Almost optimum ℓ-covering of $\mathbb{Z}_n$

Ke Shi* Chao Xu†

December 9, 2022

Abstract

A subset B of ring $\mathbb{Z}_n$ is called a ℓ-covering set if $\{ab \pmod n \mid 0 \le a \le \ell, b \in B\} = \mathbb{Z}_n$. We show there exists a ℓ-covering set of $\mathbb{Z}_n$ of size $O(\frac{n}{\ell}\log n)$ for all n and ℓ, and how to construct such set. We also show examples where any ℓ-covering set must have size $\Omega(\frac{n}{\ell}\frac{\log n}{\log\log n})$. The proof uses a refined bound for relative totient function obtained through sieve theory, and existence of a large divisor with linear divisor sum.

1 Introduction

For two sets $A, B \subseteq \mathbb{Z}_n$, we let $A \cdot B = \{ab \pmod n \mid a \in A, b \in B\}$. Let $[\ell] = \{0, \ldots, \ell\}$ be the natural numbers no larger than ℓ. A subset B of ring $\mathbb{Z}_n$ is called a *ℓ-covering set* if $(\mathbb{Z}_n \cap [\ell]) \cdot B = \mathbb{Z}_n$. Let $f(n,\ell)$ be the size of the smallest ℓ-covering set of $\mathbb{Z}_n$. Equivalently, we can define a *segment* of slope i and length ℓ to be $\{ix \pmod n \mid x \in \mathbb{Z}_n \cap [\ell]\}$, and we are interested in finding a set of segments that covers $\mathbb{Z}_n$.

ℓ-covering sets are more than just a mathematical curiosity. It was used for flash storage related problems, including covering codes [7, 8, 9], rewriting schemes [6], and generalizations to $\mathbb{Z}_n^d$ [7]. An ℓ-covering is also useful in algorithm design. Since we can *compress* a segment by dividing everything by its slope, algorithm where the running time depends on the size of the numbers in the input can be improved. A implicit but involved application of ℓ-covering was in the first significant improvement to modular subset sum problem [10].

The major question is finding the right bound for $f(n,\ell)$. The trivial lower bound is $f(n,\ell) \ge \frac{n}{\ell}$. On the upper bound of $f(n,\ell)$, there are multiple studies where ℓ is a small constant, or n has lots of structure, like being a prime number or maintaining certain divisibility conditions [7, 8, 9]. A fully general non-trivial upper bound for all ℓ and n was first established by Chen et.al., which shows an explicit construction of an $O(\frac{n(\log n)^{\omega(n)}}{\ell^{1/2}})$ size ℓ-covering set. They also showed $f(n,\ell) \le \frac{n^{1+o(1)}}{\ell^{1/2}}$ using the fourth moment of character sums, but without providing a construction [2]. In the same article, the authors show $f(p,\ell) = O(\frac{p}{\ell})$ for prime p with an explicit construction. Koiliaris and Xu improved the result by a factor of $\sqrt{\ell}$ for general n and ℓ using basic number theory, and showed $f(n,\ell) = \frac{n^{1+o(1)}}{\ell}$ [10]. An ℓ-covering set of the same size can also be found in $O(n\ell)$ time. The value hidden in $o(1)$ could be as large as $\Omega(\frac{1}{\log\log n})$, so it is relatively far from the lower bound. However, a closer inspection of their result shows $f(n,\ell) = O(\frac{n}{\ell}\log n \log\log n)$ if ℓ is neither too large nor too small. That is, if $t \le \ell \le n/t$, where $t = n^{\Omega(\frac{1}{\log\log n})}$. See fig. 1.1 for comparison of the results.

*School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China 610054. self.ke.shi@gmail.com

†Corresponding author. School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China 610054. the.chao.xu@gmail.com

	Size of ℓ-covering	Construction Time
Chen et. al. [2]	$O\left(\frac{n(\log n)^{\omega(n)}}{\ell^{1/2}}\right)$	$\tilde{O}\left(\frac{n(\log n)^{\omega(n)}}{\ell^{1/2}}\right)$
Chen et. al. [2]	$\frac{n^{1+o(1)}}{\ell^{1/2}}$	Non-constructive
Koiliaris and Xu [10]	$\frac{n^{1+o(1)}}{\ell}$	$O(n\ell)$
theorem 4.2	$O(\frac{n}{\ell}\log n)$	$O(n\ell)$
theorem 4.4	$O(\frac{n}{\ell}\log n \log\log n)$	$\tilde{O}(\frac{n}{\ell}) + n^{o(1)}$ randomized

Figure 1.1: Comparison of results for ℓ-covering for arbitrary n and ℓ. $\omega(n)$ is the number of distinct prime factors of n.

The covering problem can be considered in a more general context. For any *semigroup* $(M,\diamond)$, define $A\diamond B = \{a\diamond b \mid a\in A, b\in B\}$. For $A\subseteq M$, we are interested in finding a small B such that $A\diamond B = M$. Here B is called an A-covering. The ℓ-covering problem is the special case where the semigroup is $(\mathbb{Z}_n,\cdot)$, and $A=\mathbb{Z}_n\cap[\ell]$. When M is a group, it was studied in [1]. In particular, they showed for a finite group $(G,\diamond)$ and any $A\subseteq G$, there exists an A-covering of size no larger than $\frac{|G|}{|A|}(\log|A|+1)$. We emphasize that our problem is over the *semigroup* $(\mathbb{Z}_n,\cdot)$, which is *not a group*, and can behave very differently. For example, if A consists of only elements divisible by 2 and n is divisible by 2, then no A-covering of $(\mathbb{Z}_n,\cdot)$ exists. It was shown that there exists A that is a set of ℓ consecutive integers, any A-covering of $(\mathbb{Z}_n,\cdot)$ has $\Omega(\frac{n}{\ell}\log n)$ size [12]. This shows the choice of the set $\mathbb{Z}_n\cap[\ell]$ is very special, as there are examples where ℓ-covering has $O(\frac{n}{\ell})$ size [2]. For reasons apparent in later part of the paper, we use ℓ-covering in a semigroup $(X,\cdot)$ to mean a $(X\cap[\ell])$-covering. In the pursuit of our main theorem, another instance of the covering problem arises and might be of independent interest. Let the semigroup be $(\mathbb{D}_n,\odot)$, where $\mathbb{D}_n$ is the set of divisors of n, and $a\odot b=\gcd(ab,n)$, where gcd is the greatest common divisor function. We are interested in finding a s-covering set of $\mathbb{D}_n$ for some $s<n$.

1.1 Our Contributions

1. We show $f(n,\ell)=O(\frac{n}{\ell}\log n)$.

2. We show that there exists a constant $c>0$ and an infinite number of n and ℓ, such that $f(n,\ell)\geq c\frac{n}{\ell}\frac{\log n}{\log\log n}$.

Our contribution also includes some interesting number theoretical side results. One is a sharper bound for the relative totient function, the other is the existence of a large divisor with linear divisor sum.

1.2 Technical overview

Our approach is similar to the one of Koiliaris and Xu [10]. We briefly describe their approach. Recall $\mathbb{Z}_n$ is the set of integers modulo n. We further define $\mathbb{Z}_{n,d}=\{x \mid \gcd(x,n)=d, x\in\mathbb{Z}_n\}$, and $\mathbb{Z}_n^*=\mathbb{Z}_{n,1}$. Let $\mathcal{S}_\ell(X)$ be the set of segments of length ℓ and slope in X. Their main idea is to convert the covering problem over the *semigroup* $(\mathbb{Z}_n,\cdot)$ to covering problems over the *group* $(\mathbb{Z}^*_{n/d},\cdot)$ for all $d\in\mathbb{D}_n$. Since $\mathbb{Z}_{n,d}$ forms a partition of $\mathbb{Z}_n$, one can reason about covering them individually. That is, covering $\mathbb{Z}_{n,d}$ by $\mathcal{S}_\ell(\mathbb{Z}_{n,d})$. This is equivalent to cover $\mathbb{Z}^*_{n/d}$ with $\mathcal{S}_\ell(\mathbb{Z}^*_{n/d})$, and then lift to a cover in $\mathbb{Z}_{n,d}$ by multiply everything by d. Hence, now we only have to work with covering problem over $(\mathbb{Z}^*_{n/d},\cdot)$ for all d and $n\geq 2$, all of which are *groups*. The covering results for groups can be readily applied [1]. Once we find the covering for each individual $(\mathbb{Z}^*_{n/d},\cdot)$, we take their union, and obtain an ℓ-covering.

The approach was sufficient to obtain $f(n,\ell) = O(\frac{n}{\ell}\log n \log\log n)$ if ℓ is neither *too small* nor *too large*. However, their result suffers when ℓ is extreme in two ways.

1. $\ell = n^{1-o(\frac{1}{\log\log n})}$: Any covering obtained would have size at least the number of divisors of n, which in the worst case can be $n^{\Omega(\frac{1}{\log\log n})}$, and dominates $\frac{n}{\ell}$.
2. $\ell = n^{o(\frac{1}{\log\log n})}$: If we are working on covering $\mathbb{Z}_n^*$, we need to know $|\mathbb{Z}_n^* \cap [\ell]|$, also known as $\varphi(n,\ell)$. Previously, the estimate for $\varphi(n,\ell)$ was insufficient when ℓ is small.

Our approach can extend the appliable range to all ℓ, and also eliminates the extra $\log\log n$ factor.

First, we improve the estimate for $\varphi(n,\ell)$. This improvement alone is sufficient to handle the cases when ℓ is relatively small compared to n.

Second, we show that, roughly, a small ℓ'-covering of $\mathbb{D}_n$ with some additional nice properties implies a small ℓ-covering of $\mathbb{Z}_n$, where ℓ' is some number not too small compared to ℓ. This change can shave off the $\log\log n$ factor.

Organization The paper is organized as follows. section 2 are the preliminaries, which contains the necessary number theory backgrounds. section 3 describes some number theoretical results on bounding $\varphi(n,\ell)$, finding a large divisor of n with linear divisor sum, and covering of $\mathbb{D}_n$. section 4 proves the main theorem that $f(n,\ell) = O(\frac{n}{\ell}\log n)$, discuss its construction, and also provides a lower bound.

2 Preliminaries

The paper has a few simple algorithmic ideas, but our methods are mainly analytical. Hence, we reserved some space in the preliminaries to set up the scene.

Let $\mathcal{X}$ be a collection of sets in some universe set U. A *set cover* of U is a collection of subsets in $\mathcal{X}$ which together covers U. Formally, $\mathcal{X}' \subseteq \mathcal{X}$ such that $U = \bigcup_{X\in\mathcal{X}'} X$. The *set cover problem* is the computational problem of finding a minimum cardinality set cover.

All multiplications in $\mathbb{Z}_n$ are modulo n, hence we will omit $\pmod n$ from now on. Recall a set of the form $\{ix \mid x \in \mathbb{Z}_n \cap [\ell]\}$ is called a *segment* of length ℓ with slope i. Note that the segment of length ℓ might have fewer than ℓ elements. Recall $\mathcal{S}_\ell(X)$ is the segments of length ℓ with slope in X, namely $\{\{ix \mid x \in \mathbb{Z}_n \cap [\ell]\} \mid i \in X\}$. Hence, finding a ℓ-covering is equivalent to set cover with segments in $\mathcal{S}_\ell(\mathbb{Z}_n)$, and the universe is $\mathbb{Z}_n$.

Set cover problem has some well-known bounds relating the size of a set cover and the frequency of element covered [11, 14].

Theorem 2.1 ([11, 14]) *Let there be a collection of t sets each with size at most a, and each element of the universe is covered by at least b of the sets, then there exists a subcollection of $O(\frac{t}{b}\log a)$ sets that covers the universe.*

The above theorem is the main combinatorial tool for bounding the size of a set cover. To obtain a cover of the specified size, the greedy algorithm is sufficient. Note the group covering theorem for finite groups in [1] is a direct consequence of this.

The base of the log is e. To avoid getting into the negatives, we take $\log(x)$ to mean $\max(\log(x), 1)$. $\tilde{O}(f(n))$, the soft O, is a shorthand for $O(f(n) \operatorname{polylog} n)$.

2.1 Number theory

We refer to some standard notation and bounds, where it can be found in various analytic number theory textbook, for example [4]. Recall $\mathbb{Z}_n$ is the set of integers modulo n, $\mathbb{Z}_{n,d} = \{x \mid \gcd(x,n) = d, x \in \mathbb{Z}_n\}$, and $\mathbb{Z}_n^* = \mathbb{Z}_{n,1}$. $\mathbb{Z}_n^*$ is the set of numbers in $\mathbb{Z}_n$ that are relatively prime to n. The notation $m|n$ means m is a divisor of n. $\pi(n)$, the *prime counting function*, is the number of primes no larger than n, and

$\pi(n) = \Theta(\frac{n}{\log n})$. $\varphi(n)$, the *Euler totient function*, defined as $\varphi(n) = |\mathbb{Z}_n^*| = n\prod_{p|n}\left(1-\frac{1}{p}\right)$, and is bounded by $\Omega(\frac{n}{\log\log n})$. $\omega(n)$, the *number of distinct prime factors* of n, has the relation $\omega(n) = O(\frac{\log n}{\log\log n})$. $d(n)$, the *divisor function*, is the number of divisors of n, and $d(n) = n^{O(\frac{1}{\log\log n})} = n^{o(1)}$. $\sigma(n)$, the *divisor sum function*, is the sum of divisors of n, and $\sigma(n) \le \frac{n^2}{\varphi(n)}$. This also implies $\sigma(n) = O(n\log\log n)$. The sum of reciprocal of primes no larger than n is $\sum_{p\le n, p \text{ prime}} \frac{1}{p} = O(\log\log n)$.

The center of our argument lies in the *relative totient function*, denoted as $\varphi(n,\ell) = |\mathbb{Z}_n^* \cap [\ell]|$.

Theorem 2.2 *Consider integers $0 \le \ell < n$, $y \in \mathbb{Z}_{n,d}$. The number of solutions $x \in \mathbb{Z}_n^*$ such that $xb \equiv y \pmod{n}$ for some $b \le \ell$ is*

$$\frac{\varphi(\frac{n}{d}, \lfloor \frac{\ell}{d} \rfloor)}{\varphi(\frac{n}{d})}\varphi(n).$$

Proof: See appendix B. □

We also need Brun's sieve from sieve theory, see appendix A.

3 Number theoretical results

This section we show some number theoretical bounds. The results are technical. The reader can skip the proofs of this section on first view.

3.1 Estimate for relative totient function

This section proves a good estimate of $\varphi(n,\ell)$ using sieve theory, the direction was hinted in [5].

Theorem 3.1 *There exists positive constant c, such that*

$$\varphi(n,\ell) = \begin{cases} \Omega(\frac{\ell}{n}\varphi(n)) & \text{if } \ell > c\log^5 n \\ \Omega(\frac{\ell}{\log \ell}) & \text{if } \ell > c\log n \end{cases}$$

Proof: *Case 1.* $\ell > c\log^5 n$.

Let z be a value we will define later.

Let $n_0 = \prod_{p|n, p<z} p$, we can see $\varphi(n,\ell)$ and $\varphi(n_0,\ell)$ are close. Indeed, for some $c_1 > 0$,

$$\begin{aligned}
|\varphi(n,\ell) - \varphi(n_0,\ell)| &= \left| \sum_{0\le m\le \ell, (m,n_0)=1} 1 - \sum_{0\le m\le \ell,(m,n)=1} 1 \right| \\
&\le \sum_{1\le m\le \ell : p|n, p\ge z, p|m} 1 \\
&\le \sum_{p|n, p\ge z} \frac{\ell}{p} \\
&\le \frac{\ell\,\omega(n)}{z} \\
&\le \frac{c_1 \ell \log n}{z\log\log n}
\end{aligned}$$

Now, we want to estimate $\varphi(n_0,\ell)$ using the Brun's sieve. The notations are from the theorem. Let $\mathcal{A} = \{1, 2, \dots, \ell\}$, $\mathcal{P} = \{p : p|n\}$, $X = |\mathcal{A}| = \ell$, the multiplicative function γ, where $\gamma(p) = 1$ if $p \in \mathcal{P}$ otherwise 0.

- *Condition (1).* For any squarefree d composed of primes of $\mathcal{P}$,

$$|R_d| = \left| \left\lfloor \frac{\ell}{p} \right\rfloor - \frac{\ell}{p} \right| \le 1 = \gamma(d).$$

- *Condition (2).* We choose $A_1 = 2$, therefore $0 \le \frac{\gamma(p)}{p} = \frac{1}{p} \le \frac{1}{2} = 1 - \frac{1}{A_1}$.

- *Condition (3).* Because $R(x) := \sum_{p<x} \frac{\log p}{p} = \log x + O(1)$ [3], we have

$$\sum_{w \le p < z} \frac{\gamma(p) \log p}{p} \le \sum_{w \le p < z} \frac{\log p}{p} = R(z) - R(w) = \log \frac{z}{w} + O(1).$$

  We choose $\kappa = 1$ and some A_2 large enough to satisfy Condition (3).

- *Condition (4).* By picking $b = 1, \lambda = \frac{2}{9}$, b is a positive integer and $0 < \frac{2}{9} e^{11/9} \approx 0.75 < 1$.

We are ready to bound $\varphi(n_0, \ell)$. Brun's sieve shows

$$\begin{aligned}
\varphi(n_0, \ell) = S(\mathcal{A}, \mathcal{P}, z) \ge & \ell \frac{\varphi(n_0)}{n_0} \left(1 - \frac{2\lambda^{2b} e^{2\lambda}}{1 - \lambda^2 e^{2+2\lambda}} \exp((2b+2) \frac{c_1}{\lambda \log z}) \right) \\
& + O(z^{2b-1+\frac{2.01}{e^{2\lambda/\kappa}-1}}) \\
\ge & \ell \frac{\varphi(n_0)}{n_0} \left(1 - 0.3574719 \exp(\frac{18c_1}{\log z}) \right) + O(z^{4.59170})
\end{aligned}$$

Which means that there exists some positive constant c_2 such that for some small $\varepsilon > 0$,

$$\varphi(n_0, \ell) \ge \ell \frac{\varphi(n_0)}{n_0} \left(1 - \frac{2}{5} \exp(\frac{18c_1}{\log z}) \right) - c_2 z^{5-\varepsilon}.$$

We choose some constant z_0 such that $\frac{2}{5} \exp(\frac{18c_1}{\log z_0}) \le \frac{1}{2}$, if $z > z_0$(we will later make sure $z > z_0$), then

$$\varphi(n_0, \ell) \ge \frac{1}{2} \ell \frac{\varphi(n_0)}{n_0} - c_2 z^{5-\varepsilon}.$$

Note if $n_1 | n_2$, then $\varphi(n_1)/n_1 \ge \varphi(n_2)/n_2$ since $\varphi(n)/n = \prod_{p|n}(1 - 1/p)$ and every prime factor of n_1 is also the prime factor of n_2. Therefore,

$$\varphi(n_0, \ell) \ge \frac{1}{2} \ell \frac{\varphi(n)}{n} - c_2 z^{5-\varepsilon}.$$

Recall there exists a c_3 such that $\frac{\varphi(n)}{n} \ge \frac{c_3}{\log \log n}$,

$$\begin{aligned}
\varphi(n, \ell) &\ge \varphi(n_0, \ell) - c_1 \frac{\ell \log n}{z \log \log n} \\
&\ge \frac{1}{2} \ell \frac{\varphi(n)}{n} - c_2 z^{5-\varepsilon} - c_1 \frac{\ell \log n}{z \log \log n} \\
&= \frac{1}{4} \ell \frac{\varphi(n)}{n} + (\frac{1}{8} \ell \frac{\varphi(n)}{n} - c_2 z^{5-\varepsilon}) + (\frac{1}{8} \ell \frac{\varphi(n)}{n} - c_1 \frac{\ell \log n}{z \log \log n}) \\
&\ge \frac{1}{4} \ell \frac{\varphi(n)}{n} + (\frac{c_3}{8} \frac{\ell}{\log \log n} - c_2 z^{5-\varepsilon}) + (\frac{c_3}{8} \frac{\ell}{\log \log n} - c_1 \frac{\ell \log n}{z \log \log n}).
\end{aligned}$$

By picking

$$z = \frac{8c_1}{c_3} \log n = C \log n,$$

we obtain

$$c_1 \frac{\ell \log n}{z \log\log n} \le \frac{c_3}{8} \frac{\ell}{\log\log n}.$$

By picking $c = 8\frac{c_2}{c_3}C^5$ and

$$\ell \ge \frac{8c_2}{c_3} C^5 \log^{5-\varepsilon} n \log\log n = c \log^{5-\varepsilon} n \log\log n,$$

we obtain

$$cz^{5-\varepsilon} \le \frac{\ell}{\log\log n}.$$

Recall for the above to be true we require $z > z_0$. Because $z = C \log n$, for $z > z_0$ for sufficiently large n.

We obtain if n is sufficiently large and $\ell \ge c \log^5 n \ge c \log^{5-\varepsilon} n \log\log n$, then $\varphi(n, \ell) \ge \frac{\ell}{4n}\varphi(n)$. Thus, for all n and $\ell \ge c \log^5 n$, $\varphi(n, \ell) = \Omega(\ell \frac{\varphi(n)}{n})$.

Case 2. $\ell > c \log n$.

Observe that for all $\ell \le n$, $\varphi(n, \ell) \ge 1 + \pi(\ell) - \omega(n)$. This is because the primes no larger than ℓ are relatively prime to n if it is not a factor of n, and 1 is also relatively prime to n.

We show there exists a constant c such that $\varphi(n, \ell) = \Omega(\frac{\ell}{\log \ell})$ for $\ell \ge c \log n$, by showing $\frac{1}{2}\pi(\ell) \ge \omega(n)$. There exists constant c_1, c_2 such that $\pi(\ell) \ge c_1 \frac{\ell}{\log \ell}$ and $\omega(n) \le c_2 \frac{\log n}{\log\log n}$. Therefore, we want some ℓ, such that $\frac{c_1}{2} \frac{\ell}{\log \ell} \ge c_2 \frac{\log n}{\log\log n}$. The desired relation holds as long as $\ell \ge c \log n$ for some sufficiently large c.

The constant c in two parts of the proof might be different, we pick the larger of the two to be the one in the theorem. □

As a corollary, we prove theorem 3.2.

Theorem 3.2 *There exists a constant c, such that for any n, and a divisor d of n, if $\frac{\ell}{c \log^5 n} \ge d$, then each element in $\mathbb{Z}_{n,d}$ is covered $\Omega(\frac{n}{\ell}\varphi(n))$ times by $\mathcal{S}_\ell(\mathbb{Z}_n^*)$.*

Proof: By theorem 2.2, the number of segments in $\mathcal{S}_\ell(\mathbb{Z}_n^*)$ covering some fixed element in $\mathbb{Z}_{n,d}$ is $\frac{\varphi(n/d, \ell/d)}{\varphi(n/d)}\varphi(n)$. As long as ℓ is not too small, $\varphi(n, \ell) = \Omega(\frac{\ell}{n}\varphi(n))$. In particular, by theorem 3.1, if $\lfloor \ell/d \rfloor \ge c \log^5(n/d)$, we have $\varphi(n/d, \ell/d)/\varphi(n/d) = \Omega(\frac{\ell}{n})$. Therefore, each element in $\mathbb{Z}_{n,d}$ is covered $\Omega(\frac{\ell}{n}\varphi(n))$ times. □

3.2 Large divisor with small divisor sum

Theorem 3.3 *If $r = n^{O(\frac{1}{\log\log\log n})}$, then there exists $m|n$, such that $m \ge r$, $d(m) = r^{O(\frac{1}{\log\log r})}$ and $\sigma(m) = O(m)$.*

Proof: If there is a single prime p, such that $p^e|n$ and $p^e \ge r$, then we pick $m = p^{e'}$, where e' is the smallest integer such that $p^{e'} \ge r$. One can see $d(m) = e' = O(\log r) = r^{O(\frac{1}{\log\log r})}$, also $\varphi(m) = m(1 - \frac{1}{p}) \ge \frac{m}{2}$, since $\varphi(m)\sigma(m) \le m^2$ we are done.

Otherwise, we write $n = \prod_{i=1}^k p_i^{e_i}$, where each p_i is a distinct prime number. The prime p_i are ordered by the weight $w_i = e_i p_i \log p_i$ in decreasing order. That is $w_i \ge w_{i+1}$ for all i. Let j be the smallest number such that $\prod_{i=1}^j p_i^{e_i} \ge r$. Let $m = \prod_{i=1}^j p_i^{e_i}$.

First, we show $d(m)$ is small. Let $m' = m/p_j^{e_j}$. One can see that $m' < r$ and $p_j^{e_j} < r$. So $e_j = O(\log r)$, and

$$d(m) \le (e_j + 1)d(m') = O(\log r)d(m') = r^{O(\frac{1}{\log\log r})}.$$

To show that $\sigma(m) = O(m)$, we show $\varphi(m) = \Theta(m)$. Indeed, by $\sigma(m) \le \frac{m^2}{\varphi(m)}$, we obtain $\sigma(m) = O(m)$.

For simplicity, it is easier to work with sum instead of products, so we take logarithm of everything and define $t = \log n$.

By definition, $\log r = O(\frac{\log n}{\log\log\log n}) = O(\frac{t}{\log\log t})$ and $\sum_{i=1}^{k} e_i \log p_i = t$.

Note j is the smallest number such that $\sum_{i=1}^{j} e_i \log p_i \ge \log r$. Because there is no prime p such that $p^e|n$ and $p^e \ge r$, we also have $\sum_{i=1}^{j} e_i \log p_i < 2\log r = O(\frac{t}{\log\log t})$.

Now, consider $e_1', \ldots, e_k'$, such that the following holds.

- $\sum_{i=1}^{j} e_i \log p_i = \sum_{i=1}^{j} e_i' \log p_i$, and $e_i' p_i \log p_i = c_1$ for some c_1, when $1 \le i \le j$,
- $\sum_{i=j+1}^{k} e_i \log p_i = \sum_{i=j+1}^{n} e_i' \log p_i$, and $e_i' p_i \log p_i = c_2$ for some c_2, where $j+1 \le i \le k$.

Note c_1 and c_2 can be interpreted as weighted averages over w_i. Indeed, consider sequences $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$, such that $\sum_i x_i = \sum_i y_i$. If for some non-negative $a_1, \ldots, a_n$, we have $a_i y_i = c$ for all i, j, then $c \le \max_i a_i x_i$. Indeed, there exists $x_j \ge y_j$, so $\max_i a_i x_i \ge a_j x_j \ge a_j y_j = c$. Similarly, $c \ge \min_i a_i x_i$. This shows $c_1 \ge c_2$, because $c_2 \le \max_{i=j+1}^{k} w_i = w_{j+1} \le w_j = \min_{i=1}^{j} w_i \le c_1$.

We first give a lower bound of c_2.
$\sum_{i=j+1}^{k} \frac{c_2}{p_i} = \sum_{i=j+1}^{k} e_i' \log p_i = \sum_{i=j+1}^{k} e_i \log p_i \ge t - O(\frac{t}{\log\log t}) = \Omega(t)$.
$\sum_{i=j+1}^{k} \frac{c_2}{p_i} \le c_2 \sum_{i=1}^{k} \frac{1}{p_i} \le c_2 \sum_{p \text{ prime}, p=O(t)} \frac{1}{p} = c_2 O(\log\log t)$.
This shows $c_2 O(\log\log t) = \Omega(t)$, or $c_2 = \Omega(\frac{t}{\log\log t})$.

Since $c_1 \ge c_2$, $\sum_{i=1}^{j} \frac{1}{p_i} = \sum_{i=1}^{j} \frac{e_i' \log p_i}{c_1} = \frac{O(\frac{t}{\log\log t})}{c_1} \le \frac{O(\frac{t}{\log\log t})}{c_2} = \frac{O(\frac{t}{\log\log t})}{\Omega(\frac{t}{\log\log t})} = O(1)$.

Note $\varphi(m) = m\prod_{i=1}^{j}(1-\frac{1}{p_i})$. Because $-2x < \log(1-x) < -x$ for $0 \le x \le 1/2$, so $\sum_{i=1}^{j} \log(1-\frac{1}{p_i}) \ge -2\sum_{i=1}^{j} \frac{1}{p_i} = -O(1)$. Hence $\prod_{i=1}^{j}(1-\frac{1}{p_i}) = \Omega(1)$, and $\varphi(m) = \Omega(m)$. □

A interesting number theoretical result is the direct corollary of theorem 3.3.

Corollary 3.4 *Let n be a positive integer, there exists a $m|n$ such that $m = n^{\Omega(\frac{1}{\log\log\log n})}$ and $\sigma(m) = O(m)$.*

It would be interesting to know if the above corollary is tight.

3.3 Covering of $\mathbb{D}_n$

Recall that $(\mathbb{D}_n, \odot)$ is the semigroup over the set of divisors of n, and the operation $\odot$ is defined as $a \odot b = \gcd(ab, n)$. Throughout this section, we fix a $s \le n$, and let $A := \mathbb{D}_n \cap [s]$. We are interested in finding s-coverings of $\mathbb{D}_n$, that is, finding $B \subseteq \mathbb{D}_n$ such that $(\mathbb{D}_n \cap [s]) \odot B = \mathbb{D}_n$. As we mentioned previously, the main goal is to show that a good s-covering of $\mathbb{D}_n$ lifts to a ℓ-covering of $\mathbb{Z}_n$ of small size. The criteria for a good s-covering B is two folds: the size of B should be small ($O(\frac{n}{s}\frac{1}{\log^c n})$), and the reciprocal sum of B, namely $\sum_{d\in B} \frac{1}{d}$ should also be small ($O(1)$). However, one can't hope to optimize both at the same time. Fortunately, for our application, we only need the reciprocal sum to be small when s is small.

To obtain a s-covering of $\mathbb{D}_n$, there are two natural choices of B.

1. Let $B = (\mathbb{D}_n \setminus [s]) \cup \{1\}$. If $d \leq s$, then $d = d \cdot 1$. Otherwise, if $d > s$, then $d = 1 \cdot d$. Hence, $A \odot B = \mathbb{D}_n$.

2. Let $B = \mathbb{D}_m$ for some $m|n$ and $m \geq \frac{n}{s}$. We also have $A \odot B = \mathbb{D}_n$. Indeed, consider divisor d of n, let $d_1 = \gcd(m,d) \in B$, and $d_2 = d/d_1$. $d_2 | \frac{n}{m} \leq s$, so $d_2 \in A$.

These two choices is sufficient for us to prove the following lemma. The lemma basically states there is an s-covering of $\mathbb{D}_n$ fits out requirement as long as s is not too large.

Lemma 3.5 *Let δ be a function such that $\delta(n) = \Omega(\log n)$ and $\delta(n) = O(\log^{c'} n)$ for some constant c'. There exists a constant c, such that for every $s \leq \frac{n}{\delta(n)}$, we can find $B \subseteq \mathbb{D}_n$ such that $(\mathbb{D}_n \cap [s]) \odot B = \mathbb{D}_n$, $|B| = O(\frac{n \log n}{s\delta(n)})$ and*

1. *If $s \in (0, n^{\frac{c}{\log\log n}}]$, then $\sum_{d \in B} \frac{1}{d} = O(\log\log n)$.*

2. *If $s \in (n^{\frac{c}{\log\log n}}, \frac{n}{\delta(n)}]$, then $\sum_{d \in B} \frac{1}{d} = O(1)$.*

Proof: Let $A = \mathbb{D}_n \cap [s]$. We let $B_1 = (\mathbb{D}_n \setminus [s]) \cup \{1\}$. Also, let $B_2 = \mathbb{D}_m$, where $m|n$, $d(m) = \frac{n}{s}^{O(\frac{1}{\log\log\frac{n}{s}})}$, $\sigma(m) = O(m)$. Such m exists when $s = n^{1-O(\frac{1}{\log\log\log n})}$ by setting $r = \frac{n}{s}$ in theorem 3.3. Recall both $A \odot B_1 = \mathbb{D}_n$ and $A \odot B_2 = \mathbb{D}_n$.

The proof consists of 3 different cases.

1. $s \in (0, n^{\frac{c}{\log\log n}}]$.

2. $s \in (n^{\frac{c}{\log\log n}}, n^{1-\frac{c}{\log\log n}}]$

3. $s \in (n^{1-\frac{c}{\log\log n}}, \frac{n}{f(n)}]$

For the first two cases, we let $B = B_1$.

In particular, we have $s \leq n^{1-\frac{c}{\log\log n}}$, so $\frac{n\log n}{sf(n)} = O(n^{\frac{c-\epsilon}{\log\log n}})$ for any $\epsilon > 0$. Now if we pick sufficiently large c, we would have $|B| = d(n) = n^{O(\frac{1}{\log\log n})} = O(\frac{n\log n}{sf(n)})$.

When $s \in (0, n^{\frac{c}{\log\log n}}]$, $\sum_{d\in B} \frac{1}{d} \leq \frac{1}{n}\sum_{d|n} \frac{n}{d} = \sigma(n)/n = O(\log\log n)$. Otherwise, when $s \in (n^{\frac{c}{\log\log n}}, n^{1-\frac{c}{\log\log n}}]$, each element in $B \setminus \{1\}$ is at least s, so we know that $\sum_{d\in B} \frac{1}{d} = 1 + \sum_{d \in B\setminus\{1\}} \frac{1}{d} \leq 1 + |B|\frac{1}{s} \leq 1 + \frac{n^{\frac{O(1)}{\log\log n}}}{n^{\frac{c}{\log\log n}}} = O(1)$.

Now, we consider the third case $s \in (n^{1-\frac{c}{\log\log n}}, \frac{n}{f(n)}]$. In this case we set $B = B_2$.

We first bound the size of B.

$$
\begin{aligned}
|B| &= (\frac{n}{s})^{O(\frac{1}{\log\log\frac{n}{s}})} \\
&\leq (\frac{nf(n)}{sf(n)})^{O(\frac{1}{\log\log f(n)})} \\
&\leq O(\frac{n}{sf(n)}) f(n)^{O(\frac{1}{\log\log f(n)})} \\
&\leq \frac{n}{sf(n)}(\log^{c'} n)^{O(\frac{1}{\log\log\log n})} \\
&= O(\frac{n\log n}{sf(n)})
\end{aligned}
$$

By the choice of m, we have $\sum_{d\in B} \frac{1}{d} = \frac{\sigma(m)}{m} = O(1)$. □

4 ℓ-covering

In this section, we prove our bounds in $f(n,\ell)$, provide a quick randomized construction.

4.1 Upper bound

The high level idea is to split the problem to sub-problems of covering multiple $\mathbb{Z}_{n,d}$. Can we cover $\mathbb{Z}_{n,d}$ for many distinct d, using only a few segments in $\mathcal{S}_\ell(\mathbb{Z}_n^*)$? We answer the question affirmatively by connect an s-covering of $\mathbb{D}_n$ to a ℓ-covering of $\mathbb{Z}_n$. Let $B \subseteq \mathbb{D}_n$ be any s-covering of $\mathbb{D}_n$. For each $b \in B$, we generate a cover of all $\bigcup_{d\leq s} \mathbb{Z}_{n,b\odot d}$ using $\mathcal{S}_\ell(\mathbb{Z}_{n,b})$. Let $g(n,\ell)$ to be the size of the smallest set cover of $\bigcup_{d|n,d\leq s} \mathbb{Z}_{n,d}$ using $\mathcal{S}_\ell(\mathbb{Z}_n^*)$. We obtain that

$$f(n,\ell) \leq \sum_{b\in B} g(\frac{n}{b},\ell).$$

For the rest of this section, $s = \max(1, \frac{\ell}{c\log^5 n})$, where c is the constant in theorem 3.1. We bound $g(n,\ell)$ using the fact that each element is covered many times, and theorem 2.1, the combinatorial set cover upper bound.

Theorem 4.1 *There exists a constant $c > 0$, such that*

$$g(n,\ell) = \begin{cases} O(\frac{n}{\ell}\log \ell) & \text{if } \ell \geq c\log^5 n, \\ O(\frac{\varphi(n)}{\ell}\log^2 \ell) & \text{if } c\log^5 n > \ell \geq c\log n. \end{cases}$$

Proof: By theorem 2.2, The number of times an element in $\mathbb{Z}_{n,d}$ get covered by a segment in $\mathcal{S}_\ell(\mathbb{Z}_n^*)$ is $\frac{\varphi(\frac{n}{d},\lfloor\frac{\ell}{d}\rfloor)}{\varphi(\frac{n}{d})}\varphi(n)$. We consider 2 cases.

Case 1. $\ell > c\log^5 n$. Consider a $d|n$ and $d \leq s$. Then $\lfloor\frac{\ell}{d}\rfloor = \Omega(\log^5 n)$. Hence, $\varphi(\frac{n}{d},\lfloor\frac{\ell}{d}\rfloor) = \Omega(\frac{\lfloor\frac{\ell}{d}\rfloor}{\frac{n}{d}}\varphi(\frac{n}{d})) = \Omega(\frac{\ell}{n}\varphi(\frac{n}{d}))$ by theorem 3.1. Therefore, each element in $\mathbb{Z}_{n,d}$ is covered by $\frac{\varphi(\frac{n}{d},\lfloor\frac{\ell}{d}\rfloor)}{\varphi(\frac{n}{d})}\varphi(n) = \Omega(\frac{\ell}{n}\varphi(n))$ segments in $\mathcal{S}_\ell(\mathbb{Z}_n^*)$. This is true for all element in $\bigcup_{d|n,d\leq s}\mathbb{Z}_{n,d}$.

By theorem 2.1, there exists a cover of size

$$g(n,\ell) = O\left(\frac{\varphi(n)\log\ell}{\frac{\ell}{n}\varphi(n)}\right) = O\left(\frac{n}{\ell}\log\ell\right).$$

Case 2. If $c\log^5 n > \ell \geq c\log n$, then $s = 1$, and we try to cover $\mathbb{Z}_n^*$ with $\mathcal{S}_\ell(\mathbb{Z}_n^*)$. Each element is covered by $\frac{\varphi(n,\ell)}{\varphi(n)}\varphi(n) = \Omega(\frac{\ell}{\log\ell})$ segments. By theorem 2.1, we have

$$g(n,\ell) = O\left(\frac{\varphi(n)\log\ell}{\frac{\ell}{\log\ell}}\right) = O\left(\frac{\varphi(n)}{\ell}\log^2\ell\right).$$

□

We are ready to prove our main theorem.

Theorem 4.2 (Main) *There exists an ℓ-covering set of size $O(\frac{n}{\ell}\log n)$ for all n,ℓ where $\ell < n$.*

Proof: Let B be the s-covering of $\mathbb{D}_n$ in lemma 3.5 with $\delta(n) = c\log^5 n$. Observe $s = \frac{\ell}{\delta(n)}$ and $|B| = O(\frac{n}{\ell}\log n)$.

Case 1 If $\ell < c\log n$, then we are done, since $f(n,\ell) \leq n = O(\frac{n}{\ell}\log n)$.

Case 2 Consider $c \log n \le \ell \le c \log^5 n$.

$$\begin{aligned}
f(n,\ell) &\le \sum_{d\in B} g(\frac{n}{d},\ell) \\
&\le \sum_{d\in B}\left(\varphi(n/d)\frac{(\log \ell)^2}{\ell}+1\right) \\
&\le O(\frac{n}{\ell}\log^2 \ell)+|B| \\
&= O\left(\frac{n}{\ell}(\log\log n)^2\right)+O\left(\frac{n}{\ell}\log n\right) \\
&= O\left(\frac{n}{\ell}\log n\right)
\end{aligned}$$

Case 3 Consider $\ell > c \log^5 n$.

$$\begin{aligned}
f(n,\ell) &\le \sum_{d\in B} g(\frac{n}{d},\ell) \\
&\le \sum_{d\in B} O\left(\frac{n}{d}\frac{\log \ell}{\ell}\right)+1 \\
&= |B| + O\left(\frac{n\log \ell}{\ell}\right)\sum_{d\in B}\frac{1}{d} \\
&= O\left(\frac{n}{\ell}\log n\right)+O\left(\frac{n\log \ell}{\ell}\right)\sum_{d\in B}\frac{1}{d}
\end{aligned}$$

Hence, we are concerned with the last term. We further separate into 2 cases:

Case 3.1 If $\ell < n^{\frac{c}{\log\log n}}$, then $\sum_{d\in B}\frac{1}{d} = O(\log\log n)$, and

$$\begin{aligned}
O\left(\frac{n\log \ell}{\ell}\sum_{d\in B}\frac{1}{d}\right) &= O\left(\frac{n\log\ell}{\ell}\log\log n\right) \\
&= O\left(\frac{n\frac{\log n}{\log\log n}\log\log n}{\ell}\right) \\
&= O\left(\frac{n\log n}{\ell}\right).
\end{aligned}$$

Case 3.2 $\ell \ge n^{\frac{c}{\log\log n}}$, then $\sum_{d\in B}\frac{1}{d} = O(1)$. Hence,

$$O\left(\frac{n\log \ell}{\ell}\sum_{d\in B}\frac{1}{d}\right) = O\left(\frac{n\log\ell}{\ell}\right) = O\left(\frac{n\log n}{\ell}\right).$$

In all cases, we obtain an ℓ-covering of $O(\frac{n\log n}{\ell})$ size. ∎

The upper bound automatically leads to a construction algorithm. First find the prime factorization in $n^{o(1)}$ time, then compute the desired B in $n^{o(1)}$ time, and then cover each $\bigcup_{d|n/b, d\le s}\mathbb{Z}_{n/b,d}$ using $\mathcal{S}_\ell(\mathbb{Z}^*_{n/b})$ for $b\in B$. If we use the linear time greedy algorithm for set cover, then the running time becomes $O(n\ell)$ [10].

One can use a randomized constructive version of theorem 2.1.

Theorem 4.3 *Let there be t sets, each element of the size n universe is covered by at least b of the sets, then there exists subset of $O(\frac{t}{b}\log n)$ size that covers the universe, and can be found with high probability using a Monte Carlo algorithm that runs in $\tilde{O}(\frac{t}{b})$ time.*

Proof (Sketch): The condition shows the standard linear programming relaxation of set cover has a feasible solution where every indicator variable for each set has value $\frac{1}{b}$. The standard randomized rounding algorithm of picking each set with probability equals $\frac{1}{b}$ independently, for $\Theta(\log n)$ rounds, would cover the universe with high probability [15]. It can be simulated through independently sample sets of size $\frac{t}{b}$ for $\Theta(\log n)$ rounds instead, which can be done in $\tilde{O}(\frac{t}{b})$ time. $\square$

The main difference is the coverage size between theorem 4.3 and theorem 2.1. Let a be the maximum size of each set, the randomized algorithm have a higher factor of $\log n$ instead of $\log a$. If we use more sophisticated rounding techniques, we can again obtain $\log a$ [13]. However, the algorithm will not be as fast. The change to $\log n$ has a consequence in the output size. In particular, following the proof of theorem 4.2, there will be an extra $\log\log n$ factor to the size of the cover.

The analysis is similar as before, and we can obtain the following theorem.

Theorem 4.4 *There exists a constant c, such that a $O(\frac{n}{\ell}\log n)$ size ℓ-covering of $\mathbb{Z}_n$ can be found in $\tilde{O}(\frac{n}{\ell}) + n^{o(1)}$ time with high probability if $\ell < n^{\frac{c}{\log\log n}}$, and the size is $O(\frac{n}{\ell}\log n \log\log n)$ otherwise.*

4.2 Lower bound

We remark our upper bound is the best possible through the combinatorial set covering property (theorem 2.1). The $\log n$ factor cannot be avoided when $\ell = n^{\Omega(1)}$. In order to obtain a better bound, stronger *number theoretical properties* has to be exploited, as it was for the case when n is a prime [2].

We show that it is unlikely we can get much stronger bounds when ℓ *is small*. For infinite many (n,ℓ) pairs, our bound is only $\log\log n$ factor away from the lower bound.

Theorem 4.5 *There exists a constant $c > 0$, where there are an infinite number of n, ℓ pairs where $f(n,\ell) \geq c\frac{n}{\ell}\frac{\log n}{\log\log n}$.*

Proof: Let n be the product of the smallest k prime numbers, then $k = \Theta(\frac{\log n}{\log\log n})$. Let ℓ be the smallest number where $\pi(\ell) = k$. Because $\pi(\ell) = \Theta(\frac{\ell}{\log \ell})$, we know $\ell = \Theta(\log n)$.

Observe that $\varphi(n,\ell) = 1$. Indeed, every number $\leq \ell$ except 1 has a common factor with n. In order to cover all elements in $\mathbb{Z}_n^* \subseteq \mathbb{Z}_n$, the ℓ-covering size is at least $\frac{\varphi(n)}{\varphi(n,\ell)} = \varphi(n) = \Omega(\frac{n}{\log\log n}) = \Omega(\frac{n}{\ell}\frac{\log n}{\log\log n})$. $\square$

References

[1] Béla Bollobás, Svante Janson, and Oliver Riordan. On covering by translates of a set. *Random Structures & Algorithms*, 38(1-2):33–67, 2011.

[2] Zhixiong Chen, Igor E. Shparlinski, and Arne Winterhof. Covering sets for limited-magnitude errors. *IEEE Transactions on Information Theory*, 60(9):5315–5321, September 2013.

[3] Alina Carmen Cojocaru, M Ram Murty, et al. *An introduction to sieve methods and their applications*, volume 66. Cambridge University Press, 2006.

[4] Harold Davenport, Hugh L Montgomery, and Ann Arbor. *Multiplicative number theory*. Graduate Texts in Mathematics. Springer, New York, NY, 3 edition, October 2000.

[5] Lucia (https://mathoverflow.net/users/38624/lucia). Bounds for relative totient function for small values. MathOverflow. URL:https://mathoverflow.net/q/252852 (version: 2016-10-23).

[6] Anxiao Jiang, Michael Langberg, Moshe Schwartz, and Jehoshua Bruck. Trajectory codes for flash memory. *IEEE Transactions on Information Theory*, 59(7):4530–4541, 2013.

[7] Torleiv Kløve. On covering sets for limited-magnitude errors. *Cryptography and Communications*, 8(3):415–433, July 2016.

[8] Torleiv Klove, Jinquan Luo, and Somaye Yari. Codes correcting single errors of limited magnitude. *IEEE Transactions on Information Theory*, 58(4):2206–2219, 2012.

[9] Torleiv Kløve and Moshe Schwartz. Linear covering codes and error-correcting codes for limited-magnitude errors. *Designs, Codes and Cryptography*, 73(2):329–354, November 2014.

[10] Konstantinos Koiliaris and Chao Xu. Faster Pseudopolynomial Time Algorithms for Subset Sum. *ACM Transactions on Algorithms*, 15(3):1–20, July 2019.

[11] L. Lovász. On the ratio of optimal integral and fractional covers. *Discrete Mathematics*, 13(4):383–390, 1975.

[12] Oliver Roche-Newton, Ilya D Shkredov, and Arne Winterhof. PACKING SETS OVER FINITE ABELIAN GROUPS. page 9, 2018.

[13] Aravind Srinivasan. Improved Approximation Guarantees for Packing and Covering Integer Programs. *SIAM Journal on Computing*, 29(2):648–670, January 1999.

[14] S.K Stein. Two combinatorial covering theorems. *Journal of Combinatorial Theory, Series A*, 16(3):391–397, May 1974.

[15] Vijay V. Vazirani. *Approximation Algorithms*. Springer, Berlin ; New York, 2001.

A Brun's sieve

Theorem A.1 (Brun's sieve [3, p.93]) *Let $\mathcal{A}$ be any set of natural number $\leq x$ (i.e. $\mathcal{A}$ is a finite set) and let $\mathcal{P}$ be a set of primes. For each prime $p \in \mathcal{P}$, Let $\mathcal{A}_p$ be the set of elements of $\mathcal{A}$ which are divisible by p. Let $\mathcal{A}_1 := A$ and for any squarefree positive integer d composed of primes of $\mathcal{P}$ let $\mathcal{A}_d := \cap_{p|d} A_p$. Let z be a positive real number and let $P(z) := \prod_{p \in \mathcal{P}, p<z} p$.*
We assume that there exist a multiplicative function $\gamma(\cdot)$ such that, for any d as above,

$$|\mathcal{A}_d| = \frac{\gamma(d)}{d} X + R_d$$

for some R_d, where

$$X := |A|.$$

We set

$$S(\mathcal{A}, \mathcal{P}, z) := |\mathcal{A} \setminus \cup_{p|P(z)} \mathcal{A}_p| = |\{a : a \in \mathcal{A}, \gcd(a, P(z)) = 1\}|$$

and

$$W(z) := \prod_{p|P(z)} (1 - \frac{\gamma(p)}{p}).$$

Supposed that
1. $|R_d| \le \gamma(d)$ *for any squarefree* d *composed of primes of* $\mathcal{P}$*;*
2.there exists a constant $A_1 \ge 1$ *such that*

$$0 \le \frac{\gamma(p)}{p} \le 1 = \frac{1}{A_1};$$

3.there exists a constant $\kappa \ge 0$ *and* $A_2 \ge 1$ *such that*

$$\sum_{w \le p < z} \frac{\gamma(p) \log p}{p} \le \kappa \log \frac{z}{w} + A_2 \quad \text{if} \quad 2 \le w \le z.$$

4.Let b *be a positive integer and let* λ *be a real number satisfying*

$$0 \le \lambda e^{1+\lambda} \le 1.$$

Then

$$\begin{aligned} S(\mathcal{A}, \mathcal{P}, z) \ge & XW(z)\{1 - \frac{2\lambda^{2b} e^{2\lambda}}{1 - \lambda^2 e^{2+2\lambda}} \exp((2b+2)\frac{c_1}{\lambda \log z})\} \\ & + O(z^{2b-1+\frac{2.01}{e^{2\lambda/\kappa}-1}}), \end{aligned}$$

where

$$c_1 := \frac{A_2}{2}\{1 + A_1(\kappa + \frac{A_2}{\log 2})\}.$$

B Proof of theorem 2.2

We first show a simple lemma.

Lemma B.1 *Let* $y \in \mathbb{Z}_n^*$*, and* $B \subseteq \mathbb{Z}_n^*$*. The number of* $x \in \mathbb{Z}_{dn}^*$ *such that* $xb \equiv y \pmod n$*, and* $b \in B$ *is* $|B|\frac{\varphi(dn)}{\varphi(n)}$.

Proof: Indeed, the theorem is the same as finding the number of solutions to $x \equiv yb^{-1} \pmod n$ where $b \in B$. For a fixed b, let $z = yb^{-1}$. We are asking the number of $x \in \mathbb{Z}_{dn}^*$ such that $x \equiv z \pmod n$. Consider the set $A = \{z + kn \mid 0 \le k \le d-1\}$. Let the distinct prime factors set of n be P_n. Note $\gcd(z, n) = 1$, thus $p \in P_n$ can't divide any element in A. Let $P_{dn} \setminus P_n = P'_d \subseteq P_d$. Let q be the product of some elements in P'_d, $q|d$, $(q, n) = 1$. Let $A_q = \{a | a \in A, q|a\}$. Consider $q|z + kn \Longleftrightarrow k \equiv -zn^{-1} \pmod q$, and note $0 \le k \le d-1, q|d$, therefore $|A_q| = \frac{d}{q}$.
We can use the principle of inclusion-exclusion to count the elements $a \in A$ such that $\gcd(a, dn) = 1$

$$\sum_{i=0}^{|P'_d|}(-1)^i \sum_{S \subseteq P'_d, |S|=i} |A_{\prod_{p\in S} p}| = \sum_{i=0}^{|P'_d|}(-1)^i \sum_{S \subseteq P'_d, |S|=i} \frac{d}{\prod_{p \in S} p} = d \prod_{p \in P'_d}(1 - \frac{1}{p}) = \frac{\varphi(dn)}{\varphi(n)}.$$

Because all the solution sets of x for different $b \in B$ are disjoint, we obtain the total number of solutions over all B is $|B|\frac{\varphi(dn)}{\varphi(n)}$. □

Now we are ready to prove the theorem. Since $x \in \mathbb{Z}_n^*$, we see that $xb \equiv y \pmod n$ if and only if $d|b$, $x\frac{b}{d} \equiv \frac{y}{d} \pmod{\frac{n}{d}}$, and $\frac{b}{d} \le \lfloor \frac{\ell}{d} \rfloor$. We can then apply lemma B.1 and obtain the number of solutions is $\varphi(n/d, \lfloor \ell/d \rfloor)\varphi(n)/\varphi(n/d)$.