Revisiting the Random Subset Sum Problem

Arthur Carvalho Walraven Da Cunha
Université Côte d'Azur, Inria Sophia Antipolis, CNRS, France

Francesco d'Amore
Aalto University, Finland
Université Côte d'Azur, Inria Sophia Antipolis, CNRS, France

Frédéric Giroire
Université Côte d'Azur, Inria Sophia Antipolis, CNRS, France

Hicham Lesfari
Université Côte d'Azur, Inria Sophia Antipolis, CNRS, France

Emanuele Natale
Université Côte d'Azur, Inria Sophia Antipolis, CNRS, France

Laurent Viennot
Inria Paris, IRIF, France

Abstract

The average properties of the well-known *Subset Sum Problem* can be studied by means of its randomised version, where we are given a target value z, random variables $X_1, \ldots, X_n$, and an error parameter $\varepsilon > 0$, and we seek a subset of the X_is whose sum approximates z up to error ε. In this setup, it has been shown that, under mild assumptions on the distribution of the random variables, a sample of size $\mathcal{O}(\log(1/\varepsilon))$ suffices to obtain, with high probability, approximations for all values in $[-1/2, 1/2]$. Recently, this result has been rediscovered outside the algorithms community, enabling meaningful progress in other fields. In this work, we present an alternative proof for this theorem, with a more direct approach and resourcing to more elementary tools.

2012 ACM Subject Classification Mathematics of computing → Stochastic processes; Mathematics of computing → Combinatoric problems

Keywords and phrases Random subset sum, Randomised method, Subset-sum, Combinatorics

Digital Object Identifier 10.4230/LIPIcs.ESA.2023.37

Related Version *arXiv Preprint*: `https://arxiv.org/abs/2204.13929`

Funding *Frédéric Giroire*: Supported by the French government through the UCA JEDI (ANR-15-IDEX-01) and EUR DS4H (ANR-17-EURE-004) Investments in the Future projects, by the PEPR CARECLOUD, and by the European Project dAIEDGE.
Emanuele Natale: Supported by the AID INRIA-DGA agreement n°2019650072.

Acknowledgements The authors are thankful to Bianca C. Araújo and the paper reviewers for the feedback on the presentation of the paper. The authors are also grateful to the OPAL infrastructure from Université Côte d'Azur for providing resources and support.

1 Introduction

In the *Subset Sum Problem (SSP)*, one is given as input a set of n integers $X = \{x_1, x_2, \ldots, x_n\}$ and a target value z, and wishes to decide if there exists a subset of X that sums to z. That is, one is to reason about a subset $S \subseteq [n]$ such that $\sum_{i \in S} x_i = z$. The special case where z is half of the sum of X is known as the *Number Partition Problem (NPP)*. The converse reduction is also rather immediate.[1]

[1] To find a subset of X summing to z, one only needs to solve the NPP for the set $X \cup \{2z, \sum_{i\in[n]} x_i\}$. By doing so, one of the parts must consist of the element $\sum_{i\in[n]} x_i$ alongside the desired subset.

31st Annual European Symposium on Algorithms (ESA 2023).
Editors: Inge Li Gørtz, Martin Farach-Colton, Simon J. Puglisi, and Grzegorz Herman; Article No. 37; pp. 37:1–37:11
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Be it in either of these forms, the SSP finds applications in a variety of fields, ranging from combinatorial number theory [31] to cryptography [20, 26]. In complexity theory, the SSP is a well-known NP-complete problem, being a common base for NP-completeness proofs. In fact, the NPP version figures among Garey and Johnson's six basic NP-hard problems [19]. Under certain circumstances, the SSP can be challenging even for heuristics that perform well for many other NP-hard problems [25, 30], and a variety of dedicated algorithms have been proposed to solve it [10, 17, 22–24]. Nonetheless, it is not hard to solve it in polynomial time if we restrict the input integers to a fixed range [5]. It suffices to recursively list all achievable sums using the first i integers: we start with $A_0 = \{0\}$ and compute A_{i+1} as $A_i \cup \{a + x_{i+1} \mid a \in A_i\}$. For integers in the range $[0, R]$, the search space has size $\mathcal{O}(nR)$.

Studying how the problem becomes hard as we consider larger ranges of integers (relative to n) requires a randomised version of the problem, the *Random Subset Sum Problem (RSSP)*, where the input values are taken as independently and identically distributed random variables. In this setup, the work [6] proved that the problem experiences a phase transition in its average complexity as the range of integers increases.

The result we approach in this work comes from related studies on the typical properties of the problem. In [27] the author proves that, under fairly general conditions, the expected minimal distance between a subset-sum and the target value is exponentially small. More specifically, they show the following result.

▶ **Theorem 1** (Lueker, 1998). *Let $X_1, \ldots, X_n$ be independent uniform random variables over $[-1, 1]$, and let $\varepsilon \in (0, 1/3)$. There exists a universal constant $C > 0$ such that, if $n \geq C \log(1/\varepsilon)$, then, with probability at least $1 - \varepsilon$, for all $z \in [-1, 1]$ there exists $S_z \subseteq [n]$ for which*

$$\left| z - \sum_{i \in S_z} X_i \right| \leq \varepsilon.$$

That is, a rather small number (of the order of $\log \frac{1}{\varepsilon}$) of random variables suffices to have a high probability of approximating not only a single target z, but all values in an interval.

Even though Theorem 1 is stated and proved for uniform random variables and target values in $[-1, 1]$, it is not hard to extend the result to a broad class of distributions[2] and a wider range of targets. This generality makes the theorem a powerful tool for the analysis of random structures and has recently proven to be particularly useful in the field of Machine Learning, taking part in a proof of the Strong Lottery Ticket Hypothesis [29] and in subsequent related works [11, 13, 14, 18], and in Federated Learning [32].

Generalisations of the RSSP have played important roles in the study of random Knapsack problems [3,4], and to random binary integer programs [7,8]. In particular, the works [2,7,8,14] recently provided an extension of Theorem 1 to multiple dimensions. As for the equivalent Random Number Partitioning Problem, [12] recently generalised [6] and the integer version of the RSSP to non-binary integer coefficients.

The simplicity and ubiquity of the SSP have granted the related results a special didactic place. Be it as a first example of an NP-complete problem [19], a path to science communication [21], or simply as a frame for the demonstration of advanced techniques [28], it has been a tool to make important, but sometimes complicated, ideas easier to communicate.

[2] Distributions whose probability density function f satisfies $f(x) \geq b$ for all $x \in [-a, a]$, for some constants $a, b > 0$ (see Corollary 3.3 from [27]).

This work offers a substantially simpler alternative to the original proof of Theorem 1 by following a general framework introduced in the context of the analysis of Rumour Spreading algorithms [15]. Originally, the work [27] approaches Theorem 1 by considering the random variable associated with the proportion of the values in the interval $[-1, 1]$ that can be approximated up to error ε by the sum of some subset of the first t variables, $X_1, \dots, X_t$.

After restricting to some specific types of subsets, they proceed to evaluate the expected per-round growth of this proportion, conditioned on the outcomes of $X_1, \dots, X_t$. Their strategy is to analyse this expected increase by martingale theory, which only becomes possible after a non-linear transformation of the variables of interest. Those operations hinder any intuition for the obtained martingale. Nonetheless, a subsequent application of the Azuma-Hoeffding bound [1] followed by a case analysis leads to the result.

The argument presented here starts in the same direction as the original one, tracking the mass of values with suitable approximations as we reveal the values of the random variables $X_1, \dots, X_n$ one by one. However, we quickly diverge from [27], managing to obtain an estimation of the expected growth of this mass without discarding any subset-sum. We eventually restrict the argument to some types of subsets, but we do so at a point where the need for such restriction is clear.

We proceed to directly analyse the estimation obtained, without any transformations. Following [15], this estimation reveals two expected behaviours in expectation, which can be analysed similarly: as we consider the first variables, the proportion of approximated values grows very fast; then, after a certain point, the proportion of non-approximable values decreases very fast.

We remark that, while Theorem 1 crucially relies on tools from martingale theory such as Azuma-Hoeffding's inequality, which are not part of standard Computer Science curricula, our argument makes use of much more elementary results[3] which should make it accessible enough for an undergraduate course on randomised algorithms.

2 Our argument

In this section, we provide an alternative argument for proving Theorem 1. It takes shape much like the pseudo-polynomial algorithm we described in the introduction. Leveraging the recursive nature of the problem, we construct a process which, at time t, describes the proportion of the interval $[-1, 1]$ that can be approximated by some subset of the first t variables.

We will show that with a suitable number of uniform variables (proportional to $\log(1/\varepsilon)$) a factor of $1 - \varepsilon/2$ of the values in $[-1, 1]$ can be approximated up to error ε. This implies that any $z \in [-1, 1]$ which cannot be approximated within error ε is at most ε away from a value that can. Therefore it is possible to approximate z up to error 2ε.

2.1 Preliminaries

Let $X_1, \dots, X_n$ be realisations of random variables as in Theorem 1, and, without loss of generality, fix $\varepsilon > 0$. We say a value $z \in \mathbb{R}$ is *ε-approximated at time t* if and only if there exists $S \subseteq [t]$ such that $|z - \sum_{i \in S} X_i| < \varepsilon$. For $0 \le t \le n$, let $f_t \colon \mathbb{R} \to \{0, 1\}$ be the indicator function for the event "z is ε-approximated at time t". Therefore, we have $f_0 = \mathbb{1}_{(-\varepsilon,\varepsilon)}$, since only the interval $(-\varepsilon, \varepsilon)$ can be approximated by an empty set of values. From there, we can

[3] Namely, the intermediate value theorem, Markov's inequality, and standard Hoeffding bounds.

exploit the recurrent nature of the problem: a value z can be ε-approximated at time $t+1$ if and only if either z or $z - X_{t+1}$ could already be approximated at time t. This implies that for all $z \in \mathbb{R}$ we have that

$$f_{t+1}(z) = f_t(z) + \big(1 - f_t(z)\big)\, f_t(z - X_{t+1}). \tag{1}$$

To keep track of the proportion of values in $[-1, 1]$ that can be ε-approximated at each step, we define, for each $0 \le t \le n$, the random variable

$$v_t = \frac{1}{2}\int_{-1}^{1} f_t(z)\,\mathrm{d}z.$$

For better readability, throughout the text we will refer to v_t simply as "the volume."

As we mentioned, it suffices to show that, with high probability, at time n, enough of the interval is ε-approximated (more precisely, that $v_n \ge 1 - \varepsilon/2$) to conclude that the entire interval is 2ε-approximated.

2.1.1 Expected behaviour

Our first lemma provides a lower bound on the expected value of v_t.

▶ **Lemma 2.** *For all $0 \le t < n$, it holds that*

$$\mathbb{E}\big[v_{t+1} \,\big|\, X_1, \ldots, X_t\big] \ge v_t \left[1 + \frac{1}{4}(1 - v_t)\right].$$

Proof. The definition of v_t and the recurrence in Equation (1) give us that

$$\begin{aligned}
\mathbb{E}\big[v_{t+1} \,\big|\, X_1, \ldots, X_t\big] &= \mathbb{E}\left[\frac{1}{2}\int_{-1}^{1} f_{t+1}(z)\,\mathrm{d}z \,\middle|\, X_1, \ldots, X_t\right] \\
&= \int_{-1}^{1} \frac{1}{2}\left(\frac{1}{2}\int_{-1}^{1} f_t(z) + \big(1 - f_t(z)\big)\, f_t(z - x)\,\mathrm{d}z\right)\mathrm{d}x \\
&= \frac{1}{2}\int_{-1}^{1} f_t(z)\,\mathrm{d}z \int_{-1}^{1} \frac{1}{2}\,\mathrm{d}x + \frac{1}{2}\int_{-1}^{1} \frac{1}{2}\int_{-1}^{1} \big(1 - f_t(z)\big)\, f_t(z - x)\,\mathrm{d}z\,\mathrm{d}x \\
&= v_t + \frac{1}{4}\int_{-1}^{1} \big(1 - f_t(z)\big) \int_{-1}^{1} f_t(z - x)\,\mathrm{d}x\,\mathrm{d}z \\
&= v_t + \frac{1}{4}\int_{-1}^{1} \big(1 - f_t(z)\big) \int_{z-1}^{z+1} f_t(y)\,\mathrm{d}y\,\mathrm{d}z,
\end{aligned}$$

where the last equality holds by substituting $y = z - x$. For the previous ones, we apply the basic properties of integrals and Fubini's theorem to change the order of integration.

We now look for a lower bound for the last integral in terms of v_t. To this end, we exploit that, since all integrands are non-negative, for all $u \in [-1/2, 1/2]$ we have that

$$\begin{aligned}
\int_{-1}^{1} \big(1 - f_t(z)\big) \int_{z-1}^{z+1} f_t(y)\,\mathrm{d}y\,\mathrm{d}z &\ge \int_{u-\frac{1}{2}}^{u+\frac{1}{2}} \big(1 - f_t(z)\big) \int_{z-1}^{z+1} f_t(y)\,\mathrm{d}y\,\mathrm{d}z \\
&\ge \int_{u-\frac{1}{2}}^{u+\frac{1}{2}} \big(1 - f_t(z)\big) \int_{u-\frac{1}{2}}^{u+\frac{1}{2}} f_t(y)\,\mathrm{d}y\,\mathrm{d}z.
\end{aligned}$$

Both inequalities come from range restrictions: in the first, we use that $u \in [-1/2, 1/2]$ implies $[u-1/2, u+1/2] \subseteq [-1, 1]$; for the second, we have that $[u-1/2, u+1/2] \subseteq [z-1, z+1]$ for all $z \in [u-1/2, u+1/2]$.

To relate the expression to v_t explicitly, we choose u in a way that the window $[u - 1/2, u + 1/2]$ entails exactly half of v_t. The existence of such u may become clear by recalling the definition of v_t. To make it formal, consider the function given by

$$h(u) = \frac{1}{2} \int_{u-\frac{1}{2}}^{u+\frac{1}{2}} f_t(y) \, \mathrm{d}y,$$

and observe that

$$\min \{h(-1/2), \, h(1/2)\} \leq \frac{v_t}{2}, \qquad \text{and} \qquad \max \{h(-1/2), \, h(1/2)\} \geq \frac{v_t}{2}.$$

Thus, by the intermediate value theorem, there exists $u^* \in [-1/2, 1/2]$ for which $h(u^*) = v_t/2$, that is, for which

$$\frac{1}{2} \int_{u^*-\frac{1}{2}}^{u^*+\frac{1}{2}} f_t(y) \, \mathrm{d}y = \frac{v_t}{2}.$$

Altogether, we can conclude that

$$\begin{aligned}
\mathbb{E}\big[v_{t+1} \,\big|\, X_1, \ldots, X_t\big] &= v_t + \frac{1}{4} \int_{-1}^{1} (1 - f_t(z)) \int_{z-1}^{z+1} f_t(y) \, \mathrm{d}y \, \mathrm{d}z \\
&\geq v_t + \frac{1}{2} \int_{u^*-\frac{1}{2}}^{u^*+\frac{1}{2}} (1 - f_t(z)) \left(\frac{1}{2} \int_{u^*-\frac{1}{2}}^{u^*+\frac{1}{2}} f_t(y) \, \mathrm{d}y \right) \mathrm{d}z \\
&= v_t + \left(\frac{1}{2} - \frac{v_t}{2} \right) \frac{v_t}{2} \\
&= v_t \left[1 + \frac{1}{4} (1 - v_t) \right]. \qquad \blacktriangleleft
\end{aligned}$$

Lemma 2 tells us that, if v_t were to behave as expected, it should grow exponentially up to $1/2$, at which point $1 - v_t$ starts to decrease exponentially. The rest of the proof follows accordingly, with Section 2.2 analysing the progress of v_t up to one half, and Section 2.3 analogously following the complementary value, $1 - v_t$, starting from one half. By building on the results from Section 2.2, we obtain fairly straightforward proofs in Section 2.3. Thus, the following subsection comprises the core of our argument.

2.2 Growth of the volume up to 1/2

Arguably, the main challenge in analysing the RSSP is the existence of over-time dependencies and deciding how to overcome it sets much of the course the proof will take. Our strategy consists in constructing another process which dominates the original one while being free of dependencies.

Let τ_1 be the first time at which the volume exceeds $1/2$, that is, let

$$\tau_1 = \min\{t \geq 0 : v_t > 1/2\}.$$

We just proved that up to time τ_1 the process v_t enjoys exponential growth in expectation. In the following lemma, we apply a basic concentration inequality to translate this property into a constant probability of exponential growth for v_t itself.

▶ **Lemma 3.** *Given* $\beta \in (0, 1/8)$*, let* $p_\beta = 1 - \frac{7}{8(1-\beta)}$*. For all integers* $0 \leq t < \tau_1$ *it holds that*

$$\Pr\left[v_{t+1} \geq v_t(1+\beta) \,\middle|\, X_1, \dots, X_t,\, t < \tau_1\right] \geq p_\beta.$$

Proof. The result shall follow easily from reverse Markov's inequality [9, Lemma 4] and the bound from Lemma 2. However, doing so requires a suitable upper bound on v_{t+1} and, while $2v_t$ would serve the purpose, such bound does not hold in general.

We overcome this limitation by fixing t and considering how much v_t would grow in the next step if we were to consider only values ε-approximated at time t that happen to lie in $[-1, 1]$ after being translated by X_{t+1}. Making it precise by the means of the recurrence in Equation (1), we define

$$\tilde{v} = \frac{1}{2} \int_{-1}^{1} \left[f_t(z) + \big(1 - f_t(z)\big)\, f_t(z - X_{t+1}) \cdot \mathbb{1}_{[-1,1]}(z - X_{t+1}) \right] \mathrm{d}z.$$

This expression differs from the one for v_{t+1} only by the inclusion of the characteristic function of $[-1, 1]$. This not only implies that $\tilde{v} \leq v_{t+1}$, but also that $\tilde{v}$ can replace v_{t+1} in the bound from Lemma 2, since the argument provided there eventually restricts itself to integrals within $[-1, 1]$, trivialising $\mathbb{1}_{[-1,1]}$. Moreover, as we obtain $\tilde{v}$ without the influence of values from outside $[-1, 1]$, we must have $\tilde{v} \leq 2v_t$. Finally, using that $t < \tau_1$ implies $v_t < 1/2$ and chaining the previous conclusions in respective order , we conclude that

$$\begin{aligned}
\Pr\left[v_{t+1} \geq v_t(1+\beta) \,\middle|\, X_1, \dots, X_t, t < \tau_1\right] &\geq \Pr\left[\tilde{v} \geq v_t(1+\beta) \,\middle|\, X_1, \dots, X_t, t < \tau_1\right] \\
&\geq \frac{\mathbb{E}[\tilde{v} \mid X_1, \dots, X_t, t < \tau_1] - v_t(1+\beta)}{2v_t - v_t(1+\beta)} \\
&\geq \frac{\frac{9}{8}v_t - v_t(1+\beta)}{2v_t - v_t(1+\beta)} \\
&= 1 - \frac{7}{8(1-\beta)},
\end{aligned}$$

where we applied the reverse Markov's inequality in the second step. ◀

The previous lemma naturally leads us to look for bounds on τ_1, that is, to estimate the time needed for the process to reach volume $1/2$. As expected, the exponential nature of the process yields a logarithmic bound.

▶ **Lemma 4.** *Let* t *be an integer and given* $\beta \in (0, 1/8)$*, let* $p_\beta = 1 - \frac{7}{8(1-\beta)}$ *and* $i^* = \left\lceil \frac{\log \frac{1}{2\varepsilon}}{\log(1+\beta)} \right\rceil$*. If* $t \geq i^*/p_\beta$*, then*

$$\Pr\left[\tau_1 \leq t\right] \geq 1 - \exp\left[-\frac{2p_\beta^2}{t} \left(t - \frac{i^*}{p_\beta} \right)^2 \right].$$

Proof. The main idea behind the proof is to define a new random variable which stochastically dominates τ_1 while being simpler to analyse. We begin by discretising the domain $(0, 1/2]$ of the volume into sub-intervals $\{I_i\}_{0 \leq i \leq i^*}$ defined as follows:

$$\begin{cases}
I_0 = (0, \varepsilon], \\
I_i = \left(\varepsilon(1+\beta)^{i-1},\, \varepsilon(1+\beta)^i\right] \text{ for } 1 \leq i < i^*, \\
I_{i^*} = \left(\varepsilon(1+\beta)^{i^*-1},\, \dfrac{1}{2}\right],
\end{cases}$$

where i^* is the smallest integer for which $\varepsilon\,(1+\beta)^{i^*} \geq 1/2$, that is, $i^* = \left\lceil \frac{\log \frac{1}{2\varepsilon}}{\log(1+\beta)} \right\rceil$.

Now, for each $i \geq 0$, we direct our interest to the number of steps required for v_t to exit the sub-interval I_i after first entering it. By Lemma 3, this amount is majorised by a geometric random variable $Y_i \sim \mathrm{Geom}(p_\beta)$. Therefore, we can conclude that τ_1 is stochastically dominated by the sum of such variables, that is, for $t \in \mathbb{N}$, we have that

$$\Pr\left[\tau_1 \geq t\right] \leq \Pr\left[\sum_{i=1}^{i^*} Y_i \geq t\right]. \tag{2}$$

Let $B_t \sim \mathrm{Bin}(t, p_\beta)$ be a binomial random variable. For the sum of geometric random variables, it holds that $\Pr\left[\sum_{i=1}^{i^*} Y_i \leq t\right] = \Pr\left[B_t \geq i^*\right]$. Since $\mathbb{E}[B_t] = tp_\beta$, the Hoeffding bound for binomial random variables [16, Theorem 1.1] implies that, for all $\lambda \geq 0$, we have that $\Pr[B_t \leq tp_\beta - \lambda] \leq \exp(-2\lambda^2/t)$. Setting t such that $tp_\beta - \lambda = i^*$, we obtain that

$$\begin{aligned}\Pr\left[\sum_{i=1}^{i^*} Y_i \geq t\right] &\leq \Pr\left[B_t \leq i^*\right] \\ &\leq \exp\left[-\frac{2}{t}\left(tp_\beta - i^*\right)^2\right] \\ &= \exp\left[-\frac{2p_\beta^2}{t}\left(t - \frac{i^*}{p_\beta}\right)^2\right],\end{aligned}$$

which holds as long as $\lambda = tp_\beta - i^* \geq 0$, that is, for all $t \geq \frac{1}{p_\beta}\left\lceil \frac{\log \frac{1}{2\varepsilon}}{\log(1+\beta)} \right\rceil$.

The thesis follows by applying this to Equation (2) and considering the complementary events. ◀

Since we are done with the analysis of the first phase, we can fix the value of β and rearrange the bound in Lemma 4 to make it easier to apply later.

▶ **Corollary 5.** *Let $\varepsilon \in (0, \frac{1}{3})$, and let t be an integer satisfying $t \geq 264 \log \frac{1}{\varepsilon}$. Then*

$$\Pr\left[\tau_1 \leq t\right] \geq 1 - \exp\left[-\frac{2}{225t}\left(t - 264 \log \frac{1}{\varepsilon}\right)^2\right].$$

Proof. Setting $\beta = \frac{1}{16}$ in Lemma 4 and, thus, $p_\beta = \frac{1}{15}$, it suffices to notice that

$$\frac{15 \log \frac{1}{2\varepsilon}}{\log \frac{17}{16}} + 15 \leq 264 \log \frac{1}{\varepsilon}.$$ ◀

2.3 Growth of the volume from 1/2

Here we study the second half of the process: from the moment the volume reaches 1/2 up to the time it gets to $1 - \varepsilon/2$. We do so by analysing the complementary stochastic process, i.e., by tracking, from time τ_1 onwards, the proportion of the interval $[-1, 1]$ that cannot be approximated up to error ε. More precisely, we consider the process $\{w_t\}_{t \geq 0}$, defined by $w_t = 1 - v_{\tau_1 + t}$.

We shall obtain results for w_t similar to those we have proved for v_t. Fortunately, building on the previous results makes those proofs quite straightforward. We start by noting that a statement analogous to Lemma 2 follows immediately from the definition of w_{t+1} and Lemma 2.

▶ **Corollary 6.** *For all $t \geq 0$, it holds that*

$$\mathbb{E}\left[w_{t+1} \,\middle|\, X_1, \ldots, X_{\tau_1+t}\right] \leq w_t \left[1 - \frac{1}{4}(1 - w_t)\right].$$

Let τ_2 the first time that w_t gets smaller than or equal to $\varepsilon/2$, that is, let

$$\tau_2 = \min\left\{t \geq 0 : w_t \leq \varepsilon/2\right\}.$$

The following lemma bounds this quantity, in analogy to Lemma 4.

▶ **Lemma 7.** *For all $t > 0$, it holds that*

$$\Pr\left[\tau_2 \leq t\right] \geq 1 - \exp\left[-\frac{1}{8}\left(t - 8\log\frac{1}{\varepsilon}\right)\right].$$

Proof. Applying that $1 - w_t = v_{\tau_1+t} > 1/2$ to Corollary 6 gives the bound

$$\mathbb{E}\left[w_{t+1} \,\middle|\, X_1, \ldots, X_{\tau_1+t}\right] \leq \frac{7}{8} w_t. \tag{3}$$

Moreover, from the conditional expectation theory, for any two random variables X and Y, we have $\mathbb{E}\left[\mathbb{E}[X \mid Y]\right] = \mathbb{E}[X]$. From this and Equation (3), we can conclude that

$$\mathbb{E}[w_t] = \mathbb{E}\Big[\mathbb{E}\left[w_t \,\middle|\, X_1, \ldots, X_{\tau_1+t-1}\right]\Big] \leq \frac{7}{8}\,\mathbb{E}\left[w_{t-1}\right],$$

which, by recursion, yields that

$$\mathbb{E}[w_t] \leq \left(\frac{7}{8}\right)^t \mathbb{E}[w_0] \leq \frac{1}{2}\left(\frac{7}{8}\right)^t.$$

Finally, by Markov's inequality,

$$\begin{aligned}\Pr\left[\tau_2 \geq t\right] &\leq \Pr\left[w_t \geq \frac{\varepsilon}{2}\right] \\ &\leq \frac{2\,\mathbb{E}[w_t]}{\varepsilon} \\ &\leq \frac{1}{\varepsilon}\left(\frac{7}{8}\right)^t,\end{aligned}$$

and, since $\log\frac{8}{7} > \frac{1}{8}$, it holds that

$$\Pr\left[\tau_2 \geq t\right] \leq \exp\left[-\frac{1}{8}\left(t - 8\log\frac{1}{\varepsilon}\right)\right].$$

The thesis follows by considering the complementary event. ◀

2.4 Putting everything together

In this section we conclude our argument, finally proving Theorem 1. We first prove a more general statement and then detail how it implies the theorem.

Let $\tau = \tau_1 + \tau_2$, the first time at which the process $\{v_t\}_{t\geq 0}$ reaches at least $1 - \varepsilon/2$.

▶ **Lemma 8.** *Let $\varepsilon \in (0, 1/3)$. There exist constants $C' > 0$ and $\kappa > 0$ such that for every $t \geq C' \log \frac{1}{\varepsilon}$, it holds that*

$$\Pr\left[\tau \leq t\right] \geq 1 - 2\exp\left[-\frac{1}{\kappa t}\left(t - C' \log \frac{1}{\varepsilon}\right)^2\right].$$

Proof. The definition of τ allows us to apply Corollary 5 and Lemma 7 quite directly. Indeed, if, for the sake of Corollary 5, we assume that $t/2 \geq 264 \log \frac{1}{\varepsilon}$, we have that

$$\begin{aligned}
\Pr\left[\tau \leq t\right] &= \Pr\left[\tau_1 + \tau_2 \leq t\right] \\
&\geq \Pr\left[\tau_1 \leq t/2, \tau_2 \leq t/2\right] \\
&\geq \Pr\left[\tau_1 \leq t/2\right] + \Pr\left[\tau_2 \leq t/2\right] - 1 \\
&\geq 1 - \exp\left[-\frac{4}{225t}\left(\frac{t}{2} - 264 \log \frac{1}{\varepsilon}\right)^2\right] - \exp\left[-\frac{1}{8}\left(\frac{t}{2} - 8 \log \frac{1}{\varepsilon}\right)\right] \\
&\geq 1 - \exp\left[-\frac{4}{225t}\left(\frac{t}{2} - 264 \log \frac{1}{\varepsilon}\right)^2\right] - \exp\left[-\frac{1}{4t}\left(\frac{t}{2} - 8 \log \frac{1}{\varepsilon}\right)^2\right],
\end{aligned}$$

where the second inequality holds by the union bound. By setting $\kappa = 225$ and $C' = 512$, we obtain the thesis. ◀

The expression in the claim of Lemma 8 can be reformulated as

$$\Pr\left[v_t \geq 1 - \frac{\varepsilon}{2}\right] \geq 1 - 2\exp\left[-\frac{1}{\kappa t}\left(t - C' \log \frac{1}{\varepsilon}\right)^2\right];$$

hence, Theorem 1 follows by taking $C \geq 3C'$ and observing that once we can approximate all but an $\varepsilon/2$ proportion of the interval $[-1, 1]$, any $z \in [-1, 1]$ either is ε-approximated itself, or is at most ε away from a value that is, which implies that z is 2ε-approximated.

References

1 Kazuoki Azuma. Weighted sums of certain dependent random variables. *Tohoku Mathematical Journal, Second Series*, 19(3):357–367, 1967.

2 Luca Becchetti, Arthur Carvalho Walraven da Cunha, Andrea Clementi, Francesco d' Amore, Hicham Lesfari, Emanuele Natale, and Luca Trevisan. On the Multidimensional Random Subset Sum Problem. report, Inria & Université Cote d'Azur, CNRS, I3S, Sophia Antipolis, France ; Sapienza Università di Roma, Rome, Italy ; Università Bocconi, Milan, Italy ; Università di Roma Tor Vergata, Rome, Italy, 2022.

3 Rene Beier and Berthold Vöcking. Random knapsack in expected polynomial time. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '03, pages 232–241. Association for Computing Machinery, 2003. doi:10.1145/780542.780578.

4 Rene Beier and Berthold Vöcking. Probabilistic analysis of knapsack core algorithms. In *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '04, pages 468–477. Society for Industrial and Applied Mathematics, 2004.

5 Richard Bellman. Dynamic programming. *Science*, 153(3731):34–37, 1966.

6 Christian Borgs, Jennifer T. Chayes, and Boris G. Pittel. Phase transition and finite-size scaling for the integer partitioning problem. *Random Struct. Algorithms*, 19(3-4):247–288, 2001. doi:10.1002/rsa.10004.

7 Sander Borst, Daniel Dadush, Sophie Huiberts, and Danish Kashaev. A nearly optimal randomized algorithm for explorable heap selection. *CoRR*, abs/2210.05982, 2022. doi:10.48550/arXiv.2210.05982.

8 Sander Borst, Daniel Dadush, Sophie Huiberts, and Samarth Tiwari. On the integrality gap of binary integer programs with Gaussian data. *Mathematical Programming*, 2022. doi:10.1007/s10107-022-01828-1.

9 Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. Randomized gossip algorithms. *IEEE transactions on information theory*, 52(6):2508–2530, 2006.

10 Karl Bringmann and Philip Wellnitz. On near-linear-time algorithms for dense subset sum. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1777–1796. SIAM, 2021.

11 Rebekka Burkholz, Nilanjana Laha, Rajarshi Mukherjee, and Alkis Gotovos. On the existence of universal lottery tickets. In *International Conference on Learning Representations*, 2022. URL: https://openreview.net/forum?id=SYB4WrJql1n.

12 Xi Chen, Yaonan Jin, Tim Randolph, and Rocco A. Servedio. Average-case subset balancing problems. In Joseph (Seffi) Naor and Niv Buchbinder, editors, *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference / Alexandria, VA, USA, January 9 - 12, 2022*, pages 743–778. SIAM, 2022. doi:10.1137/1.9781611977073.33.

13 Arthur da Cunha, Emanuele Natale, and Laurent Viennot. Proving the strong lottery ticket hypothesis for convolutional neural networks. In *International Conference on Learning Representations*, 2022. URL: https://openreview.net/forum?id=Vjki79-619-.

14 Arthur Carvalho Walraven da Cunha, Francesco d'Amore, and Emanuele Natale. Convolutional neural networks contain structured strong lottery tickets. Preprint, June 2023. URL: https://hal.science/hal-04143024.

15 Benjamin Doerr and Anatolii Kostrygin. Randomized rumor spreading revisited. In Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl, editors, *44th International Colloquium on Automata, Languages, and Programming, ICALP 2017, July 10-14, 2017, Warsaw, Poland*, volume 80 of *LIPIcs*, pages 138:1–138:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPIcs.ICALP.2017.138.

16 Devdatt P. Dubhashi and Alessandro Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press, 2009. URL: http://www.cambridge.org/gb/knowledge/isbn/item2327542/.

17 Andre Esser and Alexander May. Low weight discrete logarithms and subset sum in $2^{0.65n}$ with polynomial memory. *Cryptology ePrint Archive*, 2019.

18 Jonas Fischer and Rebekka Burkholz. Towards strong pruning for lottery tickets with non-zero biases. *CoRR*, abs/2110.11150, 2021. arXiv:2110.11150.

19 M. R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

20 Peter Gemmell and Anna M. Johnston. Analysis of a subset sum randomizer. *IACR Cryptol. ePrint Arch.*, page 18, 2001. URL: http://eprint.iacr.org/2001/018.

21 Brian Hayes. The easiest hard problem. *American Scientist*, 90:113–117, 2002.

22 Alexander Helm and Alexander May. Subset sum quantumly in 1.17ˆ n. In *13th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

23 Ce Jin, Nikhil Vyas, and Ryan Williams. Fast low-space algorithms for subset sum. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1757–1776. SIAM, 2021.

24 Ce Jin and Hongxun Wu. A simple near-linear pseudopolynomial time randomized algorithm for subset sum. *arXiv preprint arXiv:1807.11597*, 2018.

25 David S. Johnson, Cecilia R. Aragon, Lyle A. McGeoch, and Catherine Schevon. Optimization by simulated annealing: an experimental evaluation; part ii, graph coloring and number partitioning. *Operations research*, 39(3):378–406, 1991.

26 Aniket Kate and Ian Goldberg. Generalizing cryptosystems based on the subset sum problem. *Int. J. Inf. Sec.*, 10(3):189–199, 2011. doi:10.1007/s10207-011-0129-2.

27 George S. Lueker. Exponentially small bounds on the expected optimum of the partition and subset sum problems. *Random Structures and Algorithms*, 12:51–62, 1998.

28 Stephan Mertens. A physicist's approach to number partitioning. *Theor. Comput. Sci.*, 265(1-2):79–108, 2001. doi:10.1016/S0304-3975(01)00153-0.

29 Ankit Pensia, Shashank Rajput, Alliot Nagle, Harit Vishwakarma, and Dimitris S. Papailiopoulos. Optimal lottery tickets via subset sum: Logarithmic over-parameterization is sufficient. In Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin, editors, *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020. URL: https://proceedings.neurips.cc/paper/2020/hash/1b742ae215adf18b75449c6e272fd92d-Abstract.html.

30 Wheeler Ruml, J. Thomas Ngo, Joe Marks, and Stuart M Shieber. Easily searched encodings for number partitioning. *Journal of Optimization Theory and Applications*, 89(2):251–291, 1996.

31 Zhi-Wei Sun. Unification of zero-sum problems, subset sums and covers of z. *Electronic Research Announcements of The American Mathematical Society*, 9:51–60, 2003.

32 Chenghong Wang, Jieren Deng, Xianrui Meng, Yijue Wang, Ji Li, Sheng Lin, Shuo Han, Fei Miao, Sanguthevar Rajasekaran, and Caiwen Ding. A secure and efficient federated learning framework for NLP. In Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih, editors, *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*, pages 7676–7682. Association for Computational Linguistics, 2021. doi:10.18653/v1/2021.emnlp-main.606.