Weight Reduction in Distributed Protocols: New Algorithms and Analysis

Anatoliy Zinovyev
Boston University, MA, USA

Abstract

We study the problem of minimizing the total weight of (potentially many) participants of a distributed protocol, a necessary step when the original values are large but the scheme to be deployed scales poorly with the weights. We assume that α fraction of the original weights can be corrupted and we must output new weights with at most β adversarial fraction, for $\alpha < \beta$. This problem can be viewed from the prism of electing a small committee that does the heavy work, a powerful tool for making distributed protocols scalable. We solve the variant that requires giving parties potentially multiple seats in the committee and counting each seat towards the cost of the solution. Moreover, we focus on the "deterministic" version of the problem where the computed committee must be secure *for any* subset of parties that can be corrupted by the adversary; such a committee can be smaller than a randomly sampled one in some cases and is useful when security against adaptive corruptions is desired but parties in the sub-protocol speak multiple times.

Presented are new algorithms for the problem as well as analysis of prior work. We give two variants of the algorithm Swiper (PODC 2024), one that significantly improves the running time without sacrificing the quality of the output and the other improving the output for a reasonable increase in the running time. We prove, however, that all known algorithms, including our two variants of Swiper, have worst case approximation ratio $\Omega(n)$. To counter that, we give the first polynomial time algorithm with approximation factor $n/\log^2 n$ and also the first sub-exponential time exact algorithm, practical for some real-world inputs. Of theoretical interest is another polytime algorithm that we present, based on linear programming, whose output is no worse than an optimal solution to the problem with slightly different parameters.

We implemented and tested previous and new algorithms, comparing them on the stake distributions of popular proof-of-stake blockchains, and found that our second variant of Swiper computes solutions extremely close to the optimal, confirmed by our exact algorithm.

2012 ACM Subject Classification Theory of computation → Approximation algorithms analysis; Theory of computation → Cryptographic protocols; Theory of computation → Distributed algorithms

Keywords and phrases Weight reduction, distributed protocols, weighted cryptography, threshold cryptography, consensus, committee selection, adaptive corruptions, approximation algorithms, linear programming, rounding

Digital Object Identifier 10.4230/LIPIcs.DISC.2025.43

Related Version *Full Version*: `https://eprint.iacr.org/2025/1076` [42]

Supplementary Material *Software (Source Code)*: `https://github.com/tolikzinovyev/weight-reduction/tree/f1b5ce44aab9e4809bff67c5c6bff4acf1f82cb1` [41]
archived at `swh:1:dir:a48c33feba04bc98c7013c4c88b174f5e461a392`

Funding This material is based upon work supported in part by a gift from DARPA under Agreements No. HR00112020021, HR00112020023 and HR001120C0085. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Government or DARPA.
Anatoliy Zinovyev: DARPA under Agreements No. HR00112020021, HR00112020023 and HR001120C0085.

Acknowledgements The author thanks his advisor Leo Reyzin for help with preparing this paper and Nathan Klein, Rico Zenklusen for helpful discussions.

1 Introduction

Committee selection is an essential tool for making large-scale distributed protocols scalable. The idea is that, instead of all n parties doing the work, the protocol is run within a smaller committee, decided by the larger protocol; the participants outside the committee simply learn the result. To spread the work evenly or to achieve better security, multiple different committees can be created.

We are mainly interested in byzantine fault tolerant / cryptographic protocols where parties have some assigned weights w_i and the adversary can corrupt parties up to some α fraction of the total weight in the system. For example, in a proof-of-stake blockchain parties own some amount of money in the system (stake) and the assumption is that the adversary controls, say, at most 20% of all the stake. We would like to compute a "small" committee where the adversary controls less than some larger β fraction, say 1/3, of the total weight. Specifically, our goal is to assign each party a new non-negative integer weight t_i such that we minimize the total new weight $\sum_{i=1}^{n} t_i$. This is useful when the protocol in question doesn't scale well with the weights such as in secret sharing, threshold signatures or secure multiparty computation. Additionally, we focus on "deterministic" committee selection that must output a committee secure against *any* set of parties that the adversary is allowed to corrupt. Formally, the new weights $t_1, ..., t_n$ must satisfy

$$\forall A \subseteq [n] \text{ with } \sum_{i \in A} w_i \leq \alpha \sum_{i=1}^{n} w_i,\ \sum_{i \in A} t_i < \beta \sum_{i=1}^{n} t_i$$

(here A denotes the set of parties controlled by the adversary). This property is useful when one requires security against an adversary that performs adaptive corruptions throughout the protocol execution with full knowledge of $(t_1, ..., t_n)$ – for example, when the members of the committee need to speak multiple times and thus reveal themselves after the first message. Moreover, committees of this flavor can be smaller than traditional "randomized" committees, as explained in section 2. As demonstrated in [38], they also yield a convenient way to transform a wide range of unweighted protocols to their weighted version. We note that "deterministic" committee selection cannot be usefully instantiated in the unweighted setting since any committee would have to have $\Omega(n)$ members (otherwise, the adversary can corrupt the whole committee). In contrast, when parties have weights distributed non-uniformly, we can often compute a committee with the total new weight much smaller than n.

1.1 Prior work

The "deterministic" committee selection problem has a short history. To the best of our knowledge, the recent Swiper paper [38] by Tonkikh and Freitas was the first and only to study it in isolation. Near-simultaneously, other works [4, 23, 22] attempted to solve it in the context of their larger protocols. All four papers consider different variants of the problem although all aim to minimize the total new weight $\sum_{i=1}^{n} t_i$.

Swiper defines a problem they call weight restriction that is only slightly different from the one we consider: given $w_1, ..., w_n$ as input, output $t_1, ..., t_n \in \mathbb{Z}_{\geq 0}$ such that for all $A \subseteq [n]$ with $\sum_{i \in A} w_i < \alpha \sum_{i=1}^{n} w_i$, we have $\sum_{i \in A} t_i < \beta \sum_{i=1}^{n} t_i$. The only distinction is that the inequality containing α is non-strict in our paper.

The authors of Swiper prove that the problem admits an efficiently computable solution of cost $\sum_{i=1}^{n} t_i \leq O(n)$ (assuming α and β parameters are constant). Such a solution can be found by linearly scaling down the vector $(w_1, ..., w_n)$ and rounding each w_i to an integer

according to a simple rule. Swiper goes a step further and tries to compute a solution as small as possible by running a binary search on the scaling factor and testing the correctness of a vector $(t_1, ..., t_n)$ at each iteration. Its time complexity is $O(n^2 \log n)$.

We note that Swiper defines two more weight reduction problems. Their overview and the overview of the other works [4, 23, 22] that consider deterministic committee selection can be found in the full version of our paper [42, Appendix A].

1.2 Our contribution and paper outline

Section 2 gives background information on committee selection and how the problem we consider relates to other variants. In section 3 we formally define the problem and notation. Section 4 provides a detailed overview of the Swiper algorithm, useful for reading the following section.

Section 5 presents our new variants of Swiper called faster Swiper and extended Swiper. The first variant improves the time complexity of Swiper from $O(n^2 \log n)$ to $O(n+R^2)$ where $R \leq O(n)$ is the total weight in the output. This improves the running time by a $\log n$ factor in the worst case, but given that the cost of the output R is often much smaller than n, the improvement is much larger as demonstrated by our numerical evaluations. For low latency applications, our optimizations would be highly desired. Importantly, faster Swiper doesn't alter the quality of the output of Swiper. Our second variant extended Swiper improves the output quality by replacing the binary search with a linear search. Although the idea is simple, the algorithm has non-trivial time complexity of $O(n + R^{2.5})$. Numerical evaluations demonstrate a sizable reduction in the total new weight over Swiper and faster Swiper, since the two can get stuck in a local minimum; in fact, the cost of extended Swiper's output is extremely close to optimal in our tests. At the same time, its running time is comparable and often even smaller than that of Swiper. Both variants rely on two optimizations: 1) carefully bounding the running time in terms of the output cost R and 2) reusing dynamic programming computations.

We, however, prove in section 6 that all previously known algorithms, as well as our two Swiper variants, suffer from $\Omega(n)$ approximation ratio in the worst case. Our lower bound applies to a class of linear scaling based solutions which contains all but one known algorithm (which is also based on linear scaling but requires a different argument). Despite faster Swiper and extended Swiper being great "heuristic" schemes, we would still like to find an algorithm with better theoretical guarantees. We make some progress with this in subsequent sections.

Section 7 proves several facts about our weight reduction problem useful in the rest of the paper. Of particular importance is a new observation that assuming the weights $w_1, ..., w_n$ in the input are sorted, there exists an optimal solution $t_1, ..., t_n$ that is also sorted. In section 8 we show that this fact yields the first polynomial time algorithm with approximation factor $n/\left(c \log^2 n\right)$, for any $c > 0$, and the first exact (approximation factor 1) solution with time complexity $O\left(n + R^{3/2} e^{C\sqrt{R}}\right)$, where R is the cost of an optimal solution and $C \approx 2.57$. Given that R is often much smaller than n, our exact algorithm turns out to be practical for some real-world weight distributions, as shown by numerical evaluations. We believe it makes sense to use it in production when the optimal solution is expected to be small, in combination with a fallback scheme if the exact algorithm takes too long. It can also be very useful for research purposes, for example, to test some hypothesis about the problem.

Section 9 contains numerical evaluations comparing old and new algorithms in terms of the total weight in the output and the running time. It confirms our claims about the improved running time of faster Swiper and the improved output quality of extended Swiper. It also proves our sub-exponential time exact algorithm to be somewhat practical.

Finally, Appendix A describes a polynomial time algorithm based on the standard approach of representing an optimization problem as an integer linear program, relaxing it to allow fractional solutions, running an LP solver and rounding the solution to an integer vector in a problem specific manner. When denoting the cost of an optimal solution by $\mathsf{OPT}_{\leq\alpha,<\beta}$, our algorithm guarantees a solution of cost at most $\mathsf{OPT}_{\leq\alpha,<(1-\delta)\beta}$ for any arbitrarily small constant $\delta > 0$. In the full version of our paper [42, Appendix B] we, however, prove that this can be $\Omega(n)$ times larger than the optimal solution. Nevertheless, we believe the algorithm presents theoretical interest useful for future research. Moreover, such a guarantee can be meaningful when the protocol relying on committee selection can work with a continuous range of the final adversarial fractions β. Take for example Approximate Lower Bound Arguments [16] that allow one to create short threshold signatures. Its decentralized version uses committee selection and presents a configurable tradeoff between the size of the committee and the size of the final proof (e.g., a threshold signature) where the proof size would be some monotone function of β in our notation. Their committee selection algorithm is agnostic to the weight distribution, but suppose we replace it with theoretically smallest committees for any given β and plot the resulting tradeoff with the committee size on the horizontal axis and the proof size on the vertical axis. We would like to achieve this tradeoff, but since the optimal committee selection is likely infeasible, we are forced to use an approximation and get a worse curve. Using algorithms approximating the committee size for a fixed β, such as the one in subsection 8.2, would shift the tradeoff graph to the right while algorithms approximating β for a fixed committee size, such as the one in Appendix A, would shift the graph up. In the absence of a good approximation on the committee size, the second transformation could even be superior.

2 General overview of committee selection

The idea of selecting small committees can be seen as far back as [9]. It has since been used in constructing many efficient protocols for consensus [35, 36, 34, 10, 7]. For example, the proof-of-stake blockchain protocol Algorand [30, 17] modifies a classical setting $O(n^2)$ byzantine fault tolerant consensus protocol to run on committees created by selecting each unit of stake with some probability, which makes communication complexity per party independent of the number of parties. Additionally, each step of the BFT sub-protocol uses a newly sampled anonymous committee that reveals itself only when the members of the committee speak, thus achieving security against an adaptive adversary. Such an adversary can corrupt participants at any time of the protocol based on the information he possesses thus far, up to some threshold of total corrupted stake.

Committee selection can also be seen in communication efficient threshold signatures [15, 16, 24], leader election [37, 33], secure multiparty computation [8, 5, 19, 6, 18], and distributed key-generation [12].

Schemes utilizing committees typically possess a gap between the absolutely necessary requirement on the fraction of adversarial power and the smaller fraction assumed by the protocol. For instance, solving consensus in an asynchronous environment requires adversarial weight to be strictly less than 1/3 of the total weight while Algorand assumes that at most 1/5 of the total stake is controlled by the adversary. Motivated by proof-of-stake blockchains, we are interested in constructing small committees in the *weighted* setting. Given original weights $w_1, w_2, ..., w_n \in \mathbb{N}$ for n parties such that the adversary controls parties of total weight at most $\alpha \sum_{i=1}^{n} w_i$, we need to compute new weights $t_1, t_2, ..., t_n \in \mathbb{Z}_{\geq 0}$ such that the adversary can control strictly less than $\beta \sum_{i=1}^{n} t_i$ of new weight.

There are at least two dimensions of interesting variants of this problem. One is the definition of the objective function. A party in Algorand sends out one vote per step of the protocol even if it has multiple units of stake on the committee. Thus, for Algorand, we are interested only in the number of distinct parties on the committee or, equivalently, the number of non-zero t_i. On the other hand, there are schemes that cannot efficiently aggregate weights, say Shamir secret sharing. For such protocols, it may be necessary to use "virtualization" – essentially, giving w virtual identities to a party with weight $w \in \mathbb{Z}_{\geq 0}$ and running this many copies of the protocol. This can be prohibitively expensive when weights are large (e.g., in proof-of-stake blockchains weights can be on the order of 2^{64}), and thus, for such protocols we are interested in constructing committees with $\sum_{i=1}^{n} t_i$ as small as possible – the goal we pursue in this paper. We note that there have been many works trying to make cryptographic protocols scale better with the weights [3, 4, 27, 20, 15, 16, 24]. Still, protocols such as those in [22] only improve the dependence on weights for the verifiers while the provers' running time is still linear in the weights, and protocols such as those in [26] only improve complexities to $O(W)$ down from $O(W \cdot \lambda)$, for a total weight W and security parameter λ.

Another configuration knob for the problem of committee selection is, what kind of security do we require? Traditionally, especially in a setting where the adversary is static, meaning he chooses parties to corrupt before the start of the protocol, a random committee is chosen such that the (unknown) corrupted set of parties owns only a small fraction of the elected weight. More formally, the security requirement of this "randomized" variant is that

$$\begin{aligned} &\forall (w_1, ..., w_n),\ \forall A \subseteq [n] \text{ with } \sum_{i \in A} w_i \leq \alpha \sum_{i=1}^{n} w_i, \\ &\Pr\left[\sum_{i \in A} t_i < \beta \sum_{i=1}^{n} t_i : (t_1, ..., t_n) \leftarrow \text{ComputeCommittee}(w_1, ..., w_n)\right] \text{ is large} \end{aligned} \tag{1}$$

(we assume that α and β are constants). Instead, this work considers a stronger security notion where the computed committee must be secure *for any* subset of the parties that can be corrupted. A more formal statement of this "deterministic" version of the problem is as follows.

$$\begin{aligned} &\forall (w_1, ..., w_n), \quad \text{if } (t_1, ..., t_n) = \text{ComputeCommittee}(w_1, ..., w_n) \text{ then} \\ &\forall A \subseteq [n] \text{ with } \sum_{i \in A} w_i \leq \alpha \sum_{i=1}^{n} w_i,\ \sum_{i \in A} t_i < \beta \sum_{i=1}^{n} t_i \end{aligned} \tag{2}$$

(where ComputeCommittee is now deterministic). When the weaker security notion is sufficient, it is great news since selecting each unit of weight with some equal probability is a simple and efficient scheme yielding an expected constant size committee; this is generally impossible for a deterministic algorithm (since n parties with equal weights must elect a committee of size $\Omega(n)$, lest the entire committee be corrupted). On the other hand, deterministic committee selection makes it easy to maintain security against an adversary who is allowed to corrupt parties adaptively throughout the protocol execution since any corrupted set will have less than β fraction of new weight in the committee. While simple quorum systems can get away with using random committees with VRF as a source of secret randomness, other problems are less trivial and require intricate protocols and arguments of security [17, 29, 14, 11, 1]. The security against adaptive corruptions has been recognized important by the community, and in fact, is the reason why [22, 21] has been deployed in the Aptos blockchain in combination with a deterministic committee selection scheme (from

private communication with the authors). Moreover, there is evidence that a deterministic scheme can produce smaller committees than a randomized one. Take, for example, the standard randomized procedure of selecting each unit of weight with some equal probability. As the requirement on how large the probability in Equation 1 must be increases, the expected number of selected units of weight approaches infinity and the expected number of elected parties approaches n. Thus, for a sufficiently large mandated success probability, the chosen committee would be worse than that of a deterministic algorithm. Another example is the randomized schemes presented in [28] called Fait Accompli that include larger parties in the committee deterministically to either reduce the number of parties in the committee or their total new weight. From their formulas and plots, one can see that as the required probability of failure decreases, the committee must grow. Indeed, it seems inherent that when we require $\sum_{i\in A} t_i$ in Equation 1 to be increasingly concentrated, its expectation must approach infinity; it would be a good exercise to prove that as the probability of failure goes to 0, whenever at least one t_i is non-deterministic, $E\left[\sum_{i=1}^{n} t_i\right]$ must approach infinity and $E\left[|\{i : t_i > 0\}|\right]$ must approach or exceed the number of parties in an optimal deterministic committee.

In our work, we focus on the "deterministic" flavor of committee selection with the objective to minimize the total weight $\sum_{i=1}^{n} t_i$, and believe this problem is important for the reasons outlined above. Our paper can be viewed as a continuation of the Fait Accompli work [28] by Gaži, Kiayias and Russell and the Swiper work [38] by Tonkikh and Freitas that frame committee selection as an optimization problem with parties' weights in the input. Fait Accompli shows randomized schemes that optimize either $\sum_{i=1}^{n} t_i$ or $|\{i : t_i > 0\}|$, while Swiper describes algorithms for the same problem we consider, and in fact, we directly extend some of their results. Despite both works numerically showing impressive results on real-world, far from uniform, weight distributions, they say nothing about the quality of their algorithms' output compared to the best possible solution. The golden standard in approximation algorithms is to provide one that yields solutions of the cost within a certain multiplicative factor of the cost of an optimal solution [40], or otherwise bounded by some other "interesting" function of the input (as in [32] for ECSS). We make some progress towards this.

Finally, we note that there is no reason to require algorithms satisfying the stronger security notion in Equation 2 to be deterministic. The use of randomness in optimization often helps [40], but all algorithms in this paper are in fact deterministic.

3 Problem definition

Let $n \in \mathbb{N}$ be the number of parties, $w_1, ..., w_n \in \mathbb{N}$ be their (original) weights, and let α and β be constants independent of n with $0 < \alpha < \beta < 1$. Borrowing notation from [38], for the new weights $t_1, ..., t_n \in \mathbb{Z}_{\geq 0}$, we say party i gets t_i tickets and call $(t_1, ..., t_n)$ a ticket assignment. We call $(t_1, ..., t_n)$ valid if and only if $\sum_{i\in I} t_i < \beta \sum_{i=1}^{n} t_i$ whenever $I \subseteq [n]$ and $\sum_{i\in I} w_i \leq \alpha \sum_{i=1}^{n} w_i$. Our goal is to find a valid ticket assignment that minimizes the total number of tickets $\sum_{i=1}^{n} t_i$. We call this the $(\leq \alpha, < \beta)$-weight reduction problem and denote $\mathsf{OPT}_{\leq\alpha,<\beta}((w_1, ..., w_n))$ to be the smallest total number of tickets $\sum_{i=1}^{n} t_i$ in a valid solution $(t_1, ..., t_n)$. We can also formulate the problem using the following integer linear program.

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{n} t_i \\ \text{subject to} \quad & \sum_{i\in I} t_i < \beta \sum_{i=1}^{n} t_i \quad \forall I \subseteq [n] \text{ s.t. } \sum_{i\in I} w_i \leq \alpha \sum_{i=1}^{n} w_i \\ & t_i \in \mathbb{Z}_{\geq 0} \qquad\qquad \forall i \in [n] \end{aligned}$$

We note that the formulation in [38] uses a strict inequality where α is involved. For simplicity, especially when dealing with linear programming, we use a non-strict inequality. However, we still keep a strict inequality where β is involved since we are typically interested in solutions where the adversary controls strictly less than $1/2$ or $1/3$ of the total number of tickets and formulating the problem using a strict inequality simplifies asymptotic analysis where n approaches infinity. We call the variant in [38] the $(<\alpha, <\beta)$-weight reduction problem with optimum $\mathsf{OPT}_{<\alpha,<\beta}((w_1, ..., w_n))$, and also similarly define $(\leq\alpha, \leq\beta)$- and $(<\alpha, \leq\beta)$-weight reduction problems that have an additional requirement $\sum_{i=1}^{n} t_i \geq 1$.

We also note that we do not allow zero weights w_i in the input, but it does not change the problem since it is always optimal to set $t_i := 0$ for such i.

4 Detailed overview of Swiper

We now give a more detailed overview of the Swiper algorithm from [38, 39] since we will be modifying it later in section 5.

It works by linearly scaling down the weights and rounding to an integer according to a rule. Specifically, given weights $w_1, ..., w_n$ as input, Swiper outputs $t_1, ..., t_n$ such that for all i, either $t_i = \lfloor sw_i + c \rfloor$ or $t_i = \lfloor sw_i + c \rfloor - 1$, where c is a constant to be defined later. Swiper first runs a binary search on s to find a local minimal value for which the ticket assignment with all $t_i = \lfloor sw_i + c \rfloor$ is valid. Then the algorithm considers those i for which $(sw_i + c)$ is an integer, or equivalently those i for which t_i decreases when s is decreased by any value; call this borderline set of indices B. Finally, Swiper runs a binary search on $k \in \mathbb{N}$ to find a local minimal value in $\left[|B|\right]$ for which the ticket assignment remains valid when all but k arbitrary indices from B lose one ticket.

From a slightly different perspective, Swiper defines a set of potential solutions $\vec{t_1}, \vec{t_2}, ...$ with $\sum_{i=1}^{n} \left(\vec{t_j}\right)_i = j$ and finds a local minimal j for which $\vec{t_j}$ is a correct solution but $\vec{t_{j-1}}$ is not.

The authors observe that starting some index $M \in O(n)$, all $\vec{t_M}, \vec{t_{M+1}}, ...$ are valid. In particular, when $c = \alpha$, we have $M = \left\lfloor \frac{\alpha(1-\alpha)}{\beta-\alpha} n + 1 \right\rfloor$ (for the $(\leq\alpha, <\beta)$ weight reduction variant of the problem). By making M the right threshold of the search, Swiper always outputs a solution with at most this many tickets, though a much better solution is usually found for real-world weight distributions.

Each iteration of a binary search needs to test whether some ticket assignment $(t_1, ..., t_n)$ is valid. This is done precisely by solving the knapsack problem where weights are w_i and profits are t_i. Since Swiper only tests ticket assignments of size at most $M = O(n)$, the knapsack problem can be solved in time $O\left(n \cdot \sum_{i=1}^{n} t_i\right) = O\left(n^2\right)$ using dynamic programming that for each $0 \leq t \leq \sum_{i=1}^{n} t_i$ computes the smallest possible weight that achieves t tickets. The total time complexity of Swiper is $O(n^2 \log n)$.

5 Improving Swiper

In this section we present two variants of Swiper, one that improves the running time of the algorithm and the other that improves the output at the cost of a reasonable increase to the running time.

5.1 Faster Swiper

We first show a new algorithm named faster Swiper that improves the asymptotic running time of the original Swiper algorithm without semantically changing the output. Like the original Swiper, faster Swiper finds a ticket assignment $\vec{t}$ such that

- $\vec{t}$ is valid, equals a linear scaling of the weights according to Swiper rules (see section 4), and $\sum_i \left(\vec{t}\right)_i$ satisfies the $O(n)$ bound of the original Swiper;
- there exists a ticket assignment $\vec{t'}$ with one fewer ticket such that it equals a linear scaling of the weights according to Swiper rules but is not valid.

The stated time complexity of Swiper in [38] is $O(n^2 \log n)$ where n is the number of weights in the input. Our algorithm faster Swiper has time complexity $O(n + R^2)$ where R is the cost of the output (i.e., if the output is $\vec{t}$, $R = \sum_i \left(\vec{t}\right)_i$). Note that R is always $O(n)$ but at the same time, R can be much smaller than n in practice, and thus, looking at the asymptotic complexity in terms of two variables n and R can be useful.

It will be convenient to have access to the input weights vector sorted in the descending order. Let us assume such sorted input $(w_1, w_2, ..., w_n)$. We first explain how to achieve running time $O(R^2 \log R)$ and then improve it further to $O(R^2)$. In subsection 5.3 we will explain how to remove the assumption.

5.1.1 $O(R^2 \log R)$ time complexity

For all $i \geq 0$, let $\vec{t_i}$ be the vector with i tickets in total constructed according to Swiper rules such that for any i, $\vec{t_i}$ and $\overrightarrow{t_{i+1}}$ have the same values for all indices except for one; when multiple indices could be incremented in the transition from $\vec{t_i}$ to $\overrightarrow{t_{i+1}}$, choose the lowest index. The (uniquely defined) sequence $\left\{\vec{t_i}\right\}$ has a property that if $\left(\vec{t_i}\right)_j > 0$ then $\left(\vec{t_i}\right)_k > 0$ for all $k \leq j$. This in turn implies that if the cost of $\vec{t_i}$ is c, only the first c indices of $\vec{t_i}$ can be non-zero. Thus, if an algorithm only ever considers ticket assignments of cost at most $O(R)$, it only needs to query the first $O(R)$ elements of the (sorted) weight vector.

We make use of this fact. To get time complexity $O(R^2 \log R)$, we run binary search on i to find a value for which $\vec{t_i}$ is a valid ticket assignment but $\overrightarrow{t_{i-1}}$ is not, roughly as follows.

```
l ← 0; r ← 1;
while t_r⃗ is not valid do
    l ← r; r ← 2 · r;
while r − l > 1 do
    m ← (l + r)/2;
    if t_m⃗ is valid then
        r ← m;
    else
        l ← m;
return t_r⃗;
```

This requires $O(\log R)$ iterations of binary search, unlike in the original Swiper that does binary search on the scaling factor s potentially requiring $\Omega(\log n)$ iteration. Each iteration needs to solve a knapsack problem using dynamic programming to determine whether some $\vec{t_i}$ is valid. The time required to do this is at most the total number of tickets times the number of non-zero entries in $\vec{t_i}$, since the zero entries cannot help the adversary. Both are at most i. Hence, the first loop will run in time $O(R^2)$ and the second loop in time $O(R^2 \log R)$.

We are left to describe how to efficiently calculate $\vec{t_0}, \vec{t_1}, ..., \overrightarrow{t_{2R}}$. We use a mutable data structure called gen with a single method gen.next() such that if the i-th invocation of gen.next() returns index, then $\left(\vec{t_i}\right)_{\text{index}} = \left(\overrightarrow{t_{i-1}}\right)_{\text{index}} + 1$ (and all other indices stay the same). Such a data structure can be constructed using a binary heap as a priority queue defining the order of the produced indices by gen.next(). The heap initially contains only the first index – the index of the largest weight. When gen.next() outputs some index, this index is reinserted in the heap; if index has never been produced before, index + 1 is additionally

inserted. Therefore, during the k-th call to gen.next(), the heap contains at most $2k$ elements and thus, the call takes time $O(\log k)$. Our Swiper variant needs to query gen at most $2R$ times, and hence, will add $O(R \log R)$ to the time complexity.

5.1.2 $O(R^2)$ time complexity

To improve the running time of faster Swiper further, we notice that iterations of the binary search can reuse some of the dynamic programming (DP) computations. The structure of the binary search remains the same as in the $O(R^2 \log R)$ version, and we only need to optimize its main phase (the second while loop) since the initial phase already runs in time $O(R^2)$.

We need a data structure called DP containing (partial) DP computations for the knapsack problem. Specifically, it will have methods DP.apply(w, t) and DP.get() such that after running DP.apply on a sequence $(w_1, t_1), ..., (w_k, t_k)$, DP.get() reports the maximum number of tickets that the adversary can get $\sum_{j \in S} t_j$ provided $\sum_{j \in S} w_j \le \alpha W$ ($W = \sum_i w_i$ being the total number of tickets in the input). Since the binary search only tests ticket assignments $\vec{t_i}$ of cost at most $2R$, DP.apply will always run in time $O(R)$. Additionally, a complete DP computation requires calling DP.apply the number of times at most the number of non-zero elements of $\vec{t_i}$ which is at most $2R$.

The idea is to reuse some DP computation to have each iteration of the binary search (second loop) run in time $O(R(r-l))$. Since $(r-l)$ halves each iteration, the entire algorithm will be $O(R^2)$.

The loop will have the following invariants holding at the beginning of the loop.

- tickets_l: array containing the non-zero prefix of $\vec{t_l}$;
- deltas: array containing the indices from gen.next() to generate $\overrightarrow{t_{l+1}}, ..., \vec{t_r}$;
- dp_head: DP data structure with applied $\left(w_i, (\vec{t_l})_i\right)$ for all $i \notin$ deltas.

Essentially, dp_head will have indices applied whose values are not going to change anymore. Then testing the validity of $\vec{t_m}$ requires computing it in time $O(R)$ and applying the missing indices, contained in deltas, to a copy of dp_head in time $O(R(r-l))$. After that, the loop iteration needs to update the variables to maintain the invariants: tickets_l in time $O(R)$, deltas in constant time and dp_head in time $O(R(r-l))$ by applying a subset of indices from deltas.

5.2 Extended Swiper

We will now present a variant of Swiper that we call extended Swiper which replaces the binary search on the number of tickets in the solution with a linear search.

$i \leftarrow 1$;
while $\vec{t_i}$ is not valid **do**
  $i \leftarrow i + 1$;
return $\vec{t_i}$;

The linear search sometimes significantly improves the output in practice since the binary search gets stuck in a "local minimum". Although the idea is trivial, our contribution is an algorithm that runs the linear search in time $O(R^{2.5})$ instead of the easy $O(R^3)$, assuming the input weights are sorted.

The idea is to reuse DP computation as in subsection 5.1. We check the validity of $\vec{t_i}$ in batches. The batch number $j \ge 1$ will be $(\overrightarrow{t_{i_j+1}}, ..., \overrightarrow{t_{i_{j+1}}})$ where $i_1 = 0$ and $i_{j+1} = i_j + \Theta\left(\sqrt{i_j}\right)$. The batch j will be processed in time $O\left(i_j^2 + (i_{j+1} - i_j)^2 \cdot i_{j+1}\right) = O\left(i_j^2\right)$.

Before describing the batch algorithm, we briefly argue why the time complexity of the whole algorithm is $O(R^{2.5})$. Take any $d \in \mathbb{N} \cup \{0\}$ and let j be the largest integer such that $i_j + 1 \leq 2^d$ and k the smallest integer such that $i_k \geq 2^{d+1} - 1$. The batches $j, ..., (k-1)$, $\Theta(2^{d/2})$ in total, will each be processed in time $O(2^{2d})$ and thus together in time $O(2^{2.5d})$. Hence, the whole algorithm works in time $O\Big(\sum_{d=0}^{\lceil \log R \rceil} 2^{2.5d}\Big) = O\big(2^{2.5 \log R}\big) = O\big(R^{2.5}\big)$.

We now describe the batch algorithm. For batch j, collect indices from gen.next() that generate $\overrightarrow{t_{i_j+1}}, ..., \overrightarrow{t_{i_{j+1}}}$ into an array deltas. Create a DP data structure dp_head and call dp_head.apply on $\Big(w_k, \big(\overrightarrow{t_{i_j}}\big)_k\Big)$ for all $k \notin$ deltas, this takes time $O(i_j^2)$. Now for each index in deltas, apply it to get the next $\overrightarrow{t_i}$ and test its validity by applying $\Big(w_k, \big(\overrightarrow{t_i}\big)_k\Big)$ for distinct $k \in$ deltas to a copy of dp_head. This takes time $O\big((i_{j+1} - i_j) \cdot i_{j+1} \cdot (i_{j+1} - i_j)\big)$. Hence, the batch algorithm has running time $O\big(i_j^2 + (i_{j+1} - i_j)^2 \cdot i_{j+1}\big)$.

5.3 Arbitrarily ordered weights

We previously assumed that the weights vector in the input is sorted. We will now explain how to remove this assumption for an additive cost of $O(n + R \log R)$.

We need a mutable data structure that we call sorted_weights that has a single method sorted_weights.next() returning the next largest weight from the input. We require that initializing sorted_weights takes time $O(n)$ and that the first m calls to sorted_weights.next() take cumulative time $O(m \log m)$. This is sufficient since our algorithms would query sorted_weights.next() only $O(R)$ times.

Such a data structure can be constructed using a binary max-heap initialized in time $O(n)$ to contain all the weights and an auxiliary max-heap initially containing the reference to the main heap's root. sorted_weights.next() returns the root of the auxiliary heap and inserts in the auxiliary heap the references to the two children of the returned element in the main heap. [25]

5.4 A note on reducing time complexity further

Swiper [39] explores using a quasilinear time approximate knapsack solver in lieu of the quadratic time exact algorithm. Numerical evaluations by the authors show that this reduces the quality of the output only by a little. While we do not explore this modification, it could be interesting to employ it to potentially reduce the running time of faster Swiper and extended Swiper to $O(n) + \tilde{O}(R)$ and $O(n) + \tilde{O}(R^{1.5})$ respectively.

6 Analysis of linear scaling based algorithms

We are interested in analyzing the approximation factor of known algorithms solving the $(\leq \alpha, < \beta)$-weight reduction problem, meaning how many times larger can the computed solution be compared to an optimal solution. We notice that all previously known algorithms use linear scaling with rounding to compute ticket assignments $(t_1, ..., t_n)$ and prove that all have worst case approximation factor $\Omega(n)$. We start with the following theorem that applies to a class of linear scaling based solutions.

▶ **Theorem 1.** *Let $\alpha, \beta \in \mathbb{R}$ and $r, s \in \mathbb{N}$ such that $0 < \alpha < \beta < 1$ and $\frac{r}{s+1} < \beta \leq r/s \leq 1$. Also let $0 \leq c < 1$, $m \in \mathbb{Z}_{\geq 0}$, and suppose an algorithm solving the $(\leq \alpha, < \beta)$-weight reduction problem on input $(w_1, ..., w_n)$ always outputs $(t_1, ..., t_n)$ with the property that there*

exists $\delta > 0$ such that for all i,

$$\lfloor \delta w_i + c \rfloor - m \leq t_i \leq \lfloor \delta w_i + c \rfloor.$$

Then for infinitely many n there exists a weight vector $\vec{w} = (w_1, ..., w_n)$ such that

$$OPT_{\leq \alpha, < \beta}(\vec{w}) = s + 1$$

but the algorithm outputs a solution with the number of tickets

$$> \big(1 - c\big) \frac{\alpha n - r - 1}{r/s - \alpha} - ms.$$

The worst case approximation factor of the algorithm is thus

$$> \frac{(1-c)\alpha}{(s+1)(r/s - \alpha)} n - O(1).$$

Proof. Define $\vec{w} = (w_1, ..., w_n)$ where $w_1 = w_2 = ... = w_{n-s} = 1$ and $w_{n-s+1} = ... = w_n = w$ where

$$w = \left\lfloor \frac{\alpha n - \alpha s}{r - \alpha s} \right\rfloor = \frac{\alpha n - \alpha s}{r - \alpha s} - \varepsilon \tag{3}$$

for some $0 \leq \varepsilon < 1$. The total weight is $\sum_{i=1}^{n} w_i = n - s + sw$ and the adversary is allowed to corrupt the following weight.

$$W_{\mathsf{A}} = \lfloor \alpha W \rfloor = \lfloor \alpha n - \alpha s + \alpha s w \rfloor = \left\lfloor \alpha n - \alpha s + \frac{\alpha^2 sn - \alpha^2 s^2}{r - \alpha s} - \alpha s \varepsilon \right\rfloor =$$
$$\left\lfloor \frac{\alpha r n - \alpha r s}{r - \alpha s} - \alpha s \varepsilon \right\rfloor = \frac{\alpha r n - \alpha r s}{r - \alpha s} - \alpha s \varepsilon - \varepsilon'$$

for some $0 \leq \varepsilon' < 1$.

Then $W_{\mathsf{A}} - rw = (r - \alpha s)\varepsilon - \varepsilon'$. We would like to identify an infinite sequence of n such that $W_{\mathsf{A}} = rw$. Since $r - \alpha s > 0$, we have $W_{\mathsf{A}} - rw > -1$. If $r - \alpha s \leq 1$, then also $W_{\mathsf{A}} - rw < 1$ and we are done. Otherwise, only consider $n = s + \lceil k \cdot \frac{r - \alpha s}{\alpha} \rceil$ for $k \in \mathbb{N}$. Then $n = s + k \cdot \frac{r - \alpha s}{\alpha} + \varepsilon''$ for some $0 \leq \varepsilon'' < 1$ and $\alpha(n - s) = k(r - \alpha s) + \alpha \varepsilon''$. From Equation 3,

$$\varepsilon = \frac{\alpha n - \alpha s}{r - \alpha s} - w = \frac{k(r - \alpha s) + \alpha \varepsilon''}{r - \alpha s} - w = k - w + \frac{\alpha \varepsilon''}{r - \alpha s}.$$

Since $k - w \in \mathbb{Z}$, $0 \leq \varepsilon < 1$ and $0 \leq \frac{\alpha \varepsilon''}{r - \alpha} < 1$, we have $\varepsilon = \frac{\alpha \varepsilon''}{r - \alpha s} < \frac{1}{r - \alpha s}$. Hence, $W_{\mathsf{A}} - rw < 1$ and $W_{\mathsf{A}} = rw$.

Consider a ticket assignment $\vec{t} = (t_1, ..., t_n)$ where $t_1 = ... = t_{n-s-1} = 0$ and $t_{n-s} = ... = t_n = 1$. Since the adversary can only corrupt r parties that have a ticket, the total number of tickets is $s + 1$ and $r < \beta(s + 1)$ by an assumption in the theorem statement, this is a valid solution of cost $s + 1$. Clearly, no better solution is possible since the adversary can always corrupt r tickets (assuming they exist) to get adversarial ratio $\geq r/s$. Hence, $\mathsf{OPT}_{\leq \alpha, < \beta}(\vec{w}) = s + 1$.

We will now prove a lower bound on the cost of algorithm's output. Take any valid ticket assignment $\vec{t} = (t_1, ..., t_n)$ and notice that at least one of t_1, ..., t_{n-s} must be non-zero; otherwise the adversary can corrupt r parties with the largest number of tickets to get

adversarial ratio $\geq r/s$. Since the algorithm chooses $\delta > 0$ and for all i outputs $t_i \leq \lfloor \delta w_i + c \rfloor$, we must have $\delta \geq 1 - c$. Hence, the algorithm outputs a ticket assignment of cost at least

$$s\big(\lfloor (1-c)w + c \rfloor - m\big) > s\big((1-c)w + c - 1 - m\big) >$$
$$s\left(\big(1-c\big)\frac{\alpha n - \alpha s - 1}{r - \alpha s} - (1-c) - m\right) = s\left(\big(1-c\big)\frac{\alpha n - r - 1}{r - \alpha s} - m\right) =$$
$$\big(1-c\big)\frac{\alpha n - r - 1}{r/s - \alpha} - ms. \qquad \blacktriangleleft$$

When β is rational, simply set r and s such that $\beta = r/s$ to apply the lower bound. Otherwise, we need to find r, s such that $\frac{r}{s+1} < \beta \leq \frac{r}{s} \leq 1$ or, equivalently, $1 \leq \frac{s}{r} \leq \frac{1}{\beta} < \frac{s+1}{r}$. Setting $s = \lfloor 1/\beta \rfloor$ and $r = 1$ satisfies the inequality.

We now discuss particular weight reduction algorithms. By a simple modification of an argument in [39], we know that Swiper that solves the $(\leq \alpha, < \beta)$-weight reduction problem always outputs a solution with at most $\frac{\alpha(1-\alpha)}{\beta-\alpha}n + 1$ tickets. Thus, this is also an upper bound on its approximation factor. Since Swiper outputs solutions $(t_1, ..., t_n)$ such that for each i, $t_i = \lfloor \delta w_i + \alpha \rfloor$ or $t_i = \lfloor \delta w_i + \alpha \rfloor - 1$, by Theorem 1, the worst case approximation factor of Swiper is $\frac{\alpha(1-\alpha)}{(s+1)(r/s-\alpha)}n - O(1)$ (for appropriate r and s). Notice, that for some values of β, this is very close to the upper bound; for instance, when $\beta = 1/2$, it is a factor 3 away from the upper bound. The same analysis applies to faster Swiper and extended Swiper.

We now discuss the gcdWR algorithm from [23] (also see an overview in the full version of our paper [42, Appendix A]). It is quite similar to Swiper in that it outputs a ticket assignment $(t_1, ..., t_n)$ of the form $t_i = \lfloor \delta w_i + c \rfloor$; here, $\delta = 1/\gcd$ and $c = 1/2$. Theorem 1 gives approximation factor lower bound of $\frac{\alpha}{2(s+1)(r/s-\alpha)}n - O(1)$. We conclude that the worst case approximation factor of gcdWR is $\Theta(n)$.

Finally, we discuss the algorithms in [21, v1] and [4]. Even though they solve other variants of the weight reduction problem, we can still see what approximation factor they would give if solutions of the same form were to be used for the $(\leq \alpha, < \beta)$-weight reduction problem. [21, v1] outputs solutions of the form $t_i = \lfloor \delta w_i + 1/2 \rfloor$. By Theorem 1, its worst case approximation factor exceeds $\frac{\alpha}{2(s+1)(r/s-\alpha)}n - O(1)$ (for appropriate r and s).

[4] outputs solutions of the form $t_i = \lceil \delta w_i \rceil$. Since each w_i is positive, each t_i also is and thus the algorithm always outputs a solution of cost at least n. We are left to find an example where the optimal solution is small. If there exists a weight w_i large enough that the adversary cannot corrupt it, then the optimal solution has one ticket in total. Hence, the worst case approximation factor is at least n.

7 Useful facts

In this section we present several facts that will be useful in the rest of the paper. We start with a (new) characterization of optimal solutions.

▶ **Theorem 2.** *Assume $w_1 \leq w_2 \leq ... \leq w_n$. There exists an optimal solution $(t_1, ..., t_n)$ such that $t_1 \leq t_2 \leq ... \leq t_n$.*

Proof. Let $(t_1, ..., t_n)$ be an optimal solution such that $t_i > t_j$ for some $i < j$. We will show that $(t'_1, ..., t'_n) = (t_1, ..., t_{i-1}, t_j, t_{i+1}, ..., t_{j-1}, t_i, t_{j+1}, ...t_n)$ is also an optimal solution (i.e., t_i and t_j are swapped). This has the same cost, we just need to show that it is a valid ticket assignment.

Let $I \subseteq [n]$ such that $\sum_{k \in I} w_k \leq \alpha \sum_{k=1}^{n} w_k$. Consider the following cases.

- $i \notin I$, $j \notin I$: $\sum_{k \in I} t'_k = \sum_{k \in I} t_k < \beta \sum_{k=1}^{n} t_k = \beta \sum_{k=1}^{n} t'_k$;

- $i \in I$, $j \in I$: $\sum_{k\in I} t'_k = \sum_{k\in I} t_k < \beta \sum_{k=1}^n t_k = \beta \sum_{k=1}^n t'_k$;
- $i \in I$, $j \notin I$: $\sum_{k\in I} t'_k < \sum_{k\in I} t_k < \beta \sum_{k=1}^n t_k = \beta \sum_{k=1}^n t'_k$;
- $i \notin I$, $j \in I$: let $I' = I \setminus \{j\} \cup \{i\}$, $\sum_{k\in I} t'_k = \sum_{k\in I'} t_k < \beta \sum_{k=1}^n t_k = \beta \sum_{k=1}^n t'_k$, where the inequality holds since $\sum_{k\in I'} w_k < \sum_{k\in I} w_k \le \alpha \sum_{k=1}^n w_k$.

Hence, $(t'_1, ..., t'_n)$ is a valid assignment.

Therefore, we can sort all t_i using, for example, insertion sort to get an optimal sorted assignment $(t_1, ..., t_n)$. ◀

We also prove a slightly stronger statement.

▶ **Theorem 3.** *Assume $w_1 \le w_2 \le ... \le w_n$. There exists an optimal solution $(t_1, ..., t_n)$ such that $t_1 \le t_2 \le ... \le t_n$ and, moreover, for any i, j with $w_i = w_j$, $|t_i - t_j| \le 1$.*

We present the proof on page 24 of the Appendix.

We will also find useful the following, first mentioned in [39, v2 appendix C]. Similarly to the definition of $\mathsf{OPT}_{\le\alpha,<\beta}(\vec{w})$ from section 3, define $\mathsf{OPT}^{T_\mathsf{A}}_{\le\alpha,<\beta}(\vec{w})$ to be the smallest maximum number of tickets that the adversary can obtain in a valid solution:

$$\mathsf{OPT}^{T_\mathsf{A}}_{\le\alpha,<\beta}(\vec{w}) = \min_{\substack{(t_1,...,t_n)\in\mathbb{Z}^n_{\ge 0}\\ \text{is valid}}} \max_{\substack{I\subseteq[n]\wedge\\ \sum_{i\in I} w_i \le \alpha \sum_{i=1}^n w_i}} \sum_{i\in I} t_i.$$

▶ **Theorem 4.** *For all $0 < \alpha < \beta < 1$ and $\vec{w} \in \mathbb{N}^n$, we have*

$$OPT_{\le\alpha,<\beta}(\vec{w}) = \left\lfloor \frac{OPT^{T_A}_{\le\alpha,<\beta}(\vec{w})}{\beta} \right\rfloor + 1; \quad OPT^{T_A}_{\le\alpha,<\beta}(\vec{w}) = \left\lceil \beta \cdot OPT_{\le\alpha,<\beta}(\vec{w}) \right\rceil - 1.$$

Proof. Denote $T = \mathsf{OPT}_{\le\alpha,<\beta}(\vec{w})$ and $T_\mathsf{A} = \mathsf{OPT}^{T_\mathsf{A}}_{\le\alpha,<\beta}(\vec{w})$.

We first prove that $T = \lfloor T_\mathsf{A}/\beta \rfloor + 1$. Let $(t_1, ..., t_n)$ be a valid ticket assignment where the maximum total number of tickets the adversary can obtain is T_A. Decrease all t_i as much as possible as long as $T_\mathsf{A} < \beta \sum_{i=1}^n t_i$; call the resulting ticket assignment $(t'_1, ..., t'_n)$, which must be valid. Then $T \le \sum_{i=1}^n t'_i = \lfloor T_\mathsf{A}/\beta \rfloor + 1$. Now take a valid ticket assignment $(t_1, ..., t_n)$ with $\sum_{i=1}^n t_i = T$. We have

$$T_\mathsf{A} \le \max_{\substack{I\subseteq[n]\wedge\\ \sum_{i\in I} w_i \le \alpha \sum_{i=1}^n w_i}} \sum_{i\in I} t_i < \beta T$$

and thus $T \ge \lfloor T_\mathsf{A}/\beta \rfloor + 1$. Hence, $T = \lfloor T_\mathsf{A}/\beta \rfloor + 1$.

We will now prove that $T_\mathsf{A} = \lceil \beta T \rceil - 1$. Since $T = \lfloor T_\mathsf{A}/\beta \rfloor + 1$, $T > T_\mathsf{A}/\beta$ which implies $T_\mathsf{A} < \beta T$ which implies $T_\mathsf{A} \le \lceil \beta T \rceil - 1$. Also since $T = \lfloor T_\mathsf{A}/\beta \rfloor + 1$, $T - 1 \le T_\mathsf{A}/\beta$ which implies $T_\mathsf{A} \ge \beta T - \beta \ge T_\mathsf{A} \ge \lceil \beta T \rceil - 1 - \beta$ which implies $T_\mathsf{A} \ge \lceil \beta T \rceil - 1$. ◀

Similarly, we have the following for the $(\le \alpha, \le \beta)$-weight reduction problem.

▶ **Theorem 5.** *For all $0 < \alpha < \beta < 1$ and $\vec{w} \in \mathbb{N}^n$, we have*

$$OPT_{\le\alpha,\le\beta}(\vec{w}) = \max\left\{1, \left\lceil \frac{OPT^{T_A}_{\le\alpha,\le\beta}(\vec{w})}{\beta} \right\rceil\right\}; \quad OPT^{T_A}_{\le\alpha,\le\beta}(\vec{w}) = \left\lfloor \beta \cdot OPT_{\le\alpha,\le\beta}(\vec{w}) \right\rfloor.$$

8 Optimized bruteforce

Given that all previously known weight reduction algorithms have a worst case approximation factor $\Omega(n)$, we seek algorithms with improved approximation. In this section we present the first sub-exponential exact algorithm that is practical for some real-world weight distributions, as well as the first polynomial time algorithm with approximation factor at most $n/(c\log^2 n)$ for any $c > 0$.

8.1 Sub-exponential time exact solution

In this section we show an algorithm that solves the $(\leq \alpha, < \beta)$-weight reduction problem exactly (i.e., that finds an optimal solution) in time $O\left(n + R^{3/2}e^{C\sqrt{R}}\right)$ where $C = \pi\sqrt{6}/3 \approx 2.57$ and $R \leq O(n)$ is the cost of the output (i.e., the optimal total number of tickets). Given that R is often relatively small and definitely smaller than n, this algorithm is sometimes practical for real-world distributions, as the numerical evaluations in section 9 show. Thus, when a small solution is expected, it might make sense to run the exact algorithm for a number of steps and switch to a fallback scheme if necessary. The algorithm could also be very useful for research purposes, for example, to test some hypothesis about the problem.

Assume the weights in the input are sorted. Using the characterization of optimal solutions in Theorem 2, the algorithm for each $i = 1, 2, ...,$ enumerates all sorted ticket assignments with total number i and for each, checks its validity using the knapsack algorithm. It outputs the first found valid ticket assignment.

The number of sorted ticket assignments with total number i is bounded by the number of ways $p(i)$ to represent i as a sum of positive integers, where p is called partition function[1]. It is known that $p(k) \sim \frac{1}{4\sqrt{3}k} \exp\left(\frac{\pi\sqrt{6k}}{3}\right)$. Also, the knapsack algorithm checks the validity of a ticket assignment in time $O(i^2)$ since the ticket assignment has at most i non-zero entries. Hence, the running time of this algorithm is $O\left(\sum_{i=1}^{R} i^2 p(i)\right)$ which is $O\left(R^{3/2}e^{C\sqrt{R}}\right)$ by the following lemma.

▶ **Lemma 6.** *For any $c > 0$, $\sum_{i=1}^{k} ie^{c\sqrt{i}} = O\left(k^{3/2}e^{c\sqrt{k}}\right)$.*

We present the proof on page 24 of the Appendix.

We note that it might be possible to reduce this time complexity by a factor of $R^{1/2}$ by reusing dynamic programming computations as in section 5. Lastly, see subsection 5.3 on how to handle unsorted inputs for an additive running time increase of $O(n + R\log R)$.

8.2 $O(n/\log^2 n)$-approximation in polynomial time

Using the idea in subsection 8.1 we can similarly construct an algorithm with approximation factor $\frac{n}{c\log^2 n}$ and polynomial running time $O\left(n^{\frac{\pi\sqrt{6}}{3}\sqrt{c}\log e\sqrt{\frac{\alpha(1-\alpha)}{\beta-\alpha}}}\log^3 n + n^2\right)$ for any $c > 0$. The algorithm works by trying all sorted ticket assignments with total number at most L, to be defined below, and if no valid assignment was found, proceeding to using faster Swiper as fallback.

By [39], Swiper always finds a solution with cost at most $B = \frac{\alpha(1-\alpha)}{\beta-\alpha}n + 1$. By setting $L = \left\lceil \frac{Bc\log^2 n}{n} \right\rceil - 1$, the approximation factor of the algorithm is at most $\frac{B}{L+1} \leq \frac{n}{c\log^2 n}$.

[1] https://en.wikipedia.org/wiki/Integer_partition

Similarly to subsection 8.1, the running time of the bruteforce is at most

$$O\left(n + L^{3/2}\exp\left(\frac{\pi\sqrt{6}}{3}\sqrt{L}\right)\right) = O\left(n + \log^3 n \cdot \exp\left(\frac{\pi\sqrt{6}}{3}\sqrt{c}\log n\sqrt{\frac{B}{n}}\right)\right) =$$
$$O\left(n + n^{\frac{\pi\sqrt{6}}{3}\sqrt{c}\log e\sqrt{\frac{B}{n}}} \cdot \log^3 n\right) = O\left(n + n^{\frac{\pi\sqrt{6}}{3}\sqrt{c}\log e\sqrt{\frac{\alpha(1-\alpha)}{\beta-\alpha}}} \log^3 n\right),$$

while the running time of faster Swiper is $O(n^2)$.

9 Numerical evaluation

We implemented algorithms from this paper and from the literature in the Rust programming language and evaluated them based on the total number of tickets in the output and the running time. The stake distributions of Algorand, Aptos, Filecoin and Tezos blockchains were used as inputs, borrowed from the original Swiper implementation's git repository[2]; they were used for evaluations in the Swiper paper [38, 39].

The implemented algorithms from our work are faster Swiper (subsection 5.1), extended Swiper (subsection 5.2) and the exact algorithm from subsection 8.1. The implemented algorithms from the literature are Swiper and gcdWR. Even though the original implementation of Swiper is publicly available, it is written in Python, a much slower programming language. To make the comparison fair, we reimplemented Swiper in Rust with the two implementations having identical output and almost identical logic; the only difference is that we did not implement the quasilinear time approximate knapsack solver that Swiper runs before invoking the full solver, but [39] states that this optimization improves the running time only by a factor of 3 and only when the number of tickets in the output is large. We include both implementations of Swiper in our evaluations. The original implementation in Python was slightly modified to solve the $(\leq \alpha, < \beta)$-weight reduction problem instead of the variant where the inequality involving α is strict. The LP based algorithm from Appendix A was not implemented due to its impractically large time complexity and the unavailability of an ellipsoid method implementation with a separation oracle, and the general technical complexity of such an implementation.

All our implementations, as well as the original Swiper code in Python, are single-threaded. We assume sorted weights in the input, and thus our faster Swiper implementation has time complexity $O(R^2)$ as opposed to $O(n+R^2)$ in the full version of the algorithm, with a similar difference for the extended Swiper and exact schemes. Even though, we didn't implement subsection 5.3 to handle unsorted inputs, we believe it does not significantly affect the measured running times since the stake distributions used all have less than 43000 entries. Our code is publicly available[3].

We ran the tests on AMD Threadripper PRO 7955WX. Tables 1, 2, 3 and 4 have two values for each evaluation: the total number of tickets and the running time. The exact algorithm was modified to search for a solution with the total number tickets up to 143. In case a solution wasn't found, the tables have a blank cell for the total number of tickets and the running time reported is the time duration before the search terminated.

[2] https://github.com/DCL-TelecomParis/swiper
[3] https://github.com/tolikzinovyev/weight-reduction

9.1 Discussion

In Table 1 we indeed see that our faster Swiper algorithm is faster than the original Swiper. The speedup on Algorand's stake distribution ranges from 17x when the cost of the output is large to 13000x when the output cost is small. The huge efficiency gap in the small output case can be attributed to expensive rational number arithmetic when linearly scaling weights down in Swiper (time complexity $O(n \log n)$), which can dominate even the total cost of knapsack (time complexity $O(n^2 \log n)$). In contrast, our faster Swiper implementation does only $O(R \log R)$ rational number operations (inside gen.next() – see subsection 5.1). The cost of the output is comparable between Swiper and faster Swiper, but not exactly the same, as one or the other can get stuck in "local minimums" depending on luck.

Next, we observe that our extended Swiper algorithm sometimes significantly improves the quality of the output over Swiper. For instance, for Aptos' stake distribution (Table 2), Swiper outputs 85 tickets for parameters $\alpha = 1/4, \beta = 1/3$ whereas extended Swiper outputs 58; for $\alpha = 3/10, \beta = 1/3$, Swiper and faster Swiper output 346 tickets whereas extended Swiper outputs 277. It is guaranteed that the output of extended Swiper is never worse than that of Swiper and faster Swiper; moreover, in our measurements it is always better than that of gcdWR and is always within 6% of the optimal solution (based on the exact algorithm when it is feasible). Where the gap between α and β is not too small, the running time of extended Swiper is not significantly larger than even that of faster Swiper.

Finally, it is worth noting that while the exact algorithm has sub-exponential running time, it can nonetheless be practical when the optimal solution is small (e.g., in the case of Aptos). It could thus be used in conjunction with a fallback scheme when the optimal solution turns out larger than a threshold.

10 Conclusion

In this paper we made further progress in solving the problem of deterministic weight reduction. On the practical side, faster Swiper is a very efficient algorithm with time complexity $O(n+R^2)$ that outputs solutions comparable to the optimal on real-world inputs while extended Swiper outputs solutions extremely close to optimal in a reasonable time $O(n + R^{2.5})$. Both would be great candidates for deployment in, e.g., proof-of-stake blockchains.

On the theoretical side, we proved that all known algorithms based on linear scaling suffer from $\Omega(n)$ approximation ratio in the worst case. We presented the first polynomial time algorithm with approximation factor $n/(c \log^2 n)$, for any $c > 0$, and the first exact algorithm with sub-exponential running time $O\left(n + R^{3/2} e^{C\sqrt{R}}\right)$. Given that R is often much smaller than n, our exact algorithm can also be practical for production or research purposes. Finally, our linear programming based algorithm could be useful in the context of future research.

It could be interesting to try integrating a quasilinear time approximate knapsack solver, as done in the quasilinear mode Swiper [39], to decrease the exponent of R by one in the time complexity of faster Swiper, extended Swiper and the exact scheme. Additionally, we would like to know if there exists an algorithm with approximation factor $o(n/\log^2 n)$. The answer is likely yes, but more exploration is required.

Finally, it would be good to investigate the related variants of committee selection where we require the weaker ("randomized") security from Equation 1 or where the goal is to minimize the number of distinct committee members (see section 2). [28] makes progress in the "randomized" setting presenting schemes performing very well in practice, but, unfortunately, nothing is said about their approximation factor. It would be interesting to prove that a fully deterministic solution can be optimal when the weaker security is required.

References

1 Ittai Abraham, Eli Chouatt, Ivan Damgård, Yossi Gilad, Gilad Stern, and Sophia Yakoubov. Asynchronous algorand: Reaching agreement with near linear communication and constant expected time. Cryptology ePrint Archive, Report 2025/303, 2025. URL: https://eprint.iacr.org/2025/303.

2 Georg Anegg, Haris Angelidakis, Adam Kurpisz, and Rico Zenklusen. A technique for obtaining true approximations for k-center with covering constraints. *Math. Program.*, 192(1–2):3–27, March 2022. doi:10.1007/s10107-021-01645-y.

3 Amos Beimel and Enav Weinreb. Monotone circuits for monotone weighted threshold functions. *Information Processing Letters*, 97(1):12–18, 2006. doi:10.1016/j.ipl.2005.09.008.

4 Fabrice Benhamouda, Shai Halevi, and Lev Stambler. Weighted secret sharing from wiretap channels. In Kai-Min Chung, editor, *ITC 2023*, volume 267 of *LIPIcs*, pages 8:1–8:19. Schloss Dagstuhl, June 2023. doi:10.4230/LIPIcs.ITC.2023.8.

5 Elette Boyle, Kai-Min Chung, and Rafael Pass. Large-scale secure computation: Multi-party computation for (parallel) RAM programs. In Rosario Gennaro and Matthew J. B. Robshaw, editors, *CRYPTO 2015, Part II*, volume 9216 of *LNCS*, pages 742–762. Springer, Berlin, Heidelberg, August 2015. doi:10.1007/978-3-662-48000-7_36.

6 Elette Boyle, Ran Cohen, Deepesh Data, and Pavel Hubácek. Must the communication graph of MPC protocols be an expander? *Journal of Cryptology*, 36(3):20, July 2023. doi:10.1007/s00145-023-09460-8.

7 Elette Boyle, Ran Cohen, and Aarushi Goel. Breaking the $O(\sqrt{n})$-bit barrier: Byzantine agreement with polylog bits per party. *Journal of Cryptology*, 37(1):2, January 2024. doi:10.1007/s00145-023-09484-0.

8 Elette Boyle, Shafi Goldwasser, and Stefano Tessaro. Communication locality in secure multi-party computation - how to run sublinear algorithms in a distributed setting. In Amit Sahai, editor, *TCC 2013*, volume 7785 of *LNCS*, pages 356–376. Springer, Berlin, Heidelberg, March 2013. doi:10.1007/978-3-642-36594-2_21.

9 Gabriel Bracha. An $O(\lg n)$ expected rounds randomized byzantine generals protocol. In *17th ACM STOC*, pages 316–326. ACM Press, May 1985. doi:10.1145/22145.22180.

10 Nicolas Braud-Santoni, Rachid Guerraoui, and Florian Huc. Fast byzantine agreement. In Panagiota Fatourou and Gadi Taubenfeld, editors, *32nd ACM PODC*, pages 57–64. ACM, July 2013. doi:10.1145/2484239.2484243.

11 Ran Canetti, Sebastian Kolby, Divya Ravi, Eduardo Soria-Vazquez, and Sophia Yakoubov. Taming adaptivity in YOSO protocols: The modular way. In Guy N. Rothblum and Hoeteck Wee, editors, *TCC 2023, Part II*, volume 14370 of *LNCS*, pages 33–62. Springer, Cham, November / December 2023. doi:10.1007/978-3-031-48618-0_2.

12 John F. Canny and Stephen Sorkin. Practical large-scale distributed key generation. In Christian Cachin and Jan Camenisch, editors, *EUROCRYPT 2004*, volume 3027 of *LNCS*, pages 138–152. Springer, Berlin, Heidelberg, May 2004. doi:10.1007/978-3-540-24676-3_9.

13 Robert D. Carr, Lisa Fleischer, Vitus J. Leung, and Cynthia A. Phillips. Strengthening integrality gaps for capacitated network design and covering problems. In David B. Shmoys, editor, *11th SODA*, pages 106–115. ACM-SIAM, January 2000. URL: http://dl.acm.org/citation.cfm?id=338219.338241.

14 Ignacio Cascudo, Bernardo David, Lydia Garms, and Anders Konring. YOLO YOSO: Fast and simple encryption and secret sharing in the YOSO model. In Shweta Agrawal and Dongdai Lin, editors, *ASIACRYPT 2022, Part I*, volume 13791 of *LNCS*, pages 651–680. Springer, Cham, December 2022. doi:10.1007/978-3-031-22963-3_22.

15 Pyrros Chaidos and Aggelos Kiayias. Mithril: Stake-based threshold multisignatures. In Markulf Kohlweiss, Roberto Di Pietro, and Alastair R. Beresford, editors, *CANS 2024, Part I*, volume 14905 of *LNCS*, pages 239–263. Springer, Singapore, September 2024. doi:10.1007/978-981-97-8013-6_11.

16 Pyrros Chaidos, Aggelos Kiayias, Leonid Reyzin, and Anatoliy Zinovyev. Approximate lower bound arguments. In Marc Joye and Gregor Leander, editors, *EUROCRYPT 2024, Part IV*, volume 14654 of *LNCS*, pages 55–84. Springer, Cham, May 2024. doi:10.1007/978-3-031-58737-5_3.

17 Jing Chen and Silvio Micali. Algorand: A secure and efficient distributed ledger. *Theoretical Computer Science*, 777:155–183, 2019. In memory of Maurice Nivat, a founding father of Theoretical Computer Science - Part I. doi:10.1016/j.tcs.2019.02.001.

18 Ran Cohen, Iftach Haitner, Eran Omri, and Lior Rotem. From fairness to full security in multiparty computation. *Journal of Cryptology*, 35(1):4, January 2022. doi:10.1007/s00145-021-09415-x.

19 Varsha Dani, Valerie King, Mahnush Movahedi, Jared Saia, and Mahdi Zamani. Secure multi-party computation in large networks. *Distrib. Comput.*, 30(3):193–229, June 2017. doi:10.1007/s00446-016-0284-9.

20 Sourav Das, Philippe Camacho, Zhuolun Xiang, Javier Nieto, Benedikt Bünz, and Ling Ren. Threshold signatures from inner product argument: Succinct, weighted, and multi-threshold. In Weizhi Meng, Christian Damsgaard Jensen, Cas Cremers, and Engin Kirda, editors, *ACM CCS 2023*, pages 356–370. ACM Press, November 2023. doi:10.1145/3576915.3623096.

21 Sourav Das, Benny Pinkas, Alin Tomescu, and Zhuolun Xiang. Distributed randomness using weighted VRFs. Cryptology ePrint Archive, Report 2024/198, 2024. URL: https://eprint.iacr.org/2024/198.

22 Sourav Das, Benny Pinkas, Alin Tomescu, and Zhuolun Xiang. Distributed randomness using weighted VUFs. In Serge Fehr and Pierre-Alain Fouque, editors, *EUROCRYPT 2025, Part VII*, volume 15607 of *LNCS*, pages 314–344. Springer, Cham, May 2025. doi:10.1007/978-3-031-91098-2_12.

23 Hanwen Feng, Tiancheng Mai, and Qiang Tang. Scalable and adaptively secure any-trust distributed key generation and all-hands checkpointing. In Bo Luo, Xiaojing Liao, Jun Xu, Engin Kirda, and David Lie, editors, *ACM CCS 2024*, pages 2636–2650. ACM Press, October 2024. doi:10.1145/3658644.3690253.

24 Nils Fleischhacker, Mathias Hall-Andersen, Mark Simkin, and Benedikt Wagner. Jackpot: Non-interactive aggregatable lotteries. In Kai-Min Chung and Yu Sasaki, editors, *ASIACRYPT 2024, Part VI*, volume 15489 of *LNCS*, pages 365–397. Springer, Singapore, December 2024. doi:10.1007/978-981-96-0938-3_12.

25 G.N. Frederickson. An optimal algorithm for selection in a min-heap. *Information and Computation*, 104(2):197–214, 1993. doi:10.1006/inco.1993.1030.

26 Sanjam Garg, Abhishek Jain, Pratyay Mukherjee, Rohit Sinha, Mingyuan Wang, and Yinuo Zhang. Cryptography with weights: MPC, encryption and signatures. In Helena Handschuh and Anna Lysyanskaya, editors, *CRYPTO 2023, Part I*, volume 14081 of *LNCS*, pages 295–327. Springer, Cham, August 2023. doi:10.1007/978-3-031-38557-5_10.

27 Sanjam Garg, Abhishek Jain, Pratyay Mukherjee, Rohit Sinha, Mingyuan Wang, and Yinuo Zhang. hinTS: Threshold signatures with silent setup. In *2024 IEEE Symposium on Security and Privacy*, pages 3034–3052. IEEE Computer Society Press, May 2024. doi:10.1109/SP54263.2024.00057.

28 Peter Gazi, Aggelos Kiayias, and Alexander Russell. Fait accompli committee selection: Improving the size-security tradeoff of stake-based committees. In Weizhi Meng, Christian Damsgaard Jensen, Cas Cremers, and Engin Kirda, editors, *ACM CCS 2023*, pages 845–858. ACM Press, November 2023. doi:10.1145/3576915.3623194.

29 Craig Gentry, Shai Halevi, Hugo Krawczyk, Bernardo Magri, Jesper Buus Nielsen, Tal Rabin, and Sophia Yakoubov. YOSO: You only speak once - secure MPC with stateless ephemeral roles. In Tal Malkin and Chris Peikert, editors, *CRYPTO 2021, Part II*, volume 12826 of *LNCS*, pages 64–93, Virtual Event, August 2021. Springer, Cham. doi:10.1007/978-3-030-84245-1_3.

30 Yossi Gilad, Rotem Hemo, Silvio Micali, Georgios Vlachos, and Nickolai Zeldovich. Algorand: Scaling byzantine agreements for cryptocurrencies. In *Proceedings of the 26th Symposium on*

Operating Systems Principles, SOSP '17, pages 51–68, New York, NY, USA, 2017. Association for Computing Machinery. doi:10.1145/3132747.3132757.

31 Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2. Springer Science & Business Media, 2012. doi:10.1007/978-3-642-78240-4.

32 D. Ellis Hershkowitz, Nathan Klein, and Rico Zenklusen. Ghost value augmentation for k-edge-connectivity. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *56th ACM STOC*, pages 1853–1864. ACM Press, June 2024. doi:10.1145/3618260.3649715.

33 Bruce M. Kapron, David Kempe, Valerie King, Jared Saia, and Vishal Sanwalani. Fast asynchronous byzantine agreement and leader election with full information. *ACM Trans. Algorithms*, 6(4), September 2010. doi:10.1145/1824777.1824788.

34 Valerie King, Steven Lonargan, Jared Saia, and Amitabh Trehan. Load balanced scalable byzantine agreement through quorum building, with full information. In Marcos K. Aguilera, Haifeng Yu, Nitin H. Vaidya, Vikram Srinivasan, and Romit Roy Choudhury, editors, *Distributed Computing and Networking*, pages 203–214, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg. doi:10.1007/978-3-642-17679-1_18.

35 Valerie King and Jared Saia. From almost everywhere to everywhere: Byzantine agreement with $\tilde{O}(n^{3/2})$ bits. In Idit Keidar, editor, *Distributed Computing*, pages 464–478, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. doi:10.1007/978-3-642-04355-0_47.

36 Valerie King and Jared Saia. Breaking the $O(n^2)$ bit barrier: Scalable byzantine agreement with an adaptive adversary. *J. ACM*, 58(4), July 2011. doi:10.1145/1989727.1989732.

37 Valerie King, Jared Saia, Vishal Sanwalani, and Erik Vee. Scalable leader election. In *17th SODA*, pages 990–999. ACM-SIAM, January 2006. URL: http://dl.acm.org/citation.cfm?id=1109557.1109667.

38 Andrei Tonkikh and Luciano Freitas de Souza. Swiper: a new paradigm for efficient weighted distributed protocols. In Ran Gelles, Dennis Olivetti, and Petr Kuznetsov, editors, *43rd ACM PODC*, pages 283–294. ACM, June 2024. doi:10.1145/3662158.3662799.

39 Andrei Tonkikh and Luciano Freitas. Swiper: a new paradigm for efficient weighted distributed protocols. Cryptology ePrint Archive, Report 2023/1164, 2023. URL: https://eprint.iacr.org/2023/1164.

40 David P. Williamson and David B. Shmoys. *The Design of Approximation Algorithms*. Cambridge University Press, USA, 1st edition, 2011. doi:10.1017/CBO9780511921735.

41 Anatoliy Zinovyev. Weight reduction. Software, DARPA, swhId: swh:1:dir:a48c33feba04bc98c7013c4c88b174f5e461a392 (visited on 2025-09-05). URL: https://github.com/tolikzinovyev/weight-reduction/tree/f1b5ce44aab9e4809bff67c5c6bff4acf1f82cb1, doi:10.4230/artifacts.24587.

42 Anatoliy Zinovyev. Weight reduction in distributed protocols: new algorithms and analysis. Cryptology ePrint Archive, Paper 2025/1076, 2025. URL: https://eprint.iacr.org/2025/1076.

A Linear programming based polynomial time approximation

In this section we show that for any constant $0 < \delta < 1 - \alpha/\beta$ there exists a polynomial time approximation algorithm that returns a solution with cost at most

$$\left\lfloor \frac{\mathsf{OPT}^{T_A}_{\leq\alpha,<(1-\delta)\beta}}{\beta} \right\rfloor + 1 = \left\lfloor (1-\delta) \cdot \frac{\mathsf{OPT}^{T_A}_{\leq\alpha,<(1-\delta)\beta}}{(1-\delta)\beta} \right\rfloor + 1 \leq \left\lfloor (1-\delta) \cdot \mathsf{OPT}_{\leq\alpha,<(1-\delta)\beta} \right\rfloor + 1 \leq \mathsf{OPT}_{\leq\alpha,<(1-\delta)\beta}, \tag{4}$$

where the first inequality is due to Theorem 4.

Table 1 Evaluation on the Algorand stake distribution.

	$\beta = 1/3$			$\beta = 1/2$			
	$\alpha = 0.2$	$\alpha = 0.25$	$\alpha = 0.3$	$\alpha = 0.3$	$\alpha = 0.35$	$\alpha = 0.4$	$\alpha = 0.45$
Swiper (Python)	235	745	22393	203	383	1203	17591
	4.75s	5.14s	6.25s	4.32s	5.15s	4.74s	5.70s
Swiper (Rust)	235	745	22393	203	383	1203	17591
	1.10s	1.21s	1.77s	1.03s	1.14s	1.18s	1.64s
faster Swiper	232	745	22660	203	387	1205	17949
	84.2µs	516µs	63.4ms	108µs	272µs	1.90ms	97.2ms
extended Swiper	232	745	22384	201	383	1171	17581
	98.9µs	849µs	3.48s	136µs	497µs	4.56ms	4.52s
gcdWR	1180	6237	178906	223	407	1180	17648
	3.09ms	2.85ms	2.42ms	3.29ms	3.27ms	3.11ms	2.66ms
exact							
	6.26h	4.50h	4.45h	8.72h	8.74h	8.70h	9.25h

Table 2 Evaluation on the Aptos stake distribution.

	$\beta = 1/3$			$\beta = 1/2$			
	$\alpha = 0.2$	$\alpha = 0.25$	$\alpha = 0.3$	$\alpha = 0.3$	$\alpha = 0.35$	$\alpha = 0.4$	$\alpha = 0.45$
Swiper (Python)	22	85	346	23	29	95	279
	5.37ms	7.32ms	6.94ms	4.87ms	6.19ms	6.76ms	7.06ms
Swiper (Rust)	22	85	346	23	29	95	279
	1.61ms	2.07ms	1.97ms	1.40ms	1.87ms	1.80ms	1.98ms
faster Swiper	22	58	346	23	31	95	279
	7.93µs	13.7µs	120µs	8.02µs	8.34µs	32.8µs	160µs
extended Swiper	22	58	277	23	29	95	241
	5.17µs	13.7µs	102µs	5.22µs	8.14µs	32.6µs	147µs
gcdWR	138	246	579	34	34	138	321
	7.96µs	7.85µs	7.86µs	8.46µs	8.68µs	7.98µs	7.98µs
exact	22	55		23	29	91	
	140µs	130ms	4.46h	227µs	1.14ms	47.5s	8.75h

Table 3 Evaluation on the Filecoin stake distribution.

	$\beta = 1/3$			$\beta = 1/2$			
	$\alpha = 0.2$	$\alpha = 0.25$	$\alpha = 0.3$	$\alpha = 0.3$	$\alpha = 0.35$	$\alpha = 0.4$	$\alpha = 0.45$
Swiper (Python)	1282	3091	10288	1149	1835	3533	8133
	281ms	324ms	388ms	287ms	301ms	322ms	430ms
Swiper (Rust)	1282	3091	10288	1149	1835	3533	8133
	29.6ms	40.2ms	98.2ms	31.7ms	35.8ms	49.8ms	129ms
faster Swiper	1282	3088	10288	1149	1837	3533	8139
	2.72ms	8.21ms	65.1ms	3.46ms	4.52ms	11.8ms	37.9ms
extended Swiper	1279	3088	10282	1147	1835	3527	8133
	3.52ms	23.4ms	305ms	5.15ms	18.2ms	67.5ms	399ms
gcdWR	3700	6516	17425	1289	1882	3700	8190
	166µs	162µs	158µs	169µs	170µs	166µs	162µs
exact							
	4.49h	4.48h	4.49h	8.51h	9.05h	8.56h	8.57h

Table 4 Evaluation on the Tezos stake distribution.

	$\beta = 1/3$			$\beta = 1/2$			
	$\alpha = 0.2$	$\alpha = 0.25$	$\alpha = 0.3$	$\alpha = 0.3$	$\alpha = 0.35$	$\alpha = 0.4$	$\alpha = 0.45$
Swiper (Python)	49	133	589	37	75	143	455
	28.5ms	28.3ms	31.3ms	20.1ms	29.2ms	25.8ms	28.9ms
Swiper (Rust)	49	133	589	37	75	143	455
	3.35ms	3.79ms	4.59ms	2.72ms	3.97ms	3.43ms	4.26ms
faster Swiper	49	133	589	41	77	145	455
	11.9µs	55.0µs	383µs	12.9µs	29.7µs	66.8µs	188µs
extended Swiper	37	127	523	35	75	143	427
	6.25µs	36.4µs	533µs	6.77µs	21.8µs	54.0µs	597µs
gcdWR	148	333	1331	46	87	148	470
	13.7µs	12.7µs	11.3µs	14.7µs	13.7µs	13.7µs	12.2µs
exact	37	127		35	75	139	
	6.44ms	1.05h	5.07h	5.36ms	6.64s	6.33h	8.62h

The algorithm is based on the standard approach of representing an optimization problem as an integer linear program, relaxing it to allow fractional solutions, solving the relaxed LP in polynomial time and rounding the solution to an integer vector in a problem specific manner [40]. We will use the ellipsoid method to solve an LP with exponentially many constraints. In this mode of operation, we are required to implement a "separation oracle" that the ellipsoid method will invoke. On input some vector $\vec{x} \in \mathbb{Q}^n$ in the multidimensional space, the separation oracle either reports that $\vec{x}$ satisfies all LP constraints or returns some violated constraint. Whenever the separation oracle is implemented by a polynomial time algorithm, the ellipsoid method also runs in polynomial time; in particular, its running time is independent of the number of LP constraints. See [40, section 4.3] for an overview of the ellipsoid method with a separation oracle.

For our problem, however, we face an issue that we cannot implement a polynomial time separation oracle that calculates the prescribed output exactly. The input will be a vector $\vec{t} \in \mathbb{Q}^n$ and we need to check if it satisfies the criteria of a valid ticket assignment. While we can test a ticket assignment $(t_1, ..., t_n) \in \mathbb{Z}^n_{\geq 0}$ for validity in time $O\big(n \cdot \sum_{i=1}^n t_i\big)$, we do not know how to efficiently do so for a fractional $\vec{t}$. We first present a sketch of our weight reduction algorithm that uses an inefficient separation oracle and in subsection A.2 we show how to get around this issue.

A.1 Sketch

Assume $w_1 \leq w_2 \leq ... \leq w_n$ and define $k = \lceil 1/\delta \rceil$. We first efficiently bruteforce the set of indices i with $t_i = g$ for each $0 \leq g \leq k-2$, having Theorem 2 in mind. Now consider the following satisfiability LP with some constant T, some t_i fixed by the "guess" and the other t_i constrained to be at least $k-1$.

$$\begin{aligned}
&\text{minimize} && 0 && \\
&\text{subject to} && \sum_{i=1}^{n} t_i = T && \\
& && \sum_{i \in I} t_i \leq \big\lfloor (1-1/k)\beta T \big\rfloor && \forall I \subseteq [n] \text{ s.t. } \sum_{i \in I} w_i \leq \alpha \sum_{i=1}^{n} w_i \\
& && t_i = g_i && \forall i : 1 \leq i < j \\
& && t_i \geq k-1 && \forall i : j \leq i \leq n
\end{aligned} \tag{5}$$

For each $T \in \{1, 2, ..., B\}$, where $B = O(n)$ is a bound on $\mathsf{OPT}_{\leq \alpha, < (1-1/k)\beta}$ from [39], if there is a solution $(t'_1, ..., t'_n)$ to this LP, round it down to get $(t_1, ..., t_n)$, let $T_{\mathsf{A}} = \big\lfloor (1-1/k)\beta T \big\rfloor$, reduce all t_i as much as possible as long as $T_{\mathsf{A}} < \beta \sum_{i=1}^n t_i$ to get $(t^+_1, ..., t^+_n)$ and record $(t^+_1, ..., t^+_n)$ as the best result for the current guess $(g_1, ..., g_{j-1})$. At the end of the algorithm, return the smallest result across all guesses.

We will now prove the bound in Equation 4 on the cost of the returned solution. By Theorem 2, fix an optimal solution $(t^*_1, ..., t^*_n)$ with $t^*_1 \leq t^*_2 \leq ... \leq t^*_n$ to the $(\leq \alpha, \leq (1-1/k)\beta)$-weight reduction problem. By definition, $\sum_{i=1}^n t^*_i = \mathsf{OPT}_{\leq \alpha, \leq (1-1/k)\beta} \leq \mathsf{OPT}_{\leq \alpha, < (1-1/k)\beta}$. Let $(g_1, ..., g_{j-1}) \in \{0, ..., k-2\}^{j-1}$ be the guess such that $(g_1, ..., g_{j-1}) = (t^*_1, ..., t^*_{j-1})$ and $t^*_i \geq k-1$ for $i \geq j$.

For this guess, the algorithm will terminate the search on some T at most $\mathsf{OPT}_{\leq \alpha, \leq (1-1/k)\beta}$. This is because if $T = \mathsf{OPT}_{\leq \alpha, \leq (1-1/k)\beta}$, then $(t^*_1, ..., t^*_n)$ is a valid solution to the above LP. Therefore, the algorithm will find a solution $(t'_1, ..., t'_n)$ to the LP with $\sum_{i=1}^n t'_i = T \leq \mathsf{OPT}_{\leq \alpha, \leq (1-1/k)\beta}$.

For any $I \subseteq [n]$ that the adversary is allowed to corrupt,

$$\sum_{i \in I} t_i \leq \sum_{i \in I} t'_i \leq \lfloor (1-1/k)\beta T \rfloor = T_{\mathsf{A}} = \left\lfloor (1-1/k)\beta \sum_{i=1}^{n} t'_i \right\rfloor \leq (1-1/k)\beta \sum_{i=1}^{n} t'_i < \beta \sum_{i=1}^{n} t_i,$$

where the last inequality holds by the assumption that every t'_i is either an integer or at least $k-1$, and that at least one t'_i is non-zero. Therefore, $(t_1^+, ..., t_n^+)$ is a valid solution to the $(\leq \alpha, < \beta)$-weight reduction problem. Moreover, $T_{\mathsf{A}} \leq \lfloor (1-1/k)\beta \cdot \mathsf{OPT}_{\leq \alpha, \leq (1-1/k)\beta} \rfloor = \mathsf{OPT}^{T_{\mathsf{A}}}_{\leq \alpha, \leq (1-1/k)\beta} \leq \mathsf{OPT}^{T_{\mathsf{A}}}_{\leq \alpha, < (1-\delta)\beta}$, where the equality is due to Theorem 5. Hence,

$$\sum_{i=1}^{n} t_i^+ = \left\lfloor \frac{T_{\mathsf{A}}}{\beta} \right\rfloor + 1 \leq \left\lfloor \frac{\mathsf{OPT}^{T_{\mathsf{A}}}_{\leq \alpha, < (1-\delta)\beta}}{\beta} \right\rfloor + 1.$$

A.2 Full algorithm

We now show how to circumvent the issue that we cannot construct a polynomial time separation oracle for the LP in Equation 5. We utilize the round-or-cut framework [13, 2] where at each iteration of the ellipsoid method, we either produce a separating hyperplane between the candidate feasible solution $\vec{t} \in \mathbb{Q}^n$ and the polytope or return another point $\vec{t'}$ that is slightly different from $\vec{t}$ but that will suffice for our weight reduction algorithm.

Configure the ellipsoid method with the LP in Equation 5 (see details in [31]) and implement the separation oracle as follows. On input $(t_1, ..., t_n) \in \mathbb{Q}^n$, check the constraints in lines 1, 3 and 4 in Equation 5 as usual; instead of the constraints in line 2, check relaxed constraints

$$\sum_{i \in I} \lfloor t_i \rfloor \leq \lfloor (1-1/k)\beta T \rfloor \quad \forall I \subseteq [n] \text{ s.t. } \sum_{i \in I} w_i \leq \alpha \sum_{i=1}^{n} w_i.$$

using knapsack in polynomial time. If some check failed, we return one of constraints from the original LP in Equation 5 as a separating hyperplane; otherwise, we terminate the search and return $(t_1, ..., t_n)$ as the solution.

The result of this search will be either a point $(t_1, ..., t_n)$ that satisfies

$$\begin{aligned}
&\sum_{i=1}^{n} t_i = T \\
&\sum_{i \in I} \lfloor t_i \rfloor \leq \lfloor (1-1/k)\beta T \rfloor \quad && \forall I \subseteq [n] \text{ s.t. } \sum_{i \in I} w_i \leq \alpha \sum_{i=1}^{n} w_i \\
&t_i = g_i && \forall i : 1 \leq i < j \\
&t_i \geq k-1 && \forall i : j \leq i \leq n
\end{aligned}$$

or an error certifying that the LP in Equation 5 has no solution. The rest of the algorithm and the proof of correctness remains the same.

A.3 Time complexity

We briefly argue that the algorithm runs in polynomial time. Using Theorem 2, bruteforcing the set of indices i with $t_i = j$ for each $0 \leq j \leq k-2$ takes time $O(n^{k-1})$. Bruteforcing T adds another factor of $O(n)$ by construction. Finally, solving the LP takes polynomial time.

Despite this, the algorithm is completely impractical since for reasonable parameters, the running time would be $\Omega(n^{10})$.

A.4 Approximation factor discussion

We have already proved that the algorithm finds a solution with cost at most

$$\left\lfloor \frac{\mathsf{OPT}^{T_\mathsf{A}}_{\leq\alpha,<(1-\delta)\beta}}{\beta} \right\rfloor + 1.$$

Using Theorem 4, the approximation factor for the total number of tickets is at most

$$\frac{\left\lfloor \frac{\mathsf{OPT}^{T_\mathsf{A}}_{\leq\alpha,<(1-\delta)\beta}}{\beta} \right\rfloor + 1}{\left\lfloor \frac{\mathsf{OPT}^{T_\mathsf{A}}_{\leq\alpha,<\beta}}{\beta} \right\rfloor + 1} < \frac{\frac{\mathsf{OPT}^{T_\mathsf{A}}_{\leq\alpha,<(1-\delta)\beta}}{\beta} + 1}{\frac{\mathsf{OPT}^{T_\mathsf{A}}_{\leq\alpha,<\beta}}{\beta}} = \frac{\mathsf{OPT}^{T_\mathsf{A}}_{\leq\alpha,<(1-\delta)\beta}}{\mathsf{OPT}^{T_\mathsf{A}}_{\leq\alpha,<\beta}} + \frac{\beta}{\mathsf{OPT}^{T_\mathsf{A}}_{\leq\alpha,<\beta}}.$$

While $\beta/\mathsf{OPT}^{T_\mathsf{A}}_{\leq\alpha,<\beta}$ can be made as small as $1/\log^2 n$ by subsection 8.2, in the full version of our paper [42, Appendix B] we show that, unfortunately, $\mathsf{OPT}^{T_\mathsf{A}}_{\leq\alpha,<(1-\delta)\beta}/\mathsf{OPT}^{T_\mathsf{A}}_{\leq\alpha,<\beta}$ can be $\Omega(n)$ in the worst case.

B Proofs

Proof of Theorem 3. Let $(t_1, ..., t_n)$ be any solution and take any i, j such that $w_i = w_j$ and $t_i + 2 \leq t_j$. Define $(t'_1, ..., t'_n)$ such that $t'_i = t_i + 1$, $t'_j = t_j - 1$ and $t'_k = t_k$ for all other k. We will show that this is also a valid solution.

Let $I \subseteq [n]$ such that $\sum_{k\in I} w_k \leq \alpha \sum_{k=1}^n w_k$ and consider the following cases.

- $i \notin I$, $j \notin I$: $\sum_{k\in I} t'_k = \sum_{k\in I} t_k < \beta \sum_{k=1}^n t_k = \beta \sum_{k=1}^n t'_k$;
- $i \in I$, $j \in I$: $\sum_{k\in I} t'_k = \sum_{k\in I} t_k < \beta \sum_{k=1}^n t_k = \beta \sum_{k=1}^n t'_k$;
- $i \notin I$, $j \in I$: $\sum_{k\in I} t'_k < \sum_{k\in I} t_k < \beta \sum_{k=1}^n t_k = \beta \sum_{k=1}^n t'_k$;
- $i \in I$, $j \notin I$: let $I' = I \setminus \{i\} \cup \{j\}$, $\sum_{k\in I} t'_k < \sum_{k\in I'} t_k < \beta \sum_{k=1}^n t_k = \beta \sum_{k=1}^n t'_k$, where the inequality holds since $\sum_{k\in I'} w_k = \sum_{k\in I} w_k \leq \alpha \sum_{k=1}^n w_k$.

Hence, $(t'_1, ..., t'_n)$ is a valid assignment.

By Theorem 2, take an optimal solution $(t_1, ..., t_n)$ with $t_1 \leq t_2 \leq ... \leq t_n$. One can see that we can iteratively achieve $|t_i - t_j| \leq 1$ for any block of indices $\{l, l+1, ..., r\}$ with equal weight, while preserving the sortedness of the whole array. ◀

Proof of Theorem 6.

$$\sum_{i=1}^{k} i e^{c\sqrt{i}} \leq \int_1^{k+1} x e^{c\sqrt{x}}\, dx\ [=]$$

letting $y = c\sqrt{x}$,

$$[=] \int_{y=c}^{y=c\sqrt{k+1}} \left(\frac{y}{c}\right)^2 e^y\, d\left(\left(\frac{y}{c}\right)^2\right) = \int_c^{c\sqrt{k+1}} \left(\frac{y}{c}\right)^2 e^y \frac{2y}{c^2}\, dy =$$

$$\frac{2}{c^4} \int_c^{c\sqrt{k+1}} y^3 e^y\, dy = \frac{2}{c^4} \int_{y=c}^{y=c\sqrt{k+1}} y^3\, d(e^y)\ [=]$$

letting $z = e^y$, for sufficiently large k,

$$[=] \frac{2}{c^4} \int_{e^c}^{e^{c\sqrt{k+1}}} \ln^3 z\, dz = \frac{2}{c^4} \Big(z\big(\ln^3 z - 3\ln^2 z + 6\ln z - 6\big) \Big) \Big|_{e^c}^{e^{c\sqrt{k+1}}} \leq$$

$$\frac{2}{c^4} \cdot e^{c\sqrt{k+1}} \cdot \Big(c\sqrt{k+1}\Big)^3 = \frac{2}{c} \cdot (k+1)^{3/2} \cdot e^{c\sqrt{k+1}} = O\Big(k^{3/2} e^{c\sqrt{k}}\Big).$$

◀