EXTREMAL COPRIME COVERINGS UNDER A RECIPROCAL BUDGET AND A DICKMAN-TYPE CONJECTURE

PRZEMYSLAW CHOJECKI

ABSTRACT. Fix $C > 0$ and let n be large. Among sets $A \subseteq \{2, \dots, n\}$ of pairwise coprime integers with reciprocal budget $\sum_{a \in A} \frac{1}{a} \le C$, one asks for the choice of A that minimizes the number of integers $m \le n$ not divisible by any $a \in A$. This is Erdős Problem #783, posed in [4] (see also [1]).

We prove that for $0 \le C \le \log 2$ the optimal asymptotic uncovered density equals $1 - C$, and it is achieved (up to $o(1)$) by taking A to be a set of the largest primes $\{p : y < p \le n\}$ with $y = n^{e^{-C} + o(1)}$. Beyond $\log 2$ we record a natural construction that links the problem to friable (smooth) integers and yields an upper bound in terms of the Dickman–de Bruijn function ρ:

$$\inf_A \frac{1}{n} \#\{m \le n : \forall a \in A,\ a \nmid m\} \ \le\ \rho(e^C) + o(1) \qquad (n \to \infty).$$

We then formulate a conjecture predicting $\rho(e^C)$ is the true optimum for all fixed $C > 0$, with asymptotically extremal sets given by a tail of primes. A potential strategy, via Buchstab-type recurrences and a variational relaxation, is proposed.

1. Introduction

Let $C > 0$ be fixed and n large. We consider sets $A \subseteq \{2, \dots, n\}$ satisfying

(i) $\gcd(a, b) = 1$ for all distinct $a, b \in A$ (pairwise coprime), and

(ii) $\displaystyle\sum_{a \in A} \frac{1}{a} \le C$ (a reciprocal "budget").

Such a set A *covers* an integer $m \le n$ if $a \mid m$ for some $a \in A$. Write

$$U(A; n) := \#\big\{1 \le m \le n : \forall a \in A,\ a \nmid m\big\} \tag{1.1}$$

for the number of *uncovered* integers up to n. Erdős asked [4] for the choice of admissible A that minimizes $U(A; n)$, in the regime where C is fixed and $n \to \infty$; this appears as Erdős Problem #783 in Bloom's online database [1].

A natural candidate is the *prime tail*: take consecutive primes $n \ge q_1 > q_2 > \cdots$ in decreasing order and choose $A = \{q_1, \dots, q_k\}$ where k is maximal subject to $\sum_{i=1}^{k} \frac{1}{q_i} \le C$. Heuristically, for primes one expects a density prediction

$$\frac{U(A; n)}{n} \approx \prod_{p \in A} \Big(1 - \frac{1}{p}\Big) \approx \exp\Big(-\sum_{p \in A} \frac{1}{p}\Big),$$

and for fixed budget C one wants the approximation to be accurate, which points toward taking the $1/p$ as small as possible, i.e. primes as large as possible. However, overlaps (integers divisible by two members of A) matter. For a tail of primes above a threshold y, overlaps disappear exactly when $y \ge \sqrt{n}$, corresponding to $C \le \log 2$; this explains the special role of $\log 2$.

Date: January 23, 2026.

Main contributions. We establish the conjectured optimum in the overlap-free range.

Theorem 1.1 (Optimality for $C \le \log 2$)**.** *Fix* $0 \le C \le \log 2$. *Then, as* $n \to \infty$,

$$\inf_{\substack{A\subseteq\{2,\dots,n\}\\ (a,b)=1\ \forall a\neq b\\ \sum_{a\in A} 1/a \le C}} \frac{U(A;n)}{n} \;=\; 1 - C + o(1).$$

Moreover, this bound is achieved by a set of primes of the form $A = \{p : \ y < p \le n\}$ *with* $y = n^{e^{-C}+o(1)}$ *(equivalently, by taking the largest primes until the reciprocal budget is exhausted).*

For general C we record the prime-tail construction's connection to smooth numbers and hence to the Dickman–de Bruijn function.

Theorem 1.2 (A Dickman-type upper bound for general C)**.** *Fix* $C > 0$. *Let* $y = y(n)$ *satisfy* $\log y = (e^{-C} + o(1)) \log n$ *as* $n \to \infty$, *and set* $A = \{p \in \mathbb{P} : \ y < p \le n\}$ *(removing finitely many boundary primes if needed to enforce* $\sum_{p\in A} 1/p \le C$*). Then*

$$\frac{U(A;n)}{n} = \frac{\Psi(n,y)}{n} = \rho(e^C) + o(1),$$

where $\Psi(n, y)$ *counts* y*-smooth integers* $\le n$, *and* ρ *is the Dickman–de Bruijn function. In particular,*

$$\inf_A \frac{U(A;n)}{n} \le \rho(e^C) + o(1).$$

We then formulate a conjecture that $\rho(e^C)$ is the true optimum for all $C > 0$, and outline a possible proof strategy.

2. Setup and basic bounds

Definition 2.1. For fixed $C > 0$ and $n \in \mathbb{N}$, let $\mathcal{A}_n(C)$ denote the family of sets $A \subseteq \{2, \dots, n\}$ such that

$$\gcd(a, b) = 1 \text{ for distinct } a, b \in A, \qquad \sum_{a\in A} \frac{1}{a} \le C.$$

Define the extremal uncovered proportion

$$F_n(C) := \inf_{A\in\mathcal{A}_n(C)} \frac{U(A;n)}{n}. \tag{2.1}$$

A first bound follows from the union bound and the budget constraint.

Proposition 2.2 (A universal lower bound)**.** *For every* $C > 0$ *and* $n \ge 1$,

$$F_n(C) \ge \max(0, 1 - C).$$

More precisely, for any $A \in \mathcal{A}_n(C)$,

$$\frac{U(A;n)}{n} \;\ge\; 1 - \sum_{a\in A} \frac{1}{a} \;\ge\; 1 - C.$$

Proof. Let $M(a) := \{m \le n : \ a \mid m\}$. Then $\#M(a) = \lfloor n/a \rfloor \le n/a$, so by the union bound,

$$n - U(A;n) = \# \bigcup_{a\in A} M(a) \le \sum_{a\in A} \#M(a) \le n \sum_{a\in A} \frac{1}{a}.$$

Rearrange and use $\sum_{a\in A} 1/a \le C$. □

The point is that *every* modulus a removes about n/a integers and costs $1/a$ from the budget, so all moduli have the same first-order efficiency. The optimization is therefore driven by overlap effects among the removed sets.

3. The $C \le \log 2$ regime

In this section we prove Theorem 1.1. The key observation is that if all moduli lie in $(\sqrt{n}, n]$, then no integer $m \le n$ can be divisible by two distinct moduli from A, because the product of any two exceeds n.

3.1. A prime-tail construction with no overlaps.

Let $y \ge \sqrt{n}$ and define the prime tail

$$A_y(n) := \{p \in \mathbb{P} : \; y < p \le n\}. \tag{3.1}$$

Then the sets $M(p)$ for $p \in A_y(n)$ are disjoint, because for distinct $p, q > y$ one has $pq > y^2 \ge n$.

Lemma 3.1 (Disjointness above $\sqrt{n}$)**.** *If $y \ge \sqrt{n}$, then for distinct primes $p, q > y$ the sets $M(p)$ and $M(q)$ are disjoint. Consequently,*

$$n - U(A_y(n); n) = \sum_{y<p\le n} \left\lfloor \frac{n}{p} \right\rfloor, \qquad \text{and hence} \quad \frac{U(A_y(n); n)}{n} = 1 - \sum_{y<p\le n} \frac{1}{p} + o(1).$$

Proof. If $m \le n$ is divisible by both p and q, then $pq \mid m \le n$, impossible if $pq > n$. Thus the unions are disjoint and the covered set size is the sum of the sizes. For the asymptotic, use $\lfloor n/p \rfloor = n/p + O(1)$ and note that $\#\{p \in (y, n]\} \ll n/\log n = o(n)$. □

3.2. Tuning the budget and completing the proof.

To match a given $C \le \log 2$, we choose y so that $\sum_{y<p\le n} 1/p \approx C$. This uses Mertens' second theorem:

$$\sum_{p\le x} \frac{1}{p} = \log\log x + M + o(1) \qquad (x \to \infty), \tag{3.2}$$

where M is the Meissel–Mertens constant (see, e.g., [7, 5]).

Lemma 3.2 (Budget matching for $C \le \log 2$)**.** *Fix $0 \le C \le \log 2$ and let $n \to \infty$. Set*

$$y := \exp\big(e^{-C} \log n\big) = n^{e^{-C}}. \tag{3.3}$$

Then $y \ge \sqrt{n}$ and

$$\sum_{y<p\le n} \frac{1}{p} = C + o(1).$$

In particular, for n large one can delete finitely many boundary primes from $A_y(n)$ to obtain a set $A \subseteq \{2, \ldots, n\}$ with $\sum_{p\in A} \frac{1}{p} \le C$ and $\sum_{p\in A} \frac{1}{p} = C + o(1)$.

Proof. Since $C \le \log 2$, we have $e^{-C} \ge 1/2$, hence $y = n^{e^{-C}} \ge \sqrt{n}$. By (3.2),

$$\sum_{y<p\le n} \frac{1}{p} = \big(\log\log n + M + o(1)\big) - \big(\log\log y + M + o(1)\big) = \log\Big(\frac{\log n}{\log y}\Big) + o(1).$$

But $\log y = e^{-C} \log n$, so $\log(\log n / \log y) = \log(e^{C}) = C$. The adjustment by deleting finitely many boundary primes changes the sum by at most $O(1/y) = o(1)$. □

Proof of Theorem 1.1. The lower bound $F_n(C) \geq 1 - C$ is Proposition 2.2. For the upper bound, take $y = n^{e^{-C}}$ and $A = A_y(n)$ adjusted as in Lemma 3.2. Then $y \geq \sqrt{n}$ and Lemma 3.1 applies:

$$\frac{U(A;n)}{n} = 1 - \sum_{p \in A} \frac{1}{p} + o(1) = 1 - C + o(1).$$

Taking the infimum over A yields $F_n(C) \leq 1 - C + o(1)$, matching the lower bound. □

Remark 3.3 (Why "largest primes" appear)**.** In the proof above, any set of pairwise coprime moduli lying in $(\sqrt{n}, n]$ has disjoint covered sets. To maximize coverage under a reciprocal budget, one thus wants as many moduli as possible in that high range, and among primes this becomes exactly "take the largest primes until the budget is used".

4. Prime tails, smooth numbers, and a Dickman-type bound

For $C > \log 2$ the optimal construction is expected to involve overlaps. The prime-tail construction still has a clean interpretation in terms of friable integers.

4.1. **From excluding large primes to counting smooth numbers.** Fix $y = y(n)$ and consider $A_y(n)$ as in (3.1). An integer $m \leq n$ is *not* divisible by any prime $p \in (y, n]$ if and only if all its prime factors are $\leq y$, i.e. m is y-smooth. Thus, with $\Psi(n, y)$ denoting the number of y-smooth integers $\leq n$,

$$U(A_y(n); n) = \Psi(n, y). \tag{4.1}$$

The distribution of smooth numbers is controlled by the Dickman–de Bruijn function ρ. For completeness:

Definition 4.1 (Dickman–de Bruijn function)**.** The Dickman–de Bruijn function $\rho : [0, \infty) \to \mathbb{R}$ is defined by $\rho(u) = 1$ for $0 \leq u \leq 1$ and, for $u > 1$, by the delay differential equation

$$u\rho'(u) + \rho(u - 1) = 0,$$

together with continuity on $[0, \infty)$.

A classical theorem of Dickman, de Bruijn, and Ramaswami (see [3, 2, 5, 6]) states that for fixed $u \geq 1$,

$$\Psi\big(n, n^{1/u}\big) = n\rho(u) + O\Big(\frac{n}{\log n}\Big) \qquad (n \to \infty). \tag{4.2}$$

4.2. **Matching the budget for general C.** Using Mertens' theorem (3.2), the choice $y = n^{e^{-C}}$ satisfies

$$\sum_{y < p \leq n} \frac{1}{p} = C + o(1),$$

and after deleting finitely many primes one can enforce the constraint $\sum_{p \in A} 1/p \leq C$ while maintaining $\sum_{p \in A} 1/p = C + o(1)$. With $y = n^{e^{-C} + o(1)}$, we have

$$u := \frac{\log n}{\log y} = e^C + o(1).$$

Combining (4.1) with (4.2) yields Theorem 1.2.

Proof of Theorem 1.2. Let y satisfy $\log y = (e^{-C} + o(1)) \log n$. Let $A = \{p : y < p \le n\}$ with finitely many boundary primes removed so that $\sum_{p\in A} 1/p \le C$. Then $U(A; n) = \Psi(n, y)$ by (4.1). Write $y = n^{1/u}$ so $u = \log n/ \log y = e^C + o(1)$. By (4.2) (valid for fixed u, and stable under $o(1)$ perturbations in u),

$$\Psi(n, y) = n\rho(u) + o(n) = n\rho(e^C) + o(n),$$

which implies the desired bound. □

Remark 4.2 (Consistency with the log 2 theorem)**.** For $1 \le u \le 2$ one has the explicit formula $\rho(u) = 1-\log u$. Thus when $C \le \log 2$ (so $u = e^C \le 2$), $\rho(e^C) = 1-C$, recovering Theorem 1.1.

5. Conjectures

We now state the conjectural global picture suggested by Erdős Problem #783.

Conjecture 5.1 (Dickman optimum)**.** *For each fixed $C > 0$ the limit*

$$F(C) := \lim_{n\to\infty} F_n(C)$$

exists and equals

$$F(C) = \rho(e^C).$$

Conjecture 5.2 (Prime-tail extremisers)**.** *For each fixed $C > 0$, asymptotically extremal sets may be chosen to be prime tails: there exists $y = y(n)$ such that an optimal (or near-optimal) admissible set is*

$$A = \{p \in \mathbb{P} : y < p \le n\},$$

equivalently the set of consecutive primes $q_1 > q_2 > \cdots$ below n truncated at maximal k with $\sum_{i=1}^{k} 1/q_i \le C$.

Remark 5.3 (Largest vs. smallest primes)**.** A different natural competitor is $A = \{2, 3, \ldots, p_k\}$ where k is maximal with $\sum_{i=1}^{k} 1/p_i \le C$. Conjecture 5.2 asserts that, for minimizing $U(A; n)$, pushing the primes to the top of the range (near n) is asymptotically better. Theorem 1.1 confirms this in the entire range $C \le \log 2$.

6. A possible strategy beyond log 2

We briefly sketch a route toward Conjectures 5.1–5.2. This is not a proof, but it aims to organize the problem into a form amenable to existing analytic number theory tools.

6.1. Step 1: reduction heuristics toward primes.

Because the moduli are pairwise coprime, each $a \in A$ introduces a new set of prime factors. A composite modulus $a = p_1 \cdots p_r$ only removes multiples divisible by *all* of $p_1, \ldots, p_r$, which is a much smaller set than the union of multiples of the individual primes. At the same time, the reciprocal cost $1/a$ is also smaller than $\sum 1/p_i$, so the first-order "coverage per unit budget" remains comparable. Thus the advantage of composites, if any, could only come from reduced overlap patterns. A plausible strengthening of Conjecture 5.2 is that *primes are asymptotically optimal* among all coprime moduli, in the sense that any near-extremiser can be replaced by a prime set at negligible loss.

6.2. Step 2: from prime sets to friable counts.

If A is a set of primes, the uncovered integers are exactly those with no prime factor from A. If, further, A is a tail $\{p : y < p \le n\}$ then uncovered integers are exactly y-smooth integers, giving the explicit objective $\Psi(n, y)$. Thus Conjecture 5.2 upgrades the problem to showing that *among prime sets with a given prime-harmonic mass $\sum_{p\in A} 1/p \le C$, the uncovered count is minimized by taking A as a tail.*

6.3. Step 3: Buchstab-type recurrences and a variational relaxation. Smooth-number counts satisfy Buchstab's identity (see [5, 7]): for $x \geq 1$ and $z \geq y > 1$,

$$\Psi(x, y) = \Psi(x, z) - \sum_{y<p\leq z} \Psi(x/p, p). \tag{6.1}$$

Iterating (6.1) and passing to a continuous scale in $\log p/\log x$ leads (heuristically, and in many proven asymptotic regimes) to the Dickman equation that defines ρ.

One potential approach is to encode the choice of prime set A by a measure on the logarithmic scale: for $p \leq n$ set $t := \frac{\log p}{\log n} \in (0, 1]$ and consider a density μ supported on $(0, 1]$. The budget constraint $\sum_{p\in A} 1/p \leq C$ resembles a constraint of the form

$$\int e^{-t \log n} \, d\mu(t) \approx \int n^{-t} \, d\mu(t) \leq C,$$

while the uncovered count is driven by a recursion derived from (6.1). In this relaxation, a prime tail corresponds to μ supported on $[e^{-C}, 1]$ and leads to $u = e^C$ and hence $\rho(u)$.

To turn this into a proof, one would like:

- a *stability theorem* showing near-optimal discrete prime sets must resemble a continuous optimizer, and
- an *extremality principle* (a calculus-of-variations argument) showing that, under the appropriate recursion, the minimizing measure is an interval (i.e. a tail).

The difficulty is that sieve recurrences like (6.1) are nonlinear in the combinatorial choice of A, and error control becomes delicate when overlaps are substantial (the "$C > \log 2$" regime).

6.4. Outlook. Theorems 1.1 and 1.2 suggest that Erdős' guess is correct at least in the overlap-free region and remains compatible with the known theory of friable integers for general C. Proving matching lower bounds for $C > \log 2$ appears to require genuinely sharp sieve inequalities tailored to the *adversarial* selection of a coprime set of moduli under a reciprocal budget.

References

[1] T. F. Bloom, *Erdős Problem #783*, https://www.erdosproblems.com/783, accessed 2026-01-23.

[2] N. G. de Bruijn, *On the number of positive integers $\leq x$ and free of prime factors $> y$*, Indagationes Mathematicae **13** (1951), 50–60.

[3] K. Dickman, *On the frequency of numbers containing prime factors of a certain relative magnitude*, Arkiv för Matematik, Astronomi och Fysik **22A** (10) (1930), 1–14.

[4] P. Erdős, *Problems and results on combinatorial number theory*, in *A survey of combinatorial theory* (Proc. Internat. Sympos., Colorado State Univ., Fort Collins, Colo., 1971), North-Holland, Amsterdam, 1973, pp. 117–138.

[5] A. Hildebrand, *Integers without large prime factors*, in *Analytic number theory* (Allerton Park, IL, 1989), Progr. Math. **85**, Birkhäuser Boston, Boston, MA, 1990, pp. 187–201. (See also the expanded survey https://tenenb.perso.math.cnrs.fr/PPP/Survey_friables.pdf.)

[6] V. Ramaswami, *On the number of positive integers less than x and free of prime divisors greater than x^c*, Bull. Amer. Math. Soc. **55** (1949), 1122–1127.

[7] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge Studies in Advanced Mathematics, 46, Cambridge University Press, Cambridge, 1995.