SIEVES OF DIMENSION $1+\epsilon$

ZARATHUSTRA BRADY

1. Introduction

Let A be a (possibly weighted) set of whole numbers, and for each positive integer d set $A_d = \{a \in A, d \mid a\}$. Let κ be a real number and by abuse of notation let $\kappa : \mathbb{N} \to \mathbb{R}$ be a multiplicative function satisfying $0 \le \kappa(p) < p$ for all p, and

$$\sum_{p \le x} \kappa(p) \frac{\log(p)}{p} = (\kappa + o(1)) \log(x).$$

Suppose that z, y are such that for every squarefree integer d, all of whose prime factors are less than z, we have

$$\left| |A_d| - \kappa(d)\frac{y}{d} \right| \le \kappa(d), \tag{1}$$

or alternatively such that for some fixed $\epsilon > 0$ and every such d we have

$$\left| |A_d| - \kappa(d)\frac{y}{d} \right| \le \kappa(d)\frac{y}{d \log(y/d)^{2\kappa+\epsilon}}. \tag{2}$$

In particular, we have $|A| = y + O(1)$ in the first case, or $|A| = y + O(y/\log(y)^{2\kappa+\epsilon})$ in the second case. We want to estimate the quantity

$$\mathcal{S}(A, z) = \left| \{a \in A, \forall p < z \ (a, p) = 1\} \right|.$$

Suppose now that $y = z^s$, s a constant, y, z going to infinity. Define sifting functions $f_\kappa(s), F_\kappa(s)$ by

$$(1 + o(1)) f_\kappa(s) y \prod_{p<z} \left(1 - \frac{\kappa(p)}{p}\right) \le \mathcal{S}(A, z) \le (1 + o(1)) F_\kappa(s) y \prod_{p<z} \left(1 - \frac{\kappa(p)}{p}\right),$$

with $f_\kappa(s)$ as large as possible (resp. $F_\kappa(s)$ as small as possible) given that the above inequality holds for all choices of A satisfying (1). Selberg [2] has shown (in a much more general context) that the functions $f_\kappa(s), F_\kappa(s)$ are continuous, monotone, and computable for $s > 1$, that they do not change if we replace (1) with (2), and that they tend to 1 exponentially as s goes to infinity.

More specifically, $f_\kappa(s)$ and $F_\kappa(s)$ can be defined as follows. Let $\mathcal{M}$ be the collection of all finite multisubsets of $[0, 1]$, and for $S \in \mathcal{M}$ let $\Sigma(S)$ be the sum of the elements of S and $|S|$ be the number of elements of S (both counted with multiplicity). When we write sums like $\sum_{A \subseteq S}$, we also count subsets A with multiplicity, so such a sum will always have $2^{|S|}$ summands. Let $\lambda : \mathcal{M} \to \mathbb{R}$ be a piecewise continuous function supported on S with $\Sigma(S) \le 1$, and define a function $\theta : \mathcal{M} \to \mathbb{R}$ by

$$\theta(S) = \sum_{A \subseteq S} \lambda(A).$$

We say that (λ, θ) forms an upper (resp. lower) bound sieve with sifting limit s if λ is supported on multisubsets of $[0, \frac{1}{s}]$, $\theta(\emptyset) = \lambda(\emptyset) \ge 1$ (resp. $\theta(\emptyset) \le 1$), and $\theta(S) \ge 0$ (resp. $\theta(S) \le 0$) for all

$S \subseteq [0, \frac{1}{s}]$ with $|S| \geq 1$. Then

$$F_\kappa(s) = \inf_{(\lambda,\theta)\geq 0} \sum_{n=0}^{\infty} \frac{\kappa^n}{n!} \int_0^{\frac{1}{s}} \cdots \int_0^{\frac{1}{s}} \theta(x_1, ..., x_n) \frac{dx_1}{x_1} \cdots \frac{dx_n}{x_n}, \tag{3}$$

where the infimum is over all upper bound sieves (λ, θ) with sifting limit $\frac{1}{s}$, and there is a similar formula for $f_\kappa(s)$ (note that when $f_\kappa(s) = 0$, we will typically have $\lambda(\emptyset) = 0$).

The Selberg upper bound sieve corresponds to choosing $\theta = \theta'^2$ for some other sieve (l, θ'), with l supported on $\Sigma(S) \leq \frac{1}{2}$. In terms of the sieve weights λ, this corresponds to

$$\lambda(S) = \sum_{A\cup B=S} l(A)l(B).$$

In order to describe the weights l, we use the following generalization of the Dickman function. For $s < 0$ we set $\rho_\kappa(s) = 0$, for $0 < s \leq 1$ we set $\rho_\kappa(s) = 1$, and for $s \geq 1$ we define $\rho_\kappa(s)$ by the differential-difference equation

$$s^\kappa \rho_\kappa'(s) = -\kappa(s-1)^{\kappa-1}\rho_\kappa(s-1),$$

or equivalently by the integral equation

$$s^\kappa \rho_\kappa(s) = \int_{s-1}^{s} \rho_\kappa(t) dt^\kappa.$$

When κ is a whole number, the function $\rho_\kappa(s)$ has a combinatorial interpretation. Let n be large, and consider the collection of all ordered pairs (π, c) where π is a permutation of $\{1, ..., n\}$ and $c : \{1, ..., n\} \to \{1, ..., \kappa\}$ is a compatible coloring of $\{1, ..., n\}$ (i.e. $c(i) = c(\pi(i))$ for all i). Choosing an ordered pair (π, c) uniformly at random, $\rho_\kappa(s)$ is the limit, as n goes to ∞, of the probability that every cycle of π has length at most $\frac{n}{s}$.

The optimal choice for the weights l is given in terms of ρ_κ by

$$l(S) = (-1)^{|S|} \frac{\int_0^{\frac{s}{2}-s\Sigma(S)} \rho_\kappa(t) dt^\kappa}{\int_0^{\frac{s}{2}} \rho_\kappa(t) dt^\kappa}.$$

When s goes to ∞ this becomes

$$l(S) \approx \begin{cases} (-1)^{|S|} & \text{if } \Sigma(S) < \frac{1}{2}, \\ 0 & \text{else,} \end{cases}$$

and when $s \leq 2$ it becomes

$$l(S) = (-1)^{|S|}(1 - 2\Sigma(S))_+^\kappa.$$

Setting

$$\sigma_\kappa(s) = \frac{\int_0^{\frac{s}{2}} \rho_\kappa(t) dt^\kappa}{e^{\gamma\kappa}\Gamma(\kappa+1)},$$

we have

$$\begin{aligned} s^{-\kappa}\sigma_\kappa(s) &= \frac{1}{(2e^\gamma)^\kappa \Gamma(\kappa+1)} & 0 < s \leq 2, \\ (s^{-\kappa}\sigma_\kappa(s))' &= -\kappa s^{-\kappa-1}\sigma_\kappa(s-2) & s \geq 2, \end{aligned}$$

and the Selberg sieve gives us the upper bound

$$F_\kappa(s) \leq \frac{1}{\sigma_\kappa(s)}.$$

The only case in which this is known to be optimal is when $\kappa = 1$ and $s \le 2$, in which case the Selberg sieve (λ^S, θ^S) is given by

$$\lambda^S(S) = (-1)^{|S|} \sum_{A \cup B = S} (-1)^{|A \cap B|}(1 - 2\Sigma(A))_+(1 - 2\Sigma(B))_+,$$

$$\theta^S(S) = \Big(\sum_{A \subseteq S} (-1)^{|A|}(1 - 2\Sigma(A))_+ \Big)^2.$$

For $\Sigma(S) \le \frac{1}{2}$, we have

$$\lambda^S(S) = (-1)^{|S|}\big(1 - 4\sum_{x \in S} x^2\big).$$

The β-sieve $(\lambda^\beta, \theta^\beta)$ is given as follows. The formula

$$\lambda^\beta(S) = \begin{cases} (-1)^{|S|} & \text{if } \forall A \subseteq S,\ |A| \text{ odd} \implies \Sigma(A) + \beta \min(A) \le 1, \\ 0 & \text{else,} \end{cases}$$

gives the upper bound sieve weights, while the lower bound sieve weights are given by the same formula with "odd" replaced by "even". Here β is chosen such that $\beta - 1$ is the largest zero of the function $q(s)$, where q solves the differential-difference equation

$$(sq(s))' = \kappa q(s) + \kappa q(s+1).$$

When κ is a half-integer, $q(s)$ is a polynomial of degree $2\kappa - 1$ and β is an algebraic number (see [1] for details). When $\kappa = 1$, we have $\beta = 2$.

The β-sieve is best understood in terms of Buchstab iteration:

$$\mathcal{S}(A, z) = |A| - \sum_{p<z} \mathcal{S}(A_p, p).$$

This leads to the inequalities

$$s^\kappa f_\kappa(s) \ge s^\kappa - \kappa \int_{t>s} t^{\kappa-1}(F_\kappa(t-1) - 1)dt,$$

$$s^\kappa F_\kappa(s) \le s^\kappa + \kappa \int_{t>s} t^{\kappa-1}(1 - f_\kappa(t-1))dt.$$

A variant of Buchstab iteration is given by

$$\mathcal{S}(A, z) = \mathcal{S}(A, w) - \sum_{w \le p < z} \mathcal{S}(A_p, p)$$

for any $w \le z$. If $y = w^t$ and we already have an upper bound sieve $(\lambda_t^+, \theta_t^+)$ with sifting limit t and lower bound sieves $(\lambda_u^-, \theta_u^-)$ with sifting limit u for $s - 1 \le u \le t - 1$, the upper bound sieve (λ', θ') we obtain from Buchstab iteration is given by

$$\lambda'(S) = \begin{cases} \lambda_t^+(S) & \text{if } S \subseteq [0, \frac{1}{t}), \\ -\lambda^-_{\frac{1}{x}-1}(T) & \text{if } S = T \cup \{x\}, T \subseteq [0, x], \frac{1}{t} \le x < \frac{1}{s}. \end{cases}$$

When $\kappa = 1$, the optimal sifting functions f, F are fixed points of Buchstab iteration. To see they are optimal, we introduce two weighted sets A^+, A^- satisfying (2). Both are supported on $[1, y]$, with the weight on n given by $1 - \lambda(n)$ in A^+ and given by $1 + \lambda(n)$ in A^-, where by $\lambda(n)$ we mean $(-1)^{\Omega(n)}$ (and not a sieve weight). Setting

$$\pi^\pm(y, z) = \mathcal{S}(A^\pm, z),$$

we have

$$\pi^{\pm}(y,z) = \pi^{\pm}(y,w) - \sum_{w<p<z} \pi^{\mp}(y/p,p),$$

and by the prime number theorem, for $1 < s < 3$ we have

$$\pi^{+}(y,z) = 2(\pi(y)-\pi(z)) = \frac{2e^{\gamma}}{s}\frac{y}{e^{\gamma}\log(z)} + O\left(\frac{y}{\log(z)^2}\right),$$

so for all $s > 1$ we have

$$\pi^{+}(y,z) = F(s)\frac{y}{e^{\gamma}\log(z)} + O\left(\frac{y}{\log(z)^2}\right),$$
$$\pi^{-}(y,z) = f(s)\frac{y}{e^{\gamma}\log(z)} + O\left(\frac{y}{\log(z)^2}\right).$$

Our strategy for constructing sieves in dimension $1+\epsilon$ is to find an optimal upper bound sieve (λ,θ) in dimension 1 (i.e., a sieve such that the expression inside the infimum on the right hand side of (3) is equal to $F(s)$) such that the sum

$$\sum_{n=1}^{\infty} \frac{1}{(n-1)!}\int_0^{\frac{1}{s}} \cdots \int_0^{\frac{1}{s}} \theta(x_1,...,x_n)\frac{dx_1}{x_1}\cdots\frac{dx_n}{x_n}$$

is as small as possible, since this sum is the rate of change of the expression inside the infimum on the right hand side of (3) at $\kappa = 1$. For (λ,θ) an optimal upper bound sieve with sifting limit $2 \le s \le 3$, set

$$a_n^{\theta} = \frac{1}{n!}\int_0^{\frac{1}{2}} \cdots \int_0^{\frac{1}{2}} \theta(x_1,...,x_n)\frac{dx_1}{x_1}\cdots\frac{dx_n}{x_n}.$$

We then have $a_0^{\theta} = 1, a_n^{\theta} \ge 0$, and

$$e^{\gamma} = F(2) = 1 + a_1^{\theta} + a_2^{\theta} + \cdots,$$

while the quantity we wish to minimize is

$$a_1^{\theta} + 2a_2^{\theta} + 3a_3^{\theta} + \cdots.$$

Note that this is the same as *maximizing* the quantity

$$2e^{\gamma} - 2 - (a_1^{\theta} + 2a_2^{\theta} + 3a_3^{\theta} + \cdots) = a_1^{\theta} - a_3^{\theta} - 2a_4^{\theta} - \cdots.$$

As a consequence, it seems that a good rule of thumb is to simply try to maximize $a_1^{\theta} = \int_0^{\frac{1}{2}} \theta(x)\frac{dx}{x}$. Letting $a_n^S = a_n^{\theta^S}, a_n^{\beta} = a_n^{\theta^{\beta}}$, we have

$$a_1^S = \frac{1}{2},\ a_2^S = \frac{\pi^2-9}{12} \approx 0.0724,\ a_3^S \approx 0.03966,$$

and

$$a_1^{\beta} = \log(3/2) \approx 0.405,\ a_2^{\beta} = \frac{\log(3/2)^2}{2} \approx 0.0822,\ a_3^{\beta} \approx 0.06705.$$

Additionally, from the analysis of the Selberg sieve we have

$$e^{\gamma} = \frac{\partial}{\partial\kappa}e^{\gamma\kappa}\Gamma(\kappa+1)\Big|_{\kappa=1} = a_1^S + 2a_2^S + 3a_3^S + \cdots.$$

2. Constraints on optimal sieves in dimension 1

The "complementary slackness" constraints on optimal solutions to linear optimization problems imply that if A is a weighted set satisfying (2) with $\mathcal{S}(A,z)$ maximal and (λ,θ) is an optimal upper bound sieve, then if for d squarefree we set $S_d = \{\frac{\log(p)}{\log(y)} \text{ s.t. } p \mid d\}$ we get

$$\begin{aligned} p \mid n, p < z, n \in A &\implies \theta(S_n) = 0, \\ \lambda(S_d) > 0 &\implies |A_d| - \frac{y}{d} = \frac{y}{d\log(y/d)^{2+\epsilon}}, \\ \lambda(S_d) < 0 &\implies |A_d| - \frac{y}{d} = -\frac{y}{d\log(y/d)^{2+\epsilon}}. \end{aligned}$$

We know that the set A^+ maximizes $\mathcal{S}(A,z)$ to first order. Since the number of $n \in A^+$ with $n \le y^{1-\epsilon}$ is small for any $\epsilon > 0$, while the number of $n \in A^+$ with $S_n \approx S$ is large if $\Sigma(S) = 1$, we conclude that, at least away from a measure zero set,

$$\Sigma(S) = 1,\ \min(S) < \frac{1}{s},\ |S| \text{ odd} \implies \theta(S) = 0 \tag{O}$$

for any optimal sieve, and it seems that any nice upper bound sieve satisfying (O) is optimal (although making this precise is tricky).

Proposition 1. *If $|S|$ is odd, $\min(S) < \frac{1}{s}$, and $\Sigma(S) = 1$, then $\theta^S(S) = 0$ and $\theta^\beta(S) = 0$ outside the measure zero subset where the three smallest elements of S are all equal.*

Proof. Although this morally follows from the fact that the Selberg sieve and the β sieve are optimal, we will give a direct proof. We have

$$\theta^S(S) = \left(\sum_{A\subseteq S} (-1)^{|A|}(1-2\Sigma(A))_+\right)^2,$$

and from $\Sigma(S) = 1$ and $|S|$ odd we have

$$(-1)^{|A|}(1-2\Sigma(A))_+ + (-1)^{|S\setminus A|}(1-2\Sigma(S\setminus A))_+ = (-1)^{|A|}(1-2\Sigma(A))$$

for $A \subseteq S$. If $S = \{x_1, ..., x_n\}$, then since $|S| \ge 2$ we have

$$\sum_{A\subseteq S}(-1)^{|A|}(1-2\Sigma(A))_+ = \frac{1}{2}\sum_{A\subseteq S}(-1)^{|A|}(1-2\Sigma(A)) = \frac{1}{2}\sum_{A\subseteq S}(-1)^{|A|} - \sum_{i=1}^{n} x_i \sum_{x_i\in A\subseteq S}(-1)^{|A|} = 0.$$

Now we turn to $\theta^\beta(S)$. Supposing that $x_1 \ge \cdots \ge x_n$, we just need to show that for all $A \subseteq S\setminus\{x_n\}$ we have $\lambda^\beta(A) \ne 0 \iff \lambda^\beta(A\cup\{x_n\}) \ne 0$. The only case in which this is not obvious is when $|A|$ is even and $\Sigma(A) + x_n + 2x_n > 1$, and in this case we have

$$\Sigma(A) > 1 - 3x_n \ge 1 - x_n - x_{n-1} - x_{n-2} = \Sigma(S \setminus \{x_n, x_{n-1}, x_{n-2}\}),$$

so in fact we must have $A = S \setminus \{x_n\}$. But then from $\lambda^\beta(A) \ne 0$ we must have

$$x_1 + \cdots + x_{n-2} + 2x_{n-2} \le 1 = x_1 + \cdots + x_{n-2} + x_{n-1} + x_n,$$

so in fact we must have $x_{n-2} = x_{n-1} = x_n$. □

Proposition 2. *If (λ,θ) is an upper bound sieve with sifting limit s satisfying (O), then for any $0 \le x < \min(\frac{1}{s}, 1-\frac{2}{s})$ we have $\lambda(x) = -1$.*

More generally, if S is a set with $\min(S) < \frac{1}{s}$ and either $|S|$ odd and $\Sigma(S) < 1-\frac{2}{s}$ or $|S|$ even and $\Sigma(S) < 1-\frac{1}{s}$, then $\theta(S) = 0$. In particular, if S is any set such that $\max(S) < \frac{1}{s}$ and $\Sigma(S) < 1-\frac{2}{s}$, then $\lambda(S) = (-1)^{|S|}$.

Proof. Note that $\frac{1-x}{2} > \frac{1}{s}$, so for any set A containing $\frac{1-x}{2}$ we have $\lambda(A) = 0$. Taking $S = \{x, \frac{1-x}{2}, \frac{1-x}{2}\}$ in (O), we have

$$0 = \theta(S) = \sum_{A \subseteq S} \lambda(A) = 1 + \lambda(x),$$

so $\lambda(x) = -1$.

The more general statement follows by a similar argument, using the fact that $\theta(A) = \theta(A \cap [0, \frac{1}{s}])$ for every set A. $\square$

Since we conjecturally have $||A_d^+| - \frac{y}{d}| \le \left(\frac{y}{d}\right)^{\frac{1}{2}+o(1)}$, it seems that the other complementary slackness condtions should be treated with some care. If we assume that some version of Pólya's conjecture is true on average, so that $|A_d^+| > \frac{y}{d}$ for most d having an even number of prime factors and $|A_d^+| < \frac{y}{d}$ for most d having an odd number of prime factors, then we might conjecture that

$$(-1)^{|S|}\lambda(S) \ge 0 \tag{A}$$

for optimal upper bound sieves which also have small error terms. It turns out that the Selberg upper bound sieve (λ^S, θ^S) does *not* satisfy condition (A): taking S to be a set consisting of 9 copies of $\frac{1}{12}$, we get

$$(-1)^9\lambda^S\left(\left\{9 \cdot \frac{1}{12}\right\}\right) = 2\binom{9}{4}\left(1 - 2\cdot\frac{4}{12}\right)\left(1 - 2\cdot\frac{5}{12}\right) - 9\binom{8}{4}\left(1 - 2\cdot\frac{5}{12}\right)^2 = -\frac{7}{2} < 0.$$

On the other hand, the Selberg upper bound sieve does not have a very good error term in comparison with the β-sieve, which does satisfy (A). Additionally, the Selberg upper bound sieve satisfies (A) for sets S with $\Sigma(S) \le \frac{1}{2}$.

Generally speaking, linear optimization problems tend to have unique solutions, corresponding to vertices of some associated polytope. When the solution is nonunique, then the problem is said to be degenerate - this corresponds to the polytope having a face which is contained in a level set of the linear function we are trying to optimize. In the case of the linear sieve (i.e. $\kappa = 1$), the problem turns out to be infinitely degenerate. From this point of view, the Selberg upper bound sieve method corresponds to restricting ourselves to some ellipsoid contained in our polytope. Since the Selberg upper bound sieve is actually optimal when $\kappa = 1$ and $s = 2$, this means it "should" correspond to some sort of interior point of the degenerate top face of our polytope. Thus if $\theta^S(S) = 0$ for sets S satisfying some simple property, then it seems likely that $\theta(S) = 0$ for any optimal upper bound sieve and any set S satisfying the same property.

Proposition 3. *For $|S| \ge 2$, $\min(S) < \frac{1}{s}$, $\Sigma(S) \le \frac{1}{2}$, we have $\theta^S(S) = \theta^\beta(S) = 0$.*

Proof. We have

$$\theta^S(S) = \left(\sum_{A \subseteq S} (-1)^{|A|}(1 - 2\Sigma(A))_+\right)^2,$$

and from $\Sigma(S) \le \frac{1}{2}$ we have $(1 - 2\Sigma(A))_+ = 1 - 2\Sigma(A)$ for $A \subseteq S$. If $S = \{x_1, ..., x_n\}$, then

$$\sum_{A \subseteq S} (-1)^{|A|}(1 - 2\Sigma(A)) = \sum_{A \subseteq S} (-1)^{|A|} - 2\sum_{i=1}^{n} x_i \sum_{x_i \in A \subseteq S} (-1)^{|A|} = 0.$$

Now we turn to $\theta^\beta(S)$. Supposing that $x_1 \ge \cdots \ge x_n$, we just need to show that for all $A \subseteq S \setminus \{x_n\}$ we have $\lambda^\beta(A) \ne 0 \iff \lambda^\beta(A \cup \{x_n\}) \ne 0$. The only case in which this is not obvious is when $|A|$ is even and $\Sigma(A) + x_n + 2x_n > 1$, and in this case we have

$$\Sigma(S) \ge nx_n \ge 2x_n > 1 - (\Sigma(A) + x_n) \ge 1 - \Sigma(S) \ge \frac{1}{2},$$

a contradiction. $\square$

Based on this, we conjecture that any optimal upper bound sieve has the property

$$(\tfrac{1}{2}) \qquad |S| \geq 2,\ \min(S) < \frac{1}{s},\ \Sigma(S) \leq \frac{1}{2} \implies \theta(S) = 0.$$

If we assume ($\frac{1}{2}$), we get the nice formula

$$\max(S) < \frac{1}{s},\ \Sigma(S) \leq \frac{1}{2} \implies \lambda(S) = (-1)^{|S|}\big(1 - \sum_{x \in S} \theta(x)\big)$$

which determines many of the sieve weights in terms of the sieve weights attached to singletons, so it seems that the most important thing to focus on is the function $\theta(x)$. By Proposition 2, we have $\theta(x) = 0$ for $0 \leq x < \min(\frac{1}{s}, 1 - \frac{2}{s})$, and since the sifting limit is s we have $\theta(x) = 1$ for $x > \frac{1}{s}$. The Selberg upper bound sieve has

$$\theta^S(x) = \begin{cases} 4x^2 & \text{if } 0 \leq x \leq \frac{1}{2}, \\ 1 & \text{else,} \end{cases}$$

while the β-sieve has

$$\theta^\beta(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq \min(\frac{1}{s}, \frac{1}{3}), \\ 1 & \text{else.} \end{cases}$$

The following conjecture is natural, if unjustified:

$$(>) \qquad x \geq y \implies \theta(x) \geq \theta(y).$$

Now we consider the support of λ. In the case of the Selberg upper bound sieve, we clearly have

$$\lambda^S(S) \neq 0 \implies \exists A \subseteq S,\ \Sigma(A) \leq \frac{1}{2},\ \Sigma(S \setminus A) \leq \frac{1}{2}.$$

The β-sieve has a more interesting constraint on its support.

Definition 1. A set S is *flexible* if for every $0 \leq x \leq 1$ there exists $A \subseteq S$ such that $\Sigma(A) \leq x$ and $\Sigma(S \setminus A) \leq 1 - x$.

Proposition 4 (from section 12.7 of [1])**.** *If for every $A \subseteq S$ we have $\Sigma(A) + \min(A) \leq 1$, then S is flexible. In particular, if $\lambda^\beta(S) \neq 0$ then S is flexible.*

Proof. Set $u = \min(S)$, $S' = S \setminus \{u\}$. By induction on $|S|$, we see that S' is flexible. Let $0 \leq x \leq 1$, and suppose that $A' \subseteq S'$ satisfies $\Sigma(A') \leq x$, $\Sigma(S' \setminus A') \leq 1 - x$. Since

$$\Sigma(A' \cup \{u\}) + \Sigma((S' \setminus A') \cup \{u\}) = \Sigma(S) + \min(S) \leq 1$$

by assumption, we must have one of $\Sigma(A' \cup \{u\}) \leq x$, $\Sigma((S' \setminus A') \cup \{u\}) \leq 1 - x$, so at least one of the choices $A = A'$ or $A = A' \cup \{u\}$ satisfies $\Sigma(A) \leq x$, $\Sigma(S \setminus A) \leq 1 - x$.

Now suppose that $\lambda^\beta(S) \neq 0$, so that for any $A \subseteq S$ with $|A|$ odd we have $\Sigma(A) + 2\min(A) \leq 1$. Then for any $A \subseteq S$ with $|A|$ even, if $A' = A \setminus \{\min(A)\}$ then $|A'|$ is odd, so

$$\Sigma(A) + \min(A) = \Sigma(A') + 2\min(A) \leq \Sigma(A') + 2\min(A') \leq 1. \qquad \square$$

It seems that as s decreases from ∞ to 2, the supports of optimal sieves get gradually less flexible, although it isn't clear what the correct weakening of flexibility should be. The following conjecture seems plausible:

$$(\mathrm{F}) \qquad S \subseteq \big[1 - \frac{2}{s}, \frac{1}{s}\big],\ \lambda(S) \neq 0 \implies \Sigma(S) \leq \frac{2}{s}.$$

3. Upper bound iteration rules

Here by an iteration rule we mean a special type of sieve, used to get new bounds on $\mathcal{S}(A,z)$ given upper and lower bounds on $\mathcal{S}(A_d,w)$ for squarefree numbers d having all their prime factors between w and z. Supposing $y = z^s = w^t$, if $(\lambda^{it},\theta^{it})$ is an upper bound sieve such that every set in the support of λ^{it} is contained in $[\frac{1}{t},\frac{1}{s}]$, then the corresponding iteration rule is given by

$$\mathcal{S}(A,z) \le \sum_{\substack{d \text{ squarefree} \\ p|d \implies w\le p<z}} \lambda^{it}(S_d)\mathcal{S}(A_d,w),$$

where $S_d = \{\frac{\log(p)}{\log(y)} \text{ s.t. } p \mid d\}$. This leads to an iterative inequality on $F_\kappa(s)$ in terms of F_κ, f_κ in an obvious way. The main advantage of using iteration rules is that it is typically very easy to check that $(\lambda^{it},\theta^{it})$ is a valid upper bound sieve. Our main concern is with iteration rules which are *optimal* when $\kappa = 1$, i.e. such that the pair of functions F, f is a fixed point of the iteration rule.

Theorem 1. *Suppose that the upper bound sieve $(\lambda^{it},\theta^{it})$ has λ^{it} supported on sets contained in $[1-\frac{2}{s},\frac{1}{s}]$, and satisfies the conditions* (O), (A), (F) *for all sets $S \subseteq [1-\frac{2}{s},\frac{1}{s}]$. Then the corresponding iteration rule is optimal in dimension $\kappa = 1$.*

Proof. Set $t = \frac{1}{1-\frac{2}{s}}$, $w = y^{\frac{1}{t}}$. By condition (A), the iteration rule is given to first order by

$$F(s)\frac{y}{e^\gamma \log(z)} \le \sum_{\substack{\mu(d)=1 \\ p|d \implies w\le p<z}} \lambda^{it}(S_d)F(t-t\Sigma(S_d))\frac{y/d}{e^\gamma\log(w)} + \sum_{\substack{\mu(d)=-1 \\ p|d \implies w\le p<z}} \lambda^{it}(S_d)f(t-t\Sigma(S_d))\frac{y/d}{e^\gamma\log(w)}.$$

The main idea is to exploit the fact that $\mathcal{S}(A^+,z) = F(s)\frac{y}{e^\gamma\log(z)} + O\left(\frac{y}{\log(z)^2}\right)$ and $\mathcal{S}(A^-,z) = f(s)\frac{y}{e^\gamma\log(z)} + O\left(\frac{y}{\log(z)^2}\right)$ for all $s > 1$. Since $\lambda^{it}(S_d) \neq 0$ implies $t - t\Sigma(S_d) \ge t - t\cdot\frac{2}{s} = 1$ by condition (F), we just need to check that

$$\mathcal{S}(A^+,z) = \sum_{\substack{\mu(d)=1 \\ p|d \implies w\le p<z}} \lambda^{it}(S_d)\mathcal{S}(A_d^+,w) + \sum_{\substack{\mu(d)=-1 \\ p|d \implies w\le p<z}} \lambda^{it}(S_d)\mathcal{S}(A_d^-,w) + O\left(\frac{y}{\log(z)^2}\right).$$

Since nonsquarefree numbers don't have a large contribution to either side, and since A_d^+ is supported on numbers with an odd number of prime factors while A_d^- is supported on numbers with an even number of prime factors, this follows from condition (O). $\square$

We can describe the sieve weights produced by an iteration rule as follows. For every u, let (λ_u^+,θ_u^+) be an upper bound sieve with sifting limit u and let (λ_u^-,θ_u^-) be a lower bound sieve with sifting limit u. Let $(\lambda^{it},\theta^{it})$ be our iteration rule sieve, with λ^{it} supported on sets contained in $[\frac{1}{t},\frac{1}{s}]$. Then the resulting upper bound sieve (λ,θ) is given by

$$\lambda(S) = \begin{cases} \lambda^{it}(S\cap[\frac{1}{t},\frac{1}{s}])\lambda^+_{t-t\Sigma(S\cap[\frac{1}{t},\frac{1}{s}])}(S\setminus[\frac{1}{t},\frac{1}{s}]) & \text{if } \lambda^{it}(S\cap[\frac{1}{t},\frac{1}{s}]) \ge 0, \\ \lambda^{it}(S\cap[\frac{1}{t},\frac{1}{s}])\lambda^-_{t-t\Sigma(S\cap[\frac{1}{t},\frac{1}{s}])}(S\setminus[\frac{1}{t},\frac{1}{s}]) & \text{if } \lambda^{it}(S\cap[\frac{1}{t},\frac{1}{s}]) \le 0. \end{cases}$$

In particular, for a singleton set we have

$$\theta(x) = \begin{cases} \theta_t^+(x) & \text{if } 0 \le x < \frac{1}{t}, \\ \theta^{it}(x) & \text{if } \frac{1}{t} \le x \le \frac{1}{s}, \\ 1 & \text{else.} \end{cases}$$

4. The range $\frac{5}{2} \le s \le 3$

We will assume throughout that we are working with an optimal upper bound sieve (λ, θ) with sifting limit $\frac{5}{2} \le s \le 3$ satisfying conditions (O), (A), (F), and trying to maximize the quantity $a_1 = \int_0^{\frac{1}{2}} \theta(x)\frac{dx}{x}$ subject to these constraints. By Theorem 1, we only need to consider the constraints involving sets S contained in $[1-\frac{2}{s}, \frac{1}{s}]$.

By condition (F), if $S \subseteq [1-\frac{2}{s}, \frac{1}{s}]$ and $\lambda(S) \neq 0$, then $|S| < 4$ since $4(1-\frac{2}{s}) \ge \frac{2}{s}$ for $s \ge \frac{5}{2}$. By condition (A) we have $\lambda(S) \le 0$ if $|S| = 3$, so if for some $x, y, z \in [1-\frac{2}{s}, \frac{1}{s}]$ we had $\lambda(x, y, z) < 0$, then we would have

$$\theta(\{k\cdot x, k\cdot y, k\cdot z\}) = \sum_{0\le a,b,c\le k} \binom{k}{a}\binom{k}{b}\binom{k}{c}\lambda(\{a\cdot x, b\cdot y, c\cdot z\}) \le k^3\lambda(x,y,z) + O(k^2) < 0$$

for k sufficiently large, a contradiction. Thus $\lambda(x, y, z) = 0$ for $x, y, z \in [1-\frac{2}{s}, \frac{1}{s}]$.

Note that for any $x, y, z \in [1-\frac{2}{s}, \frac{1}{s}]$, $\lambda(x, y, z) = 0$ implies that

$$\theta(x, y, z) = \theta(x, y) + \theta(x, z) + \theta(y, z) - \theta(x) - \theta(y) - \theta(z) + 1.$$

Applying Proposition 2 (which used condition (O)) to the set $\{x, y\}$ of size 2, we find

$$x + y \le 1 - \frac{1}{s} \implies \theta(x, y) = 0.$$

Using condition (O) directly, we also have

$$x + y + z = 1,\ x, y, z \le \frac{1}{s} \implies \theta(x) + \theta(y) + \theta(z) = \theta(x, y) + \theta(x, z) + \theta(y, z) + 1.$$

It's convenient to replace the interval $[1-\frac{2}{s}, \frac{1}{s}]$ by the interval $[0, 1]$. Let $r_s(x) = 1 - \frac{2}{s} + (\frac{3}{s} - 1)x$. Let $f(x) = \theta(r_s(x))$, and let $g(x, y) = \theta(r_s(x), r_s(y))$. Note that $r_s(\frac{2}{3}) = \frac{1}{3}$, so if $x + y + z = 2$ then $r_s(x) + r_s(y) + r_s(z) = 1$.

Theorem 2. *Suppose $f : [0, 1] \to \mathbb{R}_{\ge 0}$ and $g : [0, 1]^2 \to \mathbb{R}_{\ge 0}$ are nonnegative functions such that*

$$x + y \le 1 \implies g(x, y) = 0,$$

$$\forall x, y, z \in [0, 1] \quad f(x) + f(y) + f(z) \le g(x, y) + g(x, z) + g(y, z) + 1,$$

and

$$x + y + z = 2 \implies f(x) + f(y) + f(z) = g(x, y) + g(x, z) + g(y, z) + 1.$$

Then we have

a) *f is nondecreasing,*

b) *for every integer $n > 1$,*

$$\frac{f(\frac{1}{n}) + \cdots + f(\frac{n-1}{n})}{n-1} \le \frac{1}{3} \le \frac{f(\frac{0}{n}) + \cdots + f(\frac{n}{n})}{n+1},$$

c) *f is integrable and $\int_0^1 f(x)dx = \frac{1}{3}$,*

d) *g is nondecreasing in either argument, and moreover satisfies the inequality*

$$w \le x, y \le z \implies g(x, z) - g(w, z) \ge g(x, y) - g(w, y),$$

e) *if f, g are continuous then they come from a symmetric probability distribution μ supported on the simplex $\{a, b, c \in [0, 1]^3 \mid a + b + c = 2\}$, according to the formulae*

$$f(x) = \mathbb{P}_{\mu(a,b,c)}[a \le x],\ g(x, y) = \mathbb{P}_{\mu(a,b,c)}[a \le x \wedge b \le y]?$$

Proof. Part a): suppose $0 \le a < b \le 1$, we will show that $f(a) \le f(b)$. Choose a nonnegative integer k such that

$$2a - b < k(b-a) < b.$$

For each $0 \le i \le k$, set

$$x_{2i} = 1 - \frac{b + (k-2i)(b-a)}{2}, \; x_{2i+1} = 1 - \frac{b + (2i-k)(b-a)}{2}.$$

Note that by the choice of k we have $a + x_0 = a + x_{2k+1} < 1$ and $1 - b < x_i < 1$ for all $0 \le i \le 2k+1$. Furthermore, for each i we have $b + x_{2i} + x_{2i+1} = 2$ and $a + x_{2i-1} + x_{2i} = 2$. Thus, for each $0 \le i \le k$ we have

$$\begin{aligned} f(b) + f(x_{2i}) + f(x_{2i+1}) &= g(b, x_{2i}) + g(b, x_{2i+1}) + g(x_{2i}, x_{2i+1}) + 1, \\ f(a) + f(x_{2i}) + f(x_{2i+1}) &\le g(a, x_{2i}) + g(a, x_{2i+1}) + g(x_{2i}, x_{2i+1}) + 1, \end{aligned}$$

and for each $1 \le i \le k$ we have

$$\begin{aligned} f(b) + f(x_{2i-1}) + f(x_{2i}) &\le g(b, x_{2i-1}) + g(b, x_{2i}) + g(x_{2i-1}, x_{2i}) + 1, \\ f(a) + f(x_{2i-1}) + f(x_{2i}) &= g(a, x_{2i-1}) + g(a, x_{2i}) + g(x_{2i-1}, x_{2i}) + 1. \end{aligned}$$

Adding together the inequalities and subtracting the equalities, we get

$$\begin{aligned} f(a) &\le f(b) + g(a, x_0) + g(a, x_{2k+1}) - g(b, x_0) - g(b, x_{2k+1}) \\ &= f(b) - g(b, x_0) - g(b, x_{2k+1}) \le f(b). \end{aligned}$$

Part b): first we prove the left hand inequality. For every ordered triple of integers $0 < i, j, k < n$ satisfying $i + j + k = 2n$, we have an equality

$$f(\tfrac{i}{n}) + f(\tfrac{j}{n}) + f(\tfrac{k}{n}) = g(\tfrac{i}{n}, \tfrac{j}{n}) + g(\tfrac{i}{n}, \tfrac{k}{n}) + g(\tfrac{j}{n}, \tfrac{k}{n}) + 1.$$

Also, for every ordered triple $0 < i, j, k < n$ satisfying $i + j + k = 2n - 1$, we have the inequality

$$f(\tfrac{i}{n}) + f(\tfrac{j}{n}) + f(\tfrac{k}{n}) \le g(\tfrac{i}{n}, \tfrac{j}{n}) + g(\tfrac{i}{n}, \tfrac{k}{n}) + g(\tfrac{j}{n}, \tfrac{k}{n}) + 1.$$

Adding the inequalities and subtracting the equalities, and using $g(\frac{i}{n}, \frac{j}{n}) = 0$ when $i + j = n$, gives the left hand inequality of b). For the right hand inequality of b) one uses equalities corresponding to triples $0 \le i, j, k \le n$ with $i + j + k = 2n$, and inequalities corresponding to triples $0 \le i, j, k \le n$ with $i + j + k = 2n + 1$.

Part c) follows immediately from parts a) and b).

First we prove part d) in the case $x - w = z - y$. If $x + y = w + z \le 1$, it is immediate. If $x + y = w + z \ge 1$, it follows from

$$\begin{aligned} f(x) + f(y) + f(2-x-y) &= g(x, y) + g(x, 2-x-y) + g(y, 2-x-y) + 1, \\ f(w) + f(z) + f(2-x-y) &= g(w, z) + g(w, 2-x-y) + g(z, 2-x-y) + 1, \\ f(w) + f(y) + f(2-x-y) &\le g(w, y) + g(w, 2-x-y) + g(y, 2-x-y) + 1, \\ f(x) + f(z) + f(2-x-y) &\le g(x, z) + g(x, 2-x-y) + g(z, 2-x-y) + 1. \end{aligned}$$

Applying this repeatedly, we see that the inequality in part d) holds whenever $x - w$ is a rational multiple of $z - y$. TODO

Also: if $a \le b$, then

$$\begin{aligned} f(a) + f(1-b) + f(1) &\le g(a, 1) + g(1-b, 1) + 1, \\ f(b) + f(1-b) + f(1) &= g(b, 1) + g(1-b, 1) + 1, \end{aligned}$$

so

$$f(b) - f(a) \ge g(b, 1) - g(a, 1).$$

Additionally:

$$f(0) + f(0) + f(1) \leq 1.$$

I don't know how to prove part e) yet, but numerical computation seems to confirm it. $\square$

Thus $\theta(x)$ is increasing, and the average value of $\theta(x)$ on the interval $[1 - \frac{2}{s}, \frac{1}{s}]$ is $\frac{1}{3}$. Since $\frac{1}{x}$ is decreasing, in order to maximize $\int_{1-\frac{2}{s}}^{\frac{1}{s}} \theta(x) \frac{dx}{x}$ we must take $\theta(x) = \frac{1}{3}$ identically on this interval. In terms of λ, this is corresponds to taking $\lambda(x) = -\frac{2}{3}$, $\lambda(x, y) = \frac{1}{3}$ for all $x, y \in [1 - \frac{2}{s}, \frac{1}{s}]$, which we can easily check gives an optimal upper bound sieve iteration. For $s = \frac{5}{2}$ the resulting sieve has

$$a_1 = \int_0^{\frac{1}{2}} \theta(x) \frac{dx}{x} = \frac{\log(2)}{3} + \log\left(\frac{5}{4}\right) \approx 0.454.$$

References

[1] John Friedlander and Henryk Iwaniec. *Opera de cribro*, volume 57 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2010.

[2] Atle Selberg. *Collected papers. Vol. II*. Springer-Verlag, Berlin, 1991. With a foreword by K. Chandrasekharan.

Department of Mathematics, Stanford University, 450 Serra Mall, Bldg. 380, Stanford, CA 94305-2125

E-mail address: notzeb@stanford.edu