side in (5.23) least

$$2 \sum_t \int_{z_{e,<t}}^{z_{e,t+1}} \exp(-w)dw$$
$$= 2 \exp(-z_{e,<1}) - 2 \exp(-z_{e,<\infty}) = 2 > 1,$$

which proves the desired result for MSSC (Lemma 3.2).

MSVC. Here, in the definition of $c_z(e)$ we defined $p_t(e) = (1 - z_{e,<t}/2)^2$, which is convex for $z_{e,<t} \in [0,2]$. For MSVC, e is surely covered if $z_{e,<t} \geq 2$, and hence $p_t(e) = 0$ for such t. So we only need to consider t, where $z_{e,<t} \leq 2$. So, following the argument above,

$$\sum_t z_{e,t}(p_t(e) + p_{t+1}(e)) \geq 2 \sum_t \int_{z_{e,<t}}^{z_{e,<t+1}} (1 - w/2)^2 dw$$
$$= 2 \int_0^2 (1 - w/2)^2 dw = \frac{4}{3} > 1.$$

This proves the desired inequality for MSVC (Lemma 4.2).

6 A Tail Bound for Sum of Bernoulli R.V.s

In this section we prove Theorem 1.5. Recall that we are going to upper bound $\Pr[S \leq k - 1]$ where S is the sum of n independent Bernoulli random variables, having $\mathbb{E}[S] = k\gamma$ for $\gamma \geq 1$ and integer $k \geq 1$. We first show a weaker bound of $e^{-\gamma}(e/2)$ to illustrate the main ideas, and the full proof of Theorem 1.5 is deferred to Appendix B.

The first step is to use a classic result of Hoeffding (Lemma 6.1 below) that allows to assume that the Bernoulli variables are identical. We then use this Lemma again, to show that increasing their number n, can only increase $\Pr[S \leq k - 1]$, that we need to upper bound. This allows us to reduce the problem to bounding the left tail of a Poisson distribution. In Theorem 6.1 we show that this problem can be reduced further to the case of $\gamma = 1$. Finally, for $\gamma = 1$, the result follows a well-known relation between the median and the mean for Poisson random variables. We now give the details.

The following lemma of Hoeffding allows us to reduce the space of distributions of independent and *identical* Bernoulli random variables, i.e., a Binomial distribution

$B(n, p)$, for which we denote the cumulative distribution function (CDF) by $F(k; n, p) \stackrel{\text{def}}{=} \sum_{i=0}^k \binom{n}{i} p^i (1 - p)^{n-i}$, for $k \in Z^+$.

LEMMA 6.1. ([18], THEOREM 4) *Among all choices of p_i with $\sum_i p_i = \mu$ fixed, for $k \leq \mu$, $\Pr[S \leq k-1]$ is maximized when all the p_i's are identical, equal to $p = \mu/n$. That is,*

$$\Pr[S \leq k - 1] \leq F(k - 1; n, p) = \sum_{i=0}^{k-1} \binom{n}{i} p^i (1 - p)^{n-i}.$$

Now we apply the above Hoeffding's lemma again to reduce our problem to bounding CDF of a Poisson distribution.

LEMMA 6.2. *For $\lambda = np$, $F(k - 1; n, p) \leq e^{-\lambda} \sum_{i=0}^{k-1} \frac{\lambda^i}{i!}$.*

Proof. Define $S' \stackrel{\text{def}}{=} \sum_{i \in [2n]} X_i'$ where

$$X_i' \sim \begin{cases} Bernoulli(p), & i \in [n] \\ Bernoulli(0), & i \in [2n] \setminus [n] \end{cases} .$$

Given that $\mathbb{E}[S'] = np$, the previous lemma gives $\Pr[S' \leq k - 1] \leq F(k - 1; 2n, p/2)$. On the other hand, $\Pr[S' \leq k - 1] = F(k - 1; n, p)$. Thus we have $F(k - 1; n, p) \leq F(k - 1; 2n, p/2)$, and more generally, $F(k - 1; n, p) \leq \lim_{\alpha \to \infty} F(k - 1; \alpha n, p/\alpha)$. Now the desired bound in the lemma follows from the convergence of the binomial distribution to the Poisson distribution, with $\lambda = np$, as $n \to \infty$. $\square$

The case of $\gamma = 1$. Recall that $\gamma = np/k = \lambda/k$. For $\gamma = 1$ (i.e., $\lambda = k$), applying Lemmas 6.1 and 6.2, the proof of Theorem 1.5 reduces to showing that

$$(6.24) \qquad e^{-k} \sum_{i=0}^{k-1} k^i/i! \leq e^{-1}(e/2) = 1/2 .$$

The left hand side is the probability that a Poisson random variable with mean k has value at most $k - 1$, which follows directly from the fact that the median of the Poisson distribution with mean λ is strictly larger than $\lambda - 1$; see [22] for a precise estimate.

We prove the (weaker) bound for $\gamma > 1$. The stronger bound in Theorem 1.5 is deferred to Appendix B. The idea of the proof is to reduce the case of $\gamma > 1$ to that of $\gamma = 1$, by a certain inductive argument on k.

THEOREM 6.1. *Let $S = \sum_i X_i$, where $\{X_i \sim Bernoulli(p_i) : i \in [n]\}$ are independent Bernoulli random variables, with $p_i \in [0, 1]$. Let $k \leq \mathbb{E}[S] = \lambda$ be a positive integer. For $\gamma = \mathbb{E}[S]/k \geq 1$, we have*

$$(6.25) \qquad \Pr[S \leq k - 1] \leq e^{-\lambda} \sum_{i=0}^{k-1} \frac{\lambda^i}{i!} \leq e^{-\gamma}(e/2) .$$

Proof. Let $g(\gamma) := e^{-\gamma}(e/2)$ denote the desired upper bound. Define

$$f_k(\lambda) \stackrel{\text{def}}{=} e^{\lambda} g(\lambda/k) - \sum_{i=0}^{k-1} \frac{\lambda^i}{i!}.$$

With this notation, the theorem is equivalent to showing that $f_k(\lambda) \geq 0$ for $\lambda \geq k$. We prove this by induction on $k \geq 1$. For $k = 1$, the basis, this holds trivially as $f_1(\lambda) = e^{\lambda} g(\lambda) - 1 = e/2 - 1 \geq 0$.

Fix $k \geq 2$. For $\lambda = k$ we previously showed that $f_k(k) \geq 0$ (which is equivalent to (6.24)), so it suffices to prove $f_k(\cdot)$ is a non-decreasing function, i.e., $\frac{d}{d\lambda} f_k(\lambda) \geq 0$, for $\lambda \geq k$. By the inductive hypothesis, we have $f_{k-1}(\lambda) \geq 0$, for all $\lambda \geq k-1$; that is,

$$(6.26) \qquad \sum_{i=0}^{k-2} \frac{\lambda^i}{i!} \leq e^{\lambda} g(\lambda/(k-1)).$$

For the induction step we have

$$\frac{d}{d\lambda} f_k(\lambda) = e^{\lambda}\left(\frac{1}{k} g'(\lambda/k) + g(\lambda/k)\right) - \sum_{i=1}^{k-1} \frac{i\lambda^{i-1}}{i!}$$

$$= e^{\lambda}\left(\frac{1}{k} g'(\lambda/k) + g(\lambda/k)\right) - \sum_{i=0}^{k-2} \frac{\lambda^i}{i!}$$

$$\geq e^{\lambda}\left(\frac{1}{k} g'(\lambda/k) + g(\lambda/k) - g(\lambda/k-1)\right)$$

where the last inequality uses (6.26). Substituting $\gamma = \lambda/k$ it suffices to show for all $k \geq 2, \gamma \geq 1$

$$(6.27) \qquad \frac{g'(\gamma)}{k} + g(\gamma) - g\left(\frac{\gamma k}{k-1}\right) \geq 0$$

For $g(\gamma) = \exp(1-\gamma)/2$, the above becomes

$$-\frac{\exp(-\gamma)}{k} + \exp(-\gamma) - \exp\left(-\frac{\gamma k}{k-1}\right) \geq 0$$

The proof now follows by re-writing the LHS of the above as

$$\frac{k-1}{k} \exp\left(-\frac{\gamma k}{k-1}\right)\left(\exp\left(\frac{\gamma}{k-1}\right) - \frac{k}{k-1}\right) \geq 0$$

where the last inequality is due to $\gamma \geq 1$ and $\exp(x) \geq 1+x$ for all x. $\quad\square$

7 Min ℓ_p Set Cover

We now consider the problem of minimizing the ℓ_p norm of cover times for MSSC. We refer to this as MSSC_p. We first sketch how the previous approach gives a $\delta_p = (p+1)^{1+1/p}$ approximation. Then we focus on the greedy algorithm and show that it simultaneously provides this guarantee for every $p \geq 1$. We then show a hardness result that no better approximation exists for any $p \geq 1$, assuming P$\neq$NP.

7.1 A Randomized Rounding Algorithm
Consider the the natural LP relaxation for MSSC_p, which is similar to the LP for MSSC but with a different objective,

$$\sum_t (t^p - (t-1)^p) u_{e,t},$$

i.e., the cost is the pth power of the ℓ_p norm of the cover times. We show the following.

THEOREM 7.1. *The rounding algorithm in Section 3 with $\beta = p+1$ for the above LP is a δ_p approximation for MSSC_p.*

Proof. The analysis is similar to that of Theorem 3.1 for MSSC. Let us define

$$c_z(e) = \sum_t (t^p - (t-1)^p) \exp\left(-\sum_{v \in e} z_{v,<t}\right),$$

amd

$$c_x(e) = \sum_t t^p x_{e,t},$$

where $c_z(e)$ is an upper bound on the expected cost of e in the tentative schedule, and $c_x(e)$ is the contribution in the LP. We first show that $c_z(e) \leq (p+1)c_x(e)$ and then relate $c_z(e)$ to the expected cost in the actual schedule.

Similar to Theorem 3.1, and using the expression for $c_z(e)$ and $c_x(e)$ above, it suffices to upper bound the value of the following optimization problem in the variables a_t.

$$\text{Maximize} \quad \frac{\sum_t (t^p - (t-1)^p) e^{-(p+1)\sum_{t'<t}(a_{t'} \ln \frac{t}{t'})}}{\sum_t t^p a_t}$$

$$\text{Subject to} \quad \|a\|_1 = 1, \qquad a_t \geq 0.$$

Exactly as in Claim 3, the optimum is at an extreme point supported on a_u for some u. To compute the objective value at this point, note that the denominator is u^p and the numerator is

$$\sum_{t=1}^{u} (t^p - (t-1)^p) + \sum_{t>u} (t^p - (t-1)^p) e^{-(p+1)\ln t/u}$$

$$\leq u^p + \sum_{t>u} pt^{p-1}(u/t)^{p+1}$$

$$\leq u^p + pu^{p+1}\int_{x=u}^{\infty} \frac{1}{x^2} dx$$

$$= u^p(1+p).$$

Hence, $c_z(e) \leq (p+1)c_x(e)$. As $\beta = p+1$, converting the tentative schedule to an actual schedule will increase the expected cover time by at most $p+1$, and its p-th power by at most $(p+1)^p$. By an argument similar to that in Section 5 to account for conditioning, the expected cost of the final solution is at most $(p+1)^p$ times $\sum_e c_z(e)$ which is at most $(p+1)^{p+1} \sum_e c_x(e)$. This gives the claimed δ_p approximation. $\quad\square$

7.2 Bound for the Greedy Algorithm The greedy algorithm, at every time-slot $t = 1, 2, 3, \ldots$ schedules a vertex v that covers the most number of currently uncovered edges. Let g denote the cost of the algorithm and OPT denote the optimum value.

THEOREM 7.2. *The greedy algorithm gives a δ_p-approximation for MSSC$_p$, for every $p \geq 1$.*

We prove this using a dual fitting argument. Consider the natural LP formulation for minimizing the p-th power of the ℓ_p norm. We use continuous time for convenience (note that this can only reduce the LP objective as compared with discrete time).

$$(\text{LP}_p) \quad \text{Minimize} \quad \sum_e \int_0^\infty pt^{p-1} u_{e,t} dt$$

$$\text{Subject to} \quad \sum_v y_{v,t} \leq t \quad \forall\, t,$$

$$u_{e,t} + \sum_{v \in e} y_{v,t} \geq 1 \quad \forall\, e, t,$$

$$u_{e,t}, y_{v,t} \geq 0.$$

Here the variable $y_{v,t}$ indicates whether a vertex is scheduled by time t and $u_{e,t}$ indicates whether an edge e is not covered by time t. The objective function for an edge e now becomes $\int_0^\infty pt^{p-1} u_{e,t} dt$ (which is t_e^p if the edge e is covered at time t_e in an integral schedule).

Consider a new LP, namely LPS$_p$, obtained by replacing the inequality $\sum_v y_{v,t} \leq t$ in LP$_p$ by

$$\sum_v y_{v,t} \leq t/(p+1) \quad \forall\, t,$$

which can be interpreted as a simple time-scaling. We have the following simple relation.

LEMMA 7.1. $Opt(LPS_p) = (p+1)^p \cdot Opt(LP_p)$.

Proof. Given an optimum solution $u_{e,t}, y_{v,t}$ to LP$_p$, let $y'_{v,t} = y_{v,t/(p+1)}$ and $u'_{e,t} = u_{e,t/(p+1)}$. It is easy to verify that $u'_{e,t}, y'_{v,t}$ is a feasible solution for LPS$_p$. It suffices to upper bound the cost of LPS$_p$ for (u', y') by $(p+1)^p$ times the cost of LP_p for (u, y). Indeed,

$$\sum_e \int_0^\infty pt^{p-1} u'_{e,t} dt = \sum_e \int_0^\infty pt^{p-1} u_{e,t/(p+1)} dt$$

$$= (p+1)^p \sum_e \int_0^\infty pt^{p-1} u_{e,t} dt.$$

by a simple change of variable in the integration. $\quad\square$

Let us consider the dual to LPS$_p$, obtained using variables α_t (resp. $\beta_{t,e}$) for the first (resp. second) set of constraints. The dual is as follows.

$$(\text{DLPS}_p) \quad \text{Max} \quad \sum_e \int_0^\infty \beta_{e,t} dt - \frac{1}{p+1} \int_0^\infty t \cdot \alpha_t dt$$

$$\text{s.t.} \quad \sum_{e \ni v} \beta_{e,t} \leq \alpha_t \quad \forall\, v, t,$$

$$0 \leq \beta_{e,t} \leq pt^{p-1} \,\forall\, e, t, \quad \alpha_t \geq 0 \quad \forall\, t.$$

In the following key lemma, we use the execution of the greedy algorithm to construct a feasible dual solution.

LEMMA 7.2. *There is a feasible solution to the dual LP above with cost at least $g^p/(p+1)$.*

Proof. Consider the execution of the greedy algorithm. Note that greedy schedules at discrete time steps, while the LP above is for continuous time. Let us denote the set of edges not covered at the beginning of time-slot k by R_k, and the edges with cover time k (i.e., $R_{k+1} \setminus R_k$) by X_k respectively. In other words, X_k is exactly the set of edges covered at time k by the greedy algorithm.

Dual Feasbility. Consider the solution $\alpha_t = pt^{p-1} |X_{\lceil t \rceil}|$ and $\beta_{e,t} = pt^{p-1} \mathbb{1}[e \in R_{\lceil t \rceil}]$, for any time $t \in \mathbb{R}^+$. We claim that this is feasible. Feasibility of the dual constraint $\beta_{e,t} \leq pt^{p-1}$ follows directly by construction. Let us consider the other constraint $\sum_{e \ni v} \beta_{e,t} \leq \alpha_t$ for some vertex v and time t. Then, we claim that

$$\sum_{e \ni v} \beta_{e,t} = pt^{p-1} \sum_{e \ni v} \mathbb{1}[e \in R_{\lceil t \rceil}] \leq pt^{p-1} |X_{\lceil t \rceil}| = \alpha_t.$$

Here the inequality uses the key property of the greedy algorithm that among all vertices v, and at all times t, the vertex chosen at time $\lceil t \rceil$ by the greedy algorithm is the one that maximizes $\sum_{e \ni v} \mathbb{1}[e \in R_{\lceil t \rceil}]$.

Cost analysis. We now bound the cost of the dual solution. Let t_e denote the cover time of an edge e under the greedy algorithm. Then we have the following.

$$\sum_e \int_0^\infty \beta_{e,t} dt = \sum_e \int_0^{t_e} pt^{p-1} dt = \sum_e t_e^p = g^p,$$

$$\int_0^\infty t\alpha_t dt = \sum_{i=1}^\infty \int_{i-1}^i t\alpha_t dt = \sum_{i=1}^\infty |X_i| \int_{i-1}^i t \cdot pt^{p-1} dt$$

$$\leq \sum_{i=1}^\infty |X_i| pi^p = p \sum_e t_e^p = p \cdot g^p.$$

The last two steps follow as the greedy objective can be written in two different ways as $\sum_i |X_i| i^p$ and $\sum_e t_e^p$.

Together, these give that the dual cost is at least $g^p - \frac{p}{p+1} g^p = g^p/(p+1)$ as desired. $\quad\square$

By Lemmas 7.2 and 7.1, we have that

$$g^p/(p+1) \leq \text{Opt}(\text{DLPS}_p) \leq \text{Opt}(\text{LPS}_p)$$
$$\leq (p+1)^p \cdot \text{Opt}(\text{LP}_p) \leq (p+1)^p \cdot \text{OPT}^p \,,$$

where the second inequality is by weak duality. Hence, $g^p \leq (p+1)^{p+1} \cdot \text{OPT}^p$ completing the proof.

7.3 Hardness of Approximation We now show the following hardness result.

THEOREM 7.3. *It is NP-hard to approximate MSSC$_p$ within a ratio of $\delta_p - \epsilon$, for any $\epsilon > 0$ and $p \geq 1$.*

Our approach is based on that of [11] to show the $4 - \epsilon$ hardness of MSSC, with some change in parameters, and we sketch the argument here, focusing only on the key differences. The starting point is the following result of [10, 11].

THEOREM 7.4. *Consider the problem of picking vertices to cover edges in a regular uniform hypergraph. For any $c_0, \epsilon > 0$, it is NP-hard to distinguish between the following.*

Case I: All edges can be covered by an independent set of vertices, of size at most r.

Case II: For any $c \leq c_0 r$, every subset of c vertices covers at most $1 - (1 - 1/r)^c + \epsilon$ fraction of edges.

Let H be a regular uniform hypergraph with m edges. Fix a large constant k and construct k disjoint copies of H, indexing them as $1, 2, \ldots, k$. Make $a/(i^{p+1})$ copies of every edge in the ith copy of H where $a = (k!)^{p+1}$, denote the resulting hypergraph by H_i. Consider the instance $G = \cup_i H_i$.

Let us call the cost of an edge to be the pth power of its cover time. Then, it suffices to show that the total cost for G in case II is at least $(p+1)^{p+1}$ times the cost in case I.

Assuming Case I for H. Let $F(i)$ be the contribution of H_i to the cost. $F(i)$ can be upper bounded by noting that each additional vertex scheduled from H_i covers $ma/(ri^{p+1})$ edges of H_i.

$$F(i) \leq \frac{a}{i^{p+1}} \sum_{j=1}^{r} \left(\frac{m}{r}\right) ((i-1)r + j)^p$$
$$\leq \frac{ma}{ri^{p+1}} r(ir)^p = \frac{mar^p}{i} \,.$$

Hence, $\sum_i F(i) \lessapprox mar^p \log k$.

Assuming Case II for H. We know that no matter which $c < c_0 r$ vertices in H_i are picked, at least $(1 - 1/r)^c$ of the edges remain uncovered. Let $m_i = ma/(i^{p+1})$ be the number of edges in H_i. Now as in [11], the best way to

cover G is to first schedule vertices from H_1 until $\leq m_2$ edges remain uncovered in H_1, then schedule vertices from H_1 and H_2 alternately until $\leq m_3$ edges remain in H_1, H_2, and so on. It can be verified that in our case, this minimizes the total number of remaining uncovered edges at every time step, and hence is the optimum solution.

Let t_i denote the first time when m_i edges are left to be covered in $H_1, \ldots, H_i$. Then $t_1 = 0$ and we claim that $t_{j+1} - t_j \approx jr \ln(m_j/m_{j+1})$. This follow as during $(t_j, t_{j+1}]$, each of the first j copies has a vertex scheduled from it every j steps, and hence during x steps the number of uncovered edges in any of $H_1, H_2, \ldots, H_j$ decreases to about a $(1 - 1/r)^{x/j} \approx \exp(-x/jr)$ fraction of its value at t_j. This gives

$$t_i \approx r \sum_{j=1}^{i-1} j \ln \frac{m_j}{m_{j+1}}$$

$$= (p+1)r \sum_{j=1}^{i-1} j(\ln(j+1) - \ln j)$$

$$= (p+1)r \left((i-1)\ln i - \sum_{j=1}^{i-1} \ln j\right) \approx (p+1)ri$$

where we use that $\sum_{j=1}^{i-1} \ln j \approx (i-1)\ln i - i$.

Consider the $i(m_i - m_{i+1})$ edges covered between t_i and t_{i+1}. The cost of each of them is at least $t_i^p \approx ((p+1)ri)^p$. Hence, the total cost is at least

$$\sum_{i=1}^{k} it_i^p(m_i - m_{i+1}) \approx i(p+1)^p r^p i^p \sum_{i=1}^{k} ma \frac{(p+1)}{i^{p+2}}$$

$$= (p+1)^{p+1} mar^p \sum_{i=1}^{k} \frac{1}{i}$$

$$\approx (p+1)^{p+1} mar^p \ln k$$

which is $(p+1)^{p+1}$ times $\sum_{i}^{k} F(i)$. Hence it is NP-hard to approximate MSSC$_p$ within $\delta_p - \epsilon$.

8 The conjecture of [20] on rounding preemptive solutions

Im et al. [20] defined the notion of a *preemptive* schedule as an assignment $x_v(t)$ of vertices v to real-valued times t satisfying $\sum_v x_v(t) = 1$ for all t and $\int_0^\infty x_v(t)dt = 1$ for all v. The cover time of an edge e in the preemptive schedule is defined to be the (real-valued) time t such that $\int_0^t \sum_{v \in e} x_v(t)dt = k_e$.

Im et al. [20] conjectured that for any instance there always exists a non-preemptive schedule (which is our usual notion of a schedule) of cost at most twice the cost of any preemptive schedule.

We now give a counter-example, which shows that this gap must be at least 4. This holds even for MSSC, and the

instance we use is the same as that used by [11] for their hardness result. As the approach of [20] loses another factor 2 to obtain a preemptive schedule from the configuration LP, this implies that their approach cannot give a better than 8 approximation (even for MSSC).

We recall that Feige et al. [10, 11] showed the following.

CLAIM 6. *For every $c_0, \epsilon > 0$, there exists a uniform regular hypergraph H with every edge having $|V|/r$ vertices (for sufficiently large r dependent on c_0) such that any collection of $c \leq c_0$ vertices covers at most $1 - (1 - 1/r)^c + \epsilon$ fraction of the edges.*

Using H from Claim 6, consider the following hard instance as in [11]. Make k disjoint copies of H (for some large enough constant k), on distinct universe of elements. For each $1 \leq i \leq k$, make a/i^2 copies of every edge in the ith copy of H where $a = (k!)^2$, call the resulting hypergraph H_i. Note that H_i has ma/i^2 edges and n vertices where $m = |E(H)|, n = |V(H)|$.

Feige et al. [11] showed that for this instance any (non-preemptive) schedule must have cost $\approx 4mar \ln k$. So it suffices to exhibit a preemptive schedule with cost at most $mar \ln k$.

Consider the following preemptive schedule: For $i = 1, 2, 3 \ldots$, schedule the vertices of H_i uniformly for r time units i.e. $x_v(t) = 1/n$ for $v \in H_i, t \in [r(i-1), ri]$. Each edge $e \in H_i$ is covered by time ri because the n/r vertices of e have each been scheduled to an extent r/n. The cost of the preemptive schedule is $\sum_{i=1}^{k}(ma/i^2)(ri) \approx mar \ln k$ as claimed.

References

[1] Yossi Azar, Iftah Gamzu, and Xiaoxin Yin. Multiple intents re-ranking. In *Symposium on Theory of computing, STOC*, pages 669–678, 2009.

[2] Shivnath Babu, Rajeev Motwani, Kamesh Munagala, Itaru Nishizawa, and Jennifer Widom. Adaptive ordering of pipelined stream filters. In *ACM SIGMOD international conference on Management of data*, pages 407–418, 2004.

[3] Nikhil Bansal, Anupam Gupta, and Ravishankar Krishnaswamy. A constant factor approximation algorithm for generalized min-sum set cover. In *Symposium on Discrete Algorithms, SODA*, pages 1539–1545, 2010.

[4] Nikhil Bansal and Subhash Khot. Optimal long code test with one free bit. In *Foundations of Computer Science, FOCS*, pages 453–462, 2009.

[5] Amotz Bar-Noy, Mihir Bellare, Magnús M Halldórsson, Hadas Shachnai, and Tami Tamir. On chromatic sums and distributed resource allocation. *Information and Computation*, 140(2):183–202, 1998.

[6] Uri Barenholz, Uriel Feige, David Peleg, et al. Improved approximation for min-sum vertex cover. Technical report, MCS06-07, Computer Science and Applied Mathematics, 2006.

[7] Samuel Burer and Renato DC Monteiro. A projected gradient algorithm for solving the maxcut sdp relaxation. *Optimization methods and Software*, 15(3-4):175–200, 2001.

[8] Chandra Chekuri and Rajeev Motwani. Precedence constrained scheduling to minimize sum of weighted completion times on a single machine. *Discrete Applied Mathematics*, 98(1-2):29–38, 1999.

[9] Edith Cohen, Amos Fiat, and Haim Kaplan. Efficient sequences of trials. In *Symposium on Discrete algorithms*, pages 737–746, 2003.

[10] Uriel Feige. A threshold of ln n for approximating set cover. *Journal of the ACM (JACM)*, 45(4):634–652, 1998.

[11] Uriel Feige, László Lovász, and Prasad Tetali. Approximating min sum set cover. *Algorithmica*, 40(4):219–234, 2004.

[12] Robbert Fokkink, Thomas Lidbetter, and László A Végh. On submodular search and machine scheduling. *Mathematics of Operations Research*, 44(4):1431–1449, 2019.

[13] Daniel Golovin, Anupam Gupta, Amit Kumar, and Kanat Tangwongsan. All-norms and all-l_p-norms approximation algorithms. In *Foundations of Software Technology and Theoretical Computer Science*, 2008.

[14] Leslie A Hall, Andreas S Schulz, David B Shmoys, and Joel Wein. Scheduling to minimize average completion time: Offline and on-line approximation algorithms. *Mathematics of Operations Research*, 22(3):513–544, 1997.

[15] Felix Happach, Lisa Hellerstein, and Thomas Lidbetter. A general framework for approximating min sum ordering problems. *arXiv preprint arXiv:2004.05954*, 2020.

[16] Felix Happach and Andreas S Schulz. Approximation algorithms and LP relaxations for scheduling problems related to min-sum set cover. *arXiv preprint arXiv:2001.07011*, 2020.

[17] Refael Hassin and Asaf Levin. An approximation algorithm for the minimum latency set cover problem. In *European Symposium on Algorithms, ESA*, pages 726–733, 2005.

[18] Wassily Hoeffding. On the distribution of the number of successes in independent trials. *The Annals of Mathematical Statistics*, 27(3):713–721, 1956.

[19] Sungjin Im, Viswanath Nagarajan, and Ruben Van Der Zwaan. Minimum latency submodular cover. *ACM Transactions on Algorithms (TALG)*, 13(1):1–28, 2016.

[20] Sungjin Im, Maxim Sviridenko, and Ruben van der Zwaan. Preemptive and non-preemptive generalized min sum set cover. *Mathematical Programming*, 145(1-2):377–401, 2014.

[21] Satoru Iwata, Prasad Tetali, and Pushkar Tripathi. Approximating minimum linear ordering problems. In *APPROX*, pages 206–217, 2012.

[22] Kumar Jogdeo and Stephen M Samuels. Monotone convergence of binomial probabilities and a generalization of ramanujan's equation. *The Annals of Mathematical Statistics*, 39(4):1191–1195, 1968.

[23] Christoph Kofler, Martha Larson, and Alan Hanjalic. Intent-aware video search result optimization. *IEEE transactions on multimedia*, 16(5):1421–1433, 2014.

[24] Christoph Kofler, Martha A. Larson, and Alan Hanjalic. User intent in multimedia search: A survey of the state of the art and future challenges. *ACM Comput. Surv.*, 49(2):36:1–36:37, 2016.

[25] François Margot, Maurice Queyranne, and Yaoguang Wang.

Decompositions, network flows, and a precedence constrained single-machine scheduling problem. *Operations Research*, 51(6):981–992, 2003.

[26] Jessica McClintock, Julián Mestre, and Anthony Wirth. Precedence-constrained min sum set cover. In *28th International Symposium on Algorithms and Computation, ISAAC*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

[27] Kamesh Munagala, Shivnath Babu, Rajeev Motwani, and Jennifer Widom. The pipelined set cover problem. In *International Conference on Database Theory*, pages 83–98, 2005.

[28] N Pisaruk. The boundaries of submodular functions. *Computational mathematics and mathematical physics*, 32(12):1769–1783, 1992.

[29] Martin Skutella and David P Williamson. A note on the generalized min-sum set cover problem. *Operations Research Letters*, 39(6):433–436, 2011.

[30] Panayiotis Tsaparas, Alexandros Ntoulas, and Evimaria Terzi. Selecting a comprehensive set of reviews. In *ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–176, 2011.

[31] Gerhard J Woeginger. On the approximability of average completion time scheduling under precedence constraints. *Discrete Applied Mathematics*, 131(1):237–252, 2003.

A   Proof of Claim 5

Let $f(a)$ and $g(a)$ be defined as

$$
\begin{aligned}
f(a) &= \Big(1 + \int_1^a P(1 + s\beta \ln x)dx \\
&\quad + \int_a^\infty P(1 + \beta \ln x - \beta(1-s)\ln a)\Big) \\
g(a) &= se^{-1/(s\beta)} + a(1-s).
\end{aligned}
$$

CLAIM 7. *For any $s \in [0,1)$ and $\beta \ge 1$, the ratio $f(a)/g(a)$, for $a \ge 1$, is maximized at $a = 1$.*

Proof. To prove that $f(a)/g(a)$ is maximized at $a = 1$, it suffices to show that $(f(a)/g(a))' \le 0$ for all $a \ge 1$. To this end, we will show that $h(a) = f'(a)g(a) - g'(a)f(a)$ is negative for all $a \ge 1$. We do this by proving that $h(1) \le 0$ and $h'(a) \le 0$ for $a \ge 1$.

Let us first note that $g'(a) = 1 - s$ and $g''(a) = 0$. We now compute f'.

$$
\begin{aligned}
f'(a) &= P(1 + s\beta \ln a) - P(1 + \beta \ln a - \beta(1-s)\ln a) \\
&\quad - \int_a^\infty \beta(1-s)\frac{1}{a}P'(1 + \beta \ln x - \beta(1-s)\ln a)dx \\
&= -\int_a^\infty \beta(1-s)\frac{1}{a}P'(1 + \beta \ln x - \beta(1-s)\ln a)dx \\
&= -\int_1^\infty \beta(1-s)P'(1 + s\beta \ln a + \beta \ln t)dt
\end{aligned}
$$

where the third equality follows by substituting $x = at$.

Proving $h'(a) \le 0 \quad \forall a \ge 1$. As $h' = f''g - g''f$, and $g'' = 0$, $g(a) \ge 0$, it suffices to show that $f''(a) \le 0$ for all $a \ge 1$. Using the expression for $f'(a)$ above, gives

$$
f''(a) = -\int_1^\infty \beta^2 s(1-s)\frac{1}{a}P''(1 + s\beta \ln a + \beta \ln t)dt
$$

As P is convex in $[1, \infty)$, $P''(1 + s\beta \ln a + \beta \ln t) \ge 0$ for all $t \ge 1$ and hence $f''(a) \le 0$.

It remains to show that $h(1) \le 0$. Let us compute $f(1), g(1), f'(1), g'(1)$. We have

$$
\begin{aligned}
f(1) &= 1 + \int_1^\infty P(1 + \beta \ln x)dx, \\
f'(1) &= -\int_1^\infty \beta(1-s)P'(1 + \beta \ln x)dx.
\end{aligned}
$$

Next, $g(1) = se^{-1/s\beta} + (1-s)$ and $g'(1) = 1 - s$.

Noting that $f'(1) \ge 0$ (as P is decreasing in $[1,\infty)$ and hence $P'(x) \le 0$ for all $x \ge 1$) and $g(1) \le 1$ (using $\exp(-1/s\beta) \le 1$), we have that

$$
\begin{aligned}
h(1) &= f'(1)g(1) - f(1)g'(1) \le f'(1) - f(1)(1-s) \\
&\le -(1-s)\left(1 + \int_1^\infty \beta P'(1 + \beta \ln x) + P(1 + \beta \ln x)dx\right) \\
&= -(1-s)\left(1 + \int_1^\infty e^{\frac{y-1}{\beta}}\left(P'(y) + \frac{P(y)}{\beta}\right)dy\right) \\
&= -(1-s)\left(1 + \int_1^\infty \left(e^{\frac{y-1}{\beta}}P(y)\right)'dy\right) \\
&\le -(1-s)(1 - P(1)) \\
&\le 0,
\end{aligned}
$$

where the equality in the third line is by substituting $1 + \beta \ln x = y$.   □

B   Proof of Theorem 1.5

Following the framework of the proof of Theorem 6.1, it suffices to show that for $g(\gamma) = e^{-\gamma}(e/2)^{\exp(1-\gamma)}$,

(B.1)
$$
\frac{1}{k}g'(\gamma) + g(\gamma) - g\Big(\gamma\frac{k}{k-1}\Big) \ge 0, \quad \text{for all } \gamma \ge 1,\, k \ge 2.
$$

Substituting $f(\gamma) = \ln g(\gamma)$ and noting $g'(\gamma) = f'(\gamma)g(\gamma)$, (B.1) becomes

$$
\frac{1}{k}f'(\gamma) + 1 \ge \frac{g\Big(\gamma\frac{k}{k-1}\Big)}{g(\gamma)}.
$$

Taking the logarithm on both sides,

$$
\ln\Big(\frac{1}{k}f'(\gamma) + 1\Big) \ge f\Big(\gamma\frac{k}{k-1}\Big) - f(\gamma).
$$

Let θ denote $\ln(e/2)$. Plugging $f(\gamma) = -\gamma + e\theta e^{-\gamma}$ and $f'(\gamma) = -1 - e\theta e^{-\gamma}$, it suffices to show

(B.2)
$$\ln\left(1 - \frac{1}{k} - \frac{e\theta}{k}e^{-\gamma}\right) \geq -\frac{\gamma}{k-1} + e\theta\left(\exp\left(-\frac{\gamma k}{k-1}\right) - e^{-\gamma}\right).$$

Denoting the left hand side of (B.2) by $\ell_k(\gamma)$ and the right hand side by $r_k(\gamma)$, we first show the following.

CLAIM 8. *For any fixed $k \geq 2$, $\ell'_k(\gamma) \geq 0$ and $r'_k(\gamma) \leq 0$.*

Proof. Since $\ln$ and $-e^{-\gamma}$ are non-decreasing and $\theta > 0$, $\ell'_k(\gamma) \geq 0$. Also,

$$\begin{aligned}
r'_k(\gamma) &= \frac{-1}{k-1} + e\theta\left(\frac{-k}{k-1}\exp\left(\frac{-\gamma k}{k-1}\right) + e^{-\gamma}\right) \\
&= e\theta e^{-\gamma}\left(\frac{-e^{\gamma}}{e\theta(k-1)} - \frac{k}{k-1}\exp\left(\frac{-\gamma}{k-1}\right) + 1\right) \\
&\leq e^{-\gamma}\left(-\frac{e^{\gamma}}{k-1} - \exp\left(-\frac{\gamma}{k-1}\right) + 1\right) \\
&\leq e^{-\gamma}\left(-\frac{e^{\gamma}}{k-1} + \frac{\gamma}{k-1}\right) \\
&\leq 0,
\end{aligned}$$

where the second and third inequalities are by using $e^{-x} \geq 1 - x$ for $x \geq 0$ and $e\theta < 1$ respectively. $\square$

By Claim 8, it suffices to prove (B.2) for $\gamma = 1$. That is, for all $k \geq 2$,
(B.3)
$$\ln\left(1 - \frac{1}{k} - \frac{\theta}{k}\right) \geq -\frac{1}{k-1} + \theta\left(\exp\left(-\frac{1}{k-1}\right) - 1\right).$$

Substitute $x = 1/(k-1)$. As $k \geq 2$, $x \in [0,1]$ and $1/k = x/(1+x)$. Then by (B.3), it suffices to show that

$$\ln\left(\frac{1 - \theta x}{x+1}\right) + x - \theta(e^{-x} - 1) \geq 0.$$

Let us denote the LHS by $h(x)$ and note that $h(0) = 0$. It suffices to show $h'(x) \geq 0$ for $x \in [0,1]$.

$$\begin{aligned}
h'(x) &= \frac{-\theta}{1 - \theta x} - \frac{1}{x+1} + 1 + \theta e^{-x} \\
&\geq \frac{-\theta}{1 - \theta x} - \frac{1}{x+1} + 1 + \theta(1-x) \\
&= \frac{x(1 - \theta - \theta^2 - 2\theta x + \theta^2 x^2)}{(1 - \theta x)(x+1)},
\end{aligned}$$

once again using, $e^{-x} \geq 1 - x$.

It suffices to show that $q(x) = (1 - \theta - \theta^2 - 2\theta x + \theta^2 x^2) \geq 0$, since the other terms are non-negative. $q'(x) = 2\theta(-1 + \theta x) \leq 0$ for $x \in [0,1]$ since $\theta = 1 - \ln 2 < 1$. Noting that $q(1) = 1 - 3\theta = 3\ln 2 - 2 \geq 0$, Theorem 1.5 follows.