Asymptotics of a model problem from sieve theory

Zarathustra Brady
Sieve theory

- Take an interval \( A \) of consecutive whole numbers, such as \([5, 9] = \{5, 6, 7, 8, 9\}\).
Sieve theory

- Take an interval $A$ of consecutive whole numbers, such as $[5, 9] = \{5, 6, 7, 8, 9\}$.

- Remove the multiples of some collection of primes $\mathcal{P}$ from this interval. Call the set that remains $S(A, \mathcal{P})$. 
Sieve theory

- Take an interval $A$ of consecutive whole numbers, such as $[5, 9] = \{5, 6, 7, 8, 9\}$.

- Remove the multiples of some collection of primes $\mathcal{P}$ from this interval. Call the set that remains $S(A, \mathcal{P})$.

- For instance, if $\mathcal{P} = \{2, 3\}$, then $S([5, 9], \{2, 3\}) = \{5, 7\}$.
Sieve theory

- Take an interval $A$ of consecutive whole numbers, such as $[5, 9] = \{5, 6, 7, 8, 9\}$.

- Remove the multiples of some collection of primes $\mathcal{P}$ from this interval. Call the set that remains $S(A, \mathcal{P})$.

- For instance, if $\mathcal{P} = \{2, 3\}$, then $S([5, 9], \{2, 3\}) = \{5, 7\}$.

- The big question:
  
  What can we say about $|S(A, \mathcal{P})|$?
Sieve theory

- Take an interval $A$ of consecutive whole numbers, such as $[5, 9] = \{5, 6, 7, 8, 9\}$.

- Remove the multiples of some collection of primes $\mathcal{P}$ from this interval. Call the set that remains $S(A, \mathcal{P})$.

- For instance, if $\mathcal{P} = \{2, 3\}$, then $S([5, 9], \{2, 3\}) = \{5, 7\}$.

- The big question:
  
  What can we say about $|S(A, \mathcal{P})|$?

- Pretend that we know $\mathcal{P}$, and that we know the length of $A$, but we don’t know the endpoints of $A$. 

Probabilistic version

- Suppose we pick a uniformly random number $n$ from the interval $A$. 

Although we don't know exactly what $A$ is, we do know that

\[ \frac{1}{2} - \frac{1}{|A|} \leq P[2 \text{ divides } n] \leq \frac{1}{2} + \frac{1}{|A|}, \]

\[ \frac{1}{3} - \frac{1}{|A|} \leq P[3 \text{ divides } n] \leq \frac{1}{3} + \frac{1}{|A|}, \]

\[ \frac{1}{6} - \frac{1}{|A|} \leq P[6 \text{ divides } n] \leq \frac{1}{6} + \frac{1}{|A|}. \]

So we can say that

\[ P[n \in S(A, \{2, 3\})] \geq 1 - \left( \frac{1}{2} + \frac{1}{|A|} \right) - \left( \frac{1}{3} + \frac{1}{|A|} \right) + \left( \frac{1}{6} - \frac{1}{|A|} \right). \]
Suppose we pick a uniformly random number $n$ from the interval $A$.

Although we don’t know exactly what $A$ is, we do know that

$$\frac{1}{2} - \frac{1}{|A|} \leq \mathbb{P}[2 \text{ divides } n] \leq \frac{1}{2} + \frac{1}{|A|}.$$
Suppose we pick a uniformly random number $n$ from the interval $A$.

Although we don’t know exactly what $A$ is, we do know that

$$\frac{1}{2} - \frac{1}{|A|} \leq \mathbb{P}[2 \text{ divides } n] \leq \frac{1}{2} + \frac{1}{|A|}.$$  

We also know that

$$\frac{1}{3} - \frac{1}{|A|} \leq \mathbb{P}[3 \text{ divides } n] \leq \frac{1}{3} + \frac{1}{|A|},$$

$$1 - \frac{1}{|A|} \leq \mathbb{P}[6 \text{ divides } n] \leq 1 + \frac{1}{|A|}.$$
Probabilistic version

▶ Suppose we pick a uniformly random number \( n \) from the interval \( A \).

▶ Although we don’t know exactly what \( A \) is, we do know that

\[
\frac{1}{2} - \frac{1}{|A|} \leq \mathbb{P}[2 \text{ divides } n] \leq \frac{1}{2} + \frac{1}{|A|}.
\]

▶ We also know that

\[
\frac{1}{3} - \frac{1}{|A|} \leq \mathbb{P}[3 \text{ divides } n] \leq \frac{1}{3} + \frac{1}{|A|},
\]
\[
\frac{1}{6} - \frac{1}{|A|} \leq \mathbb{P}[6 \text{ divides } n] \leq \frac{1}{6} + \frac{1}{|A|}.
\]

▶ So we can say that

\[
\mathbb{P}[n \in S(A, \{2, 3\})] \geq 1 - \left( \frac{1}{2} + \frac{1}{|A|} \right) - \left( \frac{1}{3} + \frac{1}{|A|} \right) + \left( \frac{1}{6} - \frac{1}{|A|} \right).
\]
The naïve approaches don’t work

- If we ignore the $1/|A|$ error terms, we can use P.I.E. to predict

\[
P[n \in S(A, \mathcal{P})] \overset{?}{=} \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right).
\]
The naïve approaches don’t work

- If we ignore the $1/|A|$ error terms, we can use P.I.E. to predict

$$\mathbb{P}[n \in S(A, \mathcal{P})] \approx \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right).$$

- This guess is completely wrong!

- Take $A = [1, N]$ and take $\sqrt{N}$ to be the set of primes below $\sqrt{N}$.

- The guess above predicts that $\mathbb{P}[n \in S([1, N], \sqrt{N})] \approx \prod_{p < \sqrt{N}} \left(1 - \frac{1}{p}\right) \approx e^{-\gamma} \log(\sqrt{N})$.

- But the true value is $\mathbb{P}[n \in S([1, N], \sqrt{N})] \approx \frac{1}{\log(N)}$. 
The naïve approaches don’t work

- If we ignore the $1/|A|$ error terms, we can use P.I.E. to predict

$$P[n \in S(A, \mathcal{P})] \approx \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right).$$

- This guess is completely wrong!
- Take $A = [1, N]$ and take $\mathcal{P}_{\sqrt{N}}$ to be the set of primes below $\sqrt{N}$. 

But the true value is

$$P[n \in S([1, N], \mathcal{P}_{\sqrt{N}})] \approx \log(N).$$
The naïve approaches don’t work

- If we ignore the $1/|A|$ error terms, we can use P.I.E. to predict

$$\mathbb{P}[n \in S(A, \mathcal{P})] \approx \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right).$$

- This guess is completely wrong!

- Take $A = [1, N]$ and take $\mathcal{P}_{\sqrt{N}}$ to be the set of primes below $\sqrt{N}$.

- The guess above predicts that

$$\mathbb{P}[n \in S([1, N], \mathcal{P}_{\sqrt{N}})] \approx \prod_{p < \sqrt{N}} \left(1 - \frac{1}{p}\right) \approx \frac{e^{-\gamma}}{\log(\sqrt{N})}.$$
The naïve approaches don’t work

- If we ignore the $1/|A|$ error terms, we can use P.I.E. to predict

$$
\mathbb{P}[n \in S(A, \mathcal{P})] \approx \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right).
$$

- This guess is completely wrong!

- Take $A = [1, N]$ and take $\mathcal{P}_{\sqrt{N}}$ to be the set of primes below $\sqrt{N}$.

- The guess above predicts that

$$
\mathbb{P}[n \in S([1, N], \mathcal{P}_{\sqrt{N}})] \approx \prod_{p < \sqrt{N}} \left(1 - \frac{1}{p}\right) \approx \frac{e^{-\gamma}}{\log(\sqrt{N})}.
$$

- But the true value is

$$
\mathbb{P}[n \in S([1, N], \mathcal{P}_{\sqrt{N}})] \approx \frac{1}{\log(N)}.
$$
The naïve approaches don’t work

▶ So we can’t ignore the error terms.
The naïve approaches don’t work

- So we can’t ignore the error terms.

- Let’s be really conservative this time, and try the union bound:

\[
\mathbb{P}[n \in S(A, \mathcal{P})] \geq 1 - \sum_{p \in \mathcal{P}} \left( \frac{1}{p} + \frac{1}{|A|} \right).
\]
The na"ïve approaches don’t work

▶ So we can’t ignore the error terms.

▶ Let’s be really conservative this time, and try the union bound:

\[
P[n \in S(A, \mathcal{P})] \geq 1 - \sum_{p \in \mathcal{P}} \left( \frac{1}{p} + \frac{1}{|A|} \right).
\]

▶ Now the error terms are under control, and at first this seems to be working well...
The naïve approaches don’t work

- So we can’t ignore the error terms.

- Let’s be really conservative this time, and try the union bound:

\[
P[n \in S(A, \mathcal{P})] \geq 1 - \sum_{p \in \mathcal{P}} \left( \frac{1}{p} + \frac{1}{|A|} \right).
\]

- Now the error terms are under control, and at first this seems to be working well...

- The problem is that

\[
\sum_{p \leq N} \frac{1}{p} \approx \log(\log(N))
\]

diverges. This kills most simple variants of the above idea.
Bucketing approach

Since $\sum p \frac{1}{p}$ diverges, a good strategy is to put primes in \textit{buckets}:

$$\mathcal{P} = \mathcal{P}_1 \sqcup \mathcal{P}_2 \sqcup \cdots \sqcup \mathcal{P}_k.$$
Bucketing approach

- Since $\sum_p \frac{1}{p}$ diverges, a good strategy is to put primes in buckets:

$$\mathcal{P} = \mathcal{P}_1 \sqcup \mathcal{P}_2 \sqcup \cdots \sqcup \mathcal{P}_k.$$ 

- We choose our buckets so that each sum

$$\sum_{p \in \mathcal{P}_i} \frac{1}{p}$$

is of size $\asymp 1$.
Bucketing approach

- Since $\sum p \frac{1}{p}$ diverges, a good strategy is to put primes in buckets:
  $$\mathcal{P} = \mathcal{P}_1 \sqcup \mathcal{P}_2 \sqcup \cdots \sqcup \mathcal{P}_k.$$  

- We choose our buckets so that each sum
  $$\sum_{p \in \mathcal{P}_i} \frac{1}{p}$$
  is of size $\asymp 1$.

- This corresponds to taking buckets of the form
  $$\mathcal{P}_i = \mathcal{P} \cap [|A|^{1/s}, |A|^{1/t}].$$
Bucketing approach

- Since $\sum p \frac{1}{p}$ diverges, a good strategy is to put primes in buckets:
  \[ \mathcal{P} = \mathcal{P}_1 \sqcup \mathcal{P}_2 \sqcup \cdots \sqcup \mathcal{P}_k. \]

- We choose our buckets so that each sum
  \[ \sum_{p \in \mathcal{P}_i} \frac{1}{p} \]
  is of size $\asymp 1$.

- This corresponds to taking buckets of the form
  \[ \mathcal{P}_i = \mathcal{P} \cap [\frac{1}{|A|^{1/s}}, \frac{1}{|A|^{1/t}}]. \]

- Buckets corresponding to smaller primes $\rightarrow$ smaller error terms $\rightarrow$ naïve P.I.E. guess is a better approximation.
The model problem

- Most of the asymptotic error comes from the bucket containing the largest primes.
The model problem

- Most of the asymptotic error comes from the bucket containing the largest primes.

- The model problem asks: what if that was the only bucket?
The model problem

- Most of the asymptotic error comes from the bucket containing the largest primes.

- The model problem asks: what if that was the only bucket?

- Suppose we have

\[ p \in \mathcal{P} \implies p \in [\left| A \right|^{1/(k+1)}, \left| A \right|^{1/k}] \]
The model problem

- Most of the asymptotic error comes from the bucket containing the largest primes.

- The model problem asks: what if that was the only bucket?

- Suppose we have

\[ p \in \mathcal{P} \implies p \in [\|A\|^{1/(k+1)}, \|A\|^{1/k}]. \]

- Then for any \( p_1, \ldots, p_k \in \mathcal{P} \), we know that

\[ \mathbb{P}[p_1 \cdots p_k \text{ divides } n] = \frac{1}{p_1 \cdots p_k} + O\left(\frac{1}{\|A\|}\right). \]
The model problem

- Most of the asymptotic error comes from the bucket containing the largest primes.

- The model problem asks: what if that was the only bucket?

- Suppose we have

\[ p \in \mathcal{P} \implies p \in [|A|^{1/(k+1)}, |A|^{1/k}] \]

- Then for any \( p_1, \ldots, p_k \in \mathcal{P} \), we know that

\[ \mathbb{P}[p_1 \cdots p_k \text{ divides } n] = \frac{1}{p_1 \cdots p_k} + O\left(\frac{1}{|A|}\right) \]

- So the primes in \( \mathcal{P} \) are uncorrelated when considered at most \( k \) at a time.
Simplification of model problem

- Since the primes all have roughly the same size, we treat them as interchangeable.

\[ X = \# \{ p \in P \text{ such that } p \text{ divides } n \} \]

The expected size of \( X \) is

\[ E[X] \approx \sum_{p \in P} \frac{1}{p} \]

The second moment of \( X \) is given by

\[ E[X^2] \approx \sum_{p < q \in P} \frac{1}{pq} \approx \frac{1}{2} \left( \sum_{p \in P} \frac{1}{p} \right)^2 \]
Simplification of model problem

- Since the primes all have roughly the same size, we treat them as interchangeable.

- Define a random variable $X$ by

$$X = \# \{ p \in \mathcal{P} \text{ such that } p \text{ divides } n \}.$$
Simplification of model problem

- Since the primes all have roughly the same size, we treat them as interchangeable.

- Define a random variable $X$ by

$$X = \# \{ p \in \mathcal{P} \text{ such that } p \text{ divides } n \}.$$

- The expected size of $X$ is

$$\mathbb{E}[X] \approx \sum_{p \in \mathcal{P}} \frac{1}{p}.$$
Simplification of model problem

Since the primes all have roughly the same size, we treat them as interchangeable.

Define a random variable $X$ by

$$X = \#\{p \in \mathcal{P} \text{ such that } p \text{ divides } n\}.$$

The expected size of $X$ is

$$\mathbb{E}[X] \approx \sum_{p \in \mathcal{P}} \frac{1}{p}.$$

The second moment of $X$ is given by

$$\mathbb{E}\left[\left(\frac{X}{2}\right)^2\right] \approx \sum_{p < q \in \mathcal{P}} \frac{1}{pq} \approx \frac{1}{2} \left(\sum_{p \in \mathcal{P}} \frac{1}{p}\right)^2.$$
A Poisson imitator appears

- For each \( i \leq k \), we see that

\[
\mathbb{E} \left[ \binom{X}{i} \right] \approx \frac{\mathbb{E}[X]^i}{i!}.
\]
A Poisson imitator appears

- For each $i \leq k$, we see that

$$\mathbb{E}\left[\binom{X}{i}\right] \approx \frac{\mathbb{E}[X]^i}{i!}.$$ 

- These are exactly the first $k$ moments of a Poisson distribution!
A Poisson imitator appears

- For each \( i \leq k \), we see that

\[ \mathbb{E}\left[\binom{X}{i}\right] \approx \frac{\mathbb{E}[X]^i}{i!} . \]

- These are exactly the first \( k \) moments of a Poisson distribution!

- (We have no idea about the higher moments of \( X \).)
A Poisson imitator appears

- For each $i \leq k$, we see that
  $$\mathbb{E}\left[ \binom{X}{i} \right] \approx \frac{\mathbb{E}[X]^i}{i!}.$$  

- These are exactly the first $k$ moments of a Poisson distribution!

- (We have no idea about the higher moments of $X$.)

- We want to estimate
  $$\mathbb{P}[n \in S(A, \mathcal{P})] = \mathbb{P}[X = 0].$$
Our problem

- Forget all the previous stuff.

We have a random variable $X \in \mathbb{N}$, a Poisson parameter $\nu \in \mathbb{R}^+$, and $k \in \mathbb{N}$, s.t.

$$i \leq k \Rightarrow E[(X_i)] = \nu i!.$$ 

What are the best bounds we can put on $P[X = 0]$?

For which $\nu$, $k$ can we prove that $P[X = 0] > 0$?
Our problem

- Forget all the previous stuff.

- We have a random variable $X \in \mathbb{N}$, a Poisson parameter $\nu \in \mathbb{R}^+$, and $k \in \mathbb{N}$, s.t.

  $$i \leq k \implies \mathbb{E} \left[ \binom{X}{i} \right] = \frac{\nu^i}{i!}.$$
Our problem

- Forget all the previous stuff.

- We have a random variable $X \in \mathbb{N}$, a Poisson parameter $\nu \in \mathbb{R}^+$, and $k \in \mathbb{N}$, s.t.

  $$i \leq k \implies \mathbb{E}[\binom{X}{i}] = \frac{\nu^i}{i!}.$$ 

- What are the best bounds we can put on $\mathbb{P}[X = 0]$?
Our problem

- Forget all the previous stuff.

- We have a random variable $X \in \mathbb{N}$, a Poisson parameter $\nu \in \mathbb{R}^+$, and $k \in \mathbb{N}$, s.t.

  $$i \leq k \implies \mathbb{E}\left[\binom{X}{i}\right] = \frac{\nu^i}{i!}.$$ 

- What are the best bounds we can put on $P[X = 0]$?

- For which $\nu, k$ can we prove that $P[X = 0] > 0$?
Markov’s inequality

▶ How do we use the moment information?

Consider a polynomial \( \theta(x) \) of degree \( k \):

\[
\theta(x) = \lambda_0 + \lambda_1 x + \lambda_2 x^2 + \cdots + \lambda_k x^k.
\]

Our moment information tells us that

\[
E[\theta(X)] = \lambda_0 + \lambda_1 \nu + \lambda_2 \nu^2 + \cdots + \lambda_k \nu^k.
\]

If \( \theta(x) \leq 0 \) for \( x \in \{1, 2, \ldots\} \), we get

\[
E[\theta(X)] \leq P[X = 0] \theta(0).
\]
Markov’s inequality

▶ How do we use the moment information?

▶ Consider a polynomial \( \theta(x) \) of degree \( k \):

\[
\theta(x) = \lambda_0 + \lambda_1 x + \lambda_2 \binom{x}{2} + \cdots + \lambda_k \binom{x}{k}.
\]
Markov’s inequality

- How do we use the moment information?

- Consider a polynomial $\theta(x)$ of degree $k$:

  $$\theta(x) = \lambda_0 + \lambda_1 x + \lambda_2 \binom{x}{2} + \cdots + \lambda_k \binom{x}{k}.$$ 

- Our moment information tells us that

  $$\mathbb{E}[\theta(X)] = \lambda_0 + \lambda_1 \nu + \lambda_2 \frac{\nu^2}{2} + \cdots + \lambda_k \frac{\nu^k}{k!}.$$
Markov’s inequality

- How do we use the moment information?

- Consider a polynomial $\theta(x)$ of degree $k$:

  $$
  \theta(x) = \lambda_0 + \lambda_1 x + \lambda_2 \binom{x}{2} + \cdots + \lambda_k \binom{x}{k}.
  $$

- Our moment information tells us that

  $$
  \mathbb{E}[\theta(X)] = \lambda_0 + \lambda_1 \nu + \lambda_2 \frac{\nu^2}{2} + \cdots + \lambda_k \frac{\nu^k}{k!}.
  $$

- If $\theta(x) \leq 0$ for $x \in \{1, 2, \ldots\}$, we get

  $$
  \mathbb{E}[\theta(X)] \leq \mathbb{P}[X = 0] \theta(0).
  $$
Convex optimization

- Our proof method is to write down a polynomial $\theta(x)$ such that:
  - $\theta$ has degree at most $k$,
  - $\theta(0) = 1$,
  - for all $x \in \mathbb{N}^+$, $\theta(x) \leq 0$,
  - and to conclude that $P[X = 0] \geq E[\theta(X)] = e^{-\nu} \sum_{n=0}^{\infty} \frac{\theta(n) \nu^n}{n!}$.

- Are there any better ways to prove a lower bound on $P[X = 0]$?

A general duality result in convex optimization says that the best lower bound using this strategy is equal to the least possible value of $P[X = 0]$. 
Convex optimization

- Our proof method is to write down a polynomial $\theta(x)$ such that:
  - $\theta$ has degree at most $k$, 
- Are there any better ways to prove a lower bound on $P[X = 0]$?
- A general duality result in convex optimization says that the best lower bound using this strategy is equal to the least possible value of $P[X = 0]$. 

Convex optimization

- Our proof method is to write down a polynomial $\theta(x)$ such that:
  - $\theta$ has degree at most $k$,
  - $\theta(0) = 1$,
  - for all $x \in \mathbb{N}^+$, $\theta(x) \leq 0$, and to conclude that $P[X = 0] \geq E[\theta(X)] = e^{-\nu} \sum_n \theta(n) \nu^n n!$.

- Are there any better ways to prove a lower bound on $P[X = 0]$?

- A general duality result in convex optimization says that the best lower bound using this strategy is equal to the least possible value of $P[X = 0]$. 
Our proof method is to write down a polynomial $\theta(x)$ such that:

- $\theta$ has degree at most $k$,
- $\theta(0) = 1$,
- for all $x \in \mathbb{N}^+$, $\theta(x) \leq 0$,
Convex optimization

- Our proof method is to write down a polynomial $\theta(x)$ such that:
  - $\theta$ has degree at most $k$,
  - $\theta(0) = 1$,
  - for all $x \in \mathbb{N}^+$, $\theta(x) \leq 0$,
- and to conclude that

$$
\mathbb{P}[X = 0] \geq \mathbb{E}[\theta(X)] = e^{-\nu} \sum_n \theta(n) \frac{\nu^n}{n!}.
$$
Our proof method is to write down a polynomial $\theta(x)$ such that:

- $\theta$ has degree at most $k$,
- $\theta(0) = 1$,
- for all $x \in \mathbb{N}^+$, $\theta(x) \leq 0$,

and to conclude that

$$\mathbb{P}[X = 0] \geq \mathbb{E}[\theta(X)] = e^{-\nu} \sum_n \theta(n) \frac{\nu^n}{n!}.$$ 

Are there any better ways to prove a lower bound on $\mathbb{P}[X = 0]$?
Our proof method is to write down a polynomial $\theta(x)$ such that:
- $\theta$ has degree at most $k$,
- $\theta(0) = 1$,
- for all $x \in \mathbb{N}^+$, $\theta(x) \leq 0$,

and to conclude that

$$\mathbb{P}[X = 0] \geq \mathbb{E}[\theta(X)] = e^{-\nu} \sum_n \theta(n) \frac{\nu^n}{n!}.$$ 

Are there any better ways to prove a lower bound on $\mathbb{P}[X = 0]$?

A general duality result in convex optimization says that the best lower bound using this strategy is equal to the least possible value of $\mathbb{P}[X = 0]$. 

---

Convex optimization
Optimizing our choice of $\theta$

- Selberg was able to compute the optimal choices of $\theta$ by hand for single digit values of the degree $k$. 

How?

To ensure that $\theta(x) \leq 0$ for $x \in \mathbb{N}^+$, we write $\theta$ in terms of its roots:

$$
\theta(x) = (1-x^{r_1})(1-x^{r_2}) \cdots (1-x^{r_k}).
$$

If there are any complex roots, replacing them with their real parts strictly improves our objective function. Removing negative roots also strictly improves our objective function.

Since coefficients of $\theta$ are linear in $1/r_i$, each $r_i$ may be taken to be a whole number.
Optimizing our choice of $\theta$

- Selberg was able to compute the optimal choices of $\theta$ by hand for single digit values of the degree $k$.
- How?
Optimizing our choice of $\theta$

- Selberg was able to compute the optimal choices of $\theta$ by hand for single digit values of the degree $k$.
- How?
- To ensure that $\theta(x) \leq 0$ for $x \in \mathbb{N}^+$, we write $\theta$ in terms of its roots:

$$\theta(x) = \left(1 - \frac{x}{r_1}\right) \left(1 - \frac{x}{r_2}\right) \cdots \left(1 - \frac{x}{r_k}\right).$$
Optimizing our choice of $\theta$

- Selberg was able to compute the optimal choices of $\theta$ by hand for single digit values of the degree $k$.

- How?

- To ensure that $\theta(x) \leq 0$ for $x \in \mathbb{N}^+$, we write $\theta$ in terms of its roots:

  $$\theta(x) = \left(1 - \frac{x}{r_1}\right) \left(1 - \frac{x}{r_2}\right) \cdots \left(1 - \frac{x}{r_k}\right).$$

- If there are any complex roots, replacing them with their real parts strictly improves our objective function.
Optimizing our choice of $\theta$

- Selberg was able to compute the optimal choices of $\theta$ by hand for single digit values of the degree $k$.

- How?

- To ensure that $\theta(x) \leq 0$ for $x \in \mathbb{N}^+$, we write $\theta$ in terms of its roots:

$$\theta(x) = \left(1 - \frac{x}{r_1}\right) \left(1 - \frac{x}{r_2}\right) \cdots \left(1 - \frac{x}{r_k}\right).$$

- If there are any complex roots, replacing them with their real parts strictly improves our objective function.

- Removing negative roots also strictly improves our objective function.
Optimizing our choice of $\theta$

- Selberg was able to compute the optimal choices of $\theta$ by hand for single digit values of the degree $k$.
- How?
- To ensure that $\theta(x) \leq 0$ for $x \in \mathbb{N}^+$, we write $\theta$ in terms of its roots:

$$
\theta(x) = \left(1 - \frac{x}{r_1}\right) \left(1 - \frac{x}{r_2}\right) \cdots \left(1 - \frac{x}{r_k}\right).
$$

- If there are any complex roots, replacing them with their real parts strictly improves our objective function.
- Removing negative roots also strictly improves our objective function.
- Since coefficients of $\theta$ are linear in $1/r_i$, each $r_i$ may be taken to be a whole number.
A simplex algorithm you can run by hand

- Our function \( \theta \) can now be completely described by listing out its (integer) roots.
A simplex algorithm you can run by hand

- Our function $\theta$ can now be completely described by listing out its (integer) roots.
- Such a $\theta$ satisfies our requirements if:

$$\text{Our objective function is } e^{-\nu} \sum_{n} \theta(n) \nu^{n} n! = \sum_{i} \lambda_{i} \nu^{i} i!.$$

- We can "pivot" our choice of $\theta$ by moving one of its roots, while keeping the other roots fixed.

Proposition

If no pivot increases the objective value, then $\theta$ is (globally) optimal.
A simplex algorithm you can run by hand

- Our function $\theta$ can now be completely described by listing out its (integer) roots.
- Such a $\theta$ satisfies our requirements if:
  - 1 is the least root of $\theta$, and
A simplex algorithm you can run by hand

- Our function \( \theta \) can now be completely described by listing out its (integer) roots.
- Such a \( \theta \) satisfies our requirements if:
  - 1 is the least root of \( \theta \), and
  - the remaining roots of \( \theta \) can be paired up so that each pair of roots are at most 1 apart.
A simplex algorithm you can run by hand

- Our function $\theta$ can now be completely described by listing out its (integer) roots.
- Such a $\theta$ satisfies our requirements if:
  - 1 is the least root of $\theta$, and
  - the remaining roots of $\theta$ can be paired up so that each pair of roots are at most 1 apart.
- Our objective function is $e^{-\nu} \sum_n \theta(n) \frac{\nu^n}{n!} = \sum_i \lambda_i \frac{\nu^i}{i!}$. 
A simplex algorithm you can run by hand

- Our function $\theta$ can now be completely described by listing out its (integer) roots.
- Such a $\theta$ satisfies our requirements if:
  - 1 is the least root of $\theta$, and
  - the remaining roots of $\theta$ can be paired up so that each pair of roots are at most 1 apart.
- Our objective function is $e^{-\nu} \sum_n \theta(n) \frac{\nu^n}{n!} = \sum_i \lambda_i \frac{\nu^i}{i!}$.
- We can “pivot” our choice of $\theta$ by moving one of its roots, while keeping the other roots fixed.
A simplex algorithm you can run by hand

- Our function $\theta$ can now be completely described by listing out its (integer) roots.
- Such a $\theta$ satisfies our requirements if:
  - 1 is the least root of $\theta$, and
  - the remaining roots of $\theta$ can be paired up so that each pair of roots are at most 1 apart.
- Our objective function is $e^{-\nu} \sum_n \theta(n) \frac{\nu^n}{n!} = \sum_i \lambda_i \frac{\nu^i}{i!}$.
- We can “pivot” our choice of $\theta$ by moving one of its roots, while keeping the other roots fixed.

**Proposition**

*If no pivot increases the objective value, then $\theta$ is (globally) optimal.*
...or by computer

<table>
<thead>
<tr>
<th>$k$</th>
<th>critical $\nu_k$</th>
<th>roots of the optimal $\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1, ${3, 4}$ or $1, {4, 5}$</td>
</tr>
<tr>
<td>5</td>
<td>3.11714</td>
<td>1, ${3, 4}, {7, 8}$</td>
</tr>
<tr>
<td>7</td>
<td>4.14377</td>
<td>1, ${3, 4}, {6, 7}, {11, 12}$</td>
</tr>
<tr>
<td>9</td>
<td>5.23808</td>
<td>1, ${3, 4}, {6, 7}, {10, 11}, {14, 15}$</td>
</tr>
<tr>
<td>1001</td>
<td>$\approx 503.37$</td>
<td>1, ${3, 4}, {5, 6}, {7, 8}, ...$</td>
</tr>
<tr>
<td>2001</td>
<td>$\approx 1004$</td>
<td>1, ${3, 4}, {5, 6}, {7, 8}, ...$</td>
</tr>
</tbody>
</table>
...or by computer

<table>
<thead>
<tr>
<th>$k$</th>
<th>critical $\nu_k$</th>
<th>roots of the optimal $\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1, ${3, 4}$ or 1, ${4, 5}$</td>
</tr>
<tr>
<td>5</td>
<td>3.11714</td>
<td>1, ${3, 4}$, ${7, 8}$</td>
</tr>
<tr>
<td>7</td>
<td>4.14377</td>
<td>1, ${3, 4}$, ${6, 7}$, ${11, 12}$</td>
</tr>
<tr>
<td>9</td>
<td>5.23808</td>
<td>1, ${3, 4}$, ${6, 7}$, ${10, 11}$, ${14, 15}$</td>
</tr>
<tr>
<td>1001</td>
<td>$\approx$ 503.37</td>
<td>1, ${3, 4}$, ${5, 6}$, ${7, 8}$, $\ldots$</td>
</tr>
<tr>
<td>2001</td>
<td>$\approx$ 1004</td>
<td>1, ${3, 4}$, ${5, 6}$, ${7, 8}$, $\ldots$</td>
</tr>
</tbody>
</table>

Selberg conjectured that $\nu_k \approx \frac{k}{2}$ based on hand calculations.
...or by computer

<table>
<thead>
<tr>
<th>$k$</th>
<th>critical $\nu_k$</th>
<th>roots of the optimal $\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1, ${3, 4}$ or 1, ${4, 5}$</td>
</tr>
<tr>
<td>5</td>
<td>3.11714</td>
<td>1, ${3, 4}$, ${7, 8}$</td>
</tr>
<tr>
<td>7</td>
<td>4.14377</td>
<td>1, ${3, 4}$, ${6, 7}$, ${11, 12}$</td>
</tr>
<tr>
<td>9</td>
<td>5.23808</td>
<td>1, ${3, 4}$, ${6, 7}$, ${10, 11}$, ${14, 15}$</td>
</tr>
<tr>
<td>1001</td>
<td>$\approx$ 503.37</td>
<td>1, ${3, 4}$, ${5, 6}$, ${7, 8}$, ${10, 11}$, ${14, 15}$</td>
</tr>
<tr>
<td>2001</td>
<td>$\approx$ 1004</td>
<td>1, ${3, 4}$, ${5, 6}$, ${7, 8}$, ${10, 11}$, ${14, 15}$</td>
</tr>
</tbody>
</table>

- Selberg conjectured that $\nu_k \approx \frac{k}{2}$ based on hand calculations.
- Selberg was able to prove that

$$\left\lfloor \frac{k + 1}{2} \right\rfloor \leq \nu_k \leq k$$

for all $k$. 
Selberg’s lower bound

Selberg has a famous construction of a “good enough” sieve which is easy to work with.
Selberg’s lower bound

- Selberg has a famous construction of a “good enough” sieve which is easy to work with.
- In this context, we try $\theta$ of the form
  \[
  \theta(x) = (1 - x)f(x)^2,
  \]
  for an arbitrary polynomial $f(x)$ of degree $d = \frac{k-1}{2}$. 

By a miracle, we can optimize this quadratic form by hand!
Selberg’s lower bound

- Selberg has a famous construction of a “good enough” sieve which is easy to work with.
- In this context, we try $\theta$ of the form

$$\theta(x) = (1 - x)f(x)^2,$$

for an arbitrary polynomial $f(x)$ of degree $d = \frac{k-1}{2}$.
- The objective becomes a quadratic function of the coefficients of $f(x)$.
Selberg’s lower bound

- Selberg has a famous construction of a “good enough” sieve which is easy to work with.
- In this context, we try $\theta$ of the form

$$\theta(x) = (1 - x)f(x)^2,$$

for an arbitrary polynomial $f(x)$ of degree $d = \frac{k-1}{2}$.
- The objective becomes a quadratic function of the coefficients of $f(x)$.
- By a miracle, we can optimize this quadratic form by hand!
Selberg’s lower bound: the quadratic form

- Write out $f$ in the binomial basis as

$$f(n) = \sum_{r \leq d} \ell_r \binom{n}{r}.$$
Selberg’s lower bound: the quadratic form

- Write out $f$ in the binomial basis as

$$f(n) = \sum_{r \leq d} \ell_r \binom{n}{r}.$$ 

- We change coordinates to $y_r$ given by

$$y_r = (-1)^r \sum_{i \geq 0} \ell_{r+i} \frac{\nu^i}{i!}.$$
Selberg’s lower bound: the quadratic form

- Write out $f$ in the binomial basis as

$$f(n) = \sum_{r \leq d} \ell_r \binom{n}{r}.$$  

- We change coordinates to $y_r$ given by

$$y_r = (-1)^r \sum_{i \geq 0} \ell_{r+i} \frac{\nu^i}{i!}.$$  

- Our objective function is

$$e^{-\nu} \sum_{n \geq 0} (1 - n) f(n)^2 \frac{\nu^n}{n!} = \sum_r \frac{\nu^r}{r!} y_r^2 - \sum_r \frac{\nu^{r+1}}{r!} (y_r - y_{r+1})^2.$$
Selberg’s lower bound: the quadratic form

- Write out $f$ in the binomial basis as
  
  $$f(n) = \sum_{r \leq d} \ell_r \binom{n}{r}.$$ 

- We change coordinates to $y_r$ given by
  
  $$y_r = (-1)^r \sum_{i \geq 0} \ell_{r+i} \frac{\nu^i}{i!}.$$ 

- Our objective function is
  
  $$e^{-\nu} \sum_{n \geq 0} (1 - n)f(n)^2 \frac{\nu^n}{n!} = \sum_r \frac{\nu^r}{r!} y_r^2 - \sum_r \frac{\nu^{r+1}}{r!} (y_r - y_{r+1})^2.$$ 

- This becomes negative semidefinite when $\nu = d + 1 = \frac{k+1}{2}$. 
Can we make a better sieve?

- I want to know how much we can improve Selberg’s construction.
Can we make a better sieve?

- I want to know how much we can improve Selberg’s construction.

- **Idea:** We know the optimal $\theta$ has the form

\[
\theta(x) = (1 - x)f(x)f(x + 1)
\]

for some polynomial $f$ with integer roots.
Can we make a better sieve?

- I want to know how much we can improve Selberg’s construction.

- **Idea:** We know the optimal $\theta$ has the form

$$\theta(x) = (1 - x)f(x)f(x + 1)$$

for some polynomial $f$ with integer roots.

- What if we drop the condition that $f$ has integer roots?
Can we make a better sieve?

- I want to know how much we can improve Selberg’s construction.

- **Idea:** We know the optimal $\theta$ has the form

  \[ \theta(x) = (1 - x)f(x)f(x + 1) \]

  for some polynomial $f$ with integer roots.

- What if we drop the condition that $f$ has integer roots?

- This will *over-estimate* the best possible lower bound on $\mathbb{P}[X = 0]$. 
A more difficult quadratic form

- We use the same change of variables $y_r$ as in Selberg’s construction.
A more difficult quadratic form

- We use the same change of variables $y_r$ as in Selberg’s construction.

- Our objective function is

$$e^{-\nu} \sum_{n \geq 0} (1 - n)f(n)f(n + 1) \frac{\nu^n}{n!} =$$

$$\sum_r \frac{\nu^r}{r!} y_r(y_r - y_{r+1}) - \sum_r \frac{\nu^{r+1}}{r!} (y_r - y_{r+1})(y_r - 2y_{r+1} + y_{r+2}).$$

- Selberg had to deal with a tridiagonal matrix, I have to deal with a pentadiagonal matrix!
A more difficult quadratic form

- We use the same change of variables $y_r$ as in Selberg’s construction.

- Our objective function is

$$e^{-v} \sum_{n \geq 0} (1 - n)f(n)f(n + 1) \frac{\nu^n}{n!} =$$

$$\sum_{r} \frac{\nu^r}{r!} y_r(y_r - y_{r+1}) - \sum_{r} \frac{\nu^{r+1}}{r!} (y_r - y_{r+1})(y_r - 2y_{r+1} + y_{r+2}).$$

- Selberg had to deal with a tridiagonal matrix, I have to deal with a pentadiagonal matrix!
Everything somehow works out

- I want to prove that this horrible pentadiagonal symmetric matrix is negative semidefinite for $\nu$ large.

Theorem
For $k = 2d + 1$, we have $\nu_k \leq d + 2\sqrt{d + 1}$. 

This result is not best-possible: numerical calculations indicate it can be improved to $\nu_k \leq d + \sqrt{d^2 + O(1)}$. 
Everything somehow works out

- I want to prove that this horrible pentadiagonal symmetric matrix is negative semidefinite for $\nu$ large.
- I computed the Cholesky decomposition for numerical examples to get a hint.

Theorem
For $k = 2d + 1$, we have $\nu k \leq d + 2\sqrt{d} + 1$.

This result is not best-possible: numerical calculations indicate it can be improved to $\nu k \leq d + \sqrt{d^2 + O(1)}$. 
Everything somehow works out

- I want to prove that this horrible pentadiagonal symmetric matrix is negative semidefinite for $\nu$ large.
- I computed the Cholesky decomposition for numerical examples to get a hint.
- Eventually I found a (somewhat) clean proof that that it is negative semidefinite for $\nu \geq (\sqrt{d} + 1)^2$. 

\textbf{Theorem} For $k = 2d + 1$, we have $\nu_k \leq d + 2\sqrt{d} + 1$.

This result is not best-possible: numerical calculations indicate it can be improved to $\nu_k \leq d + \sqrt{d^2 + 1} + O(1)$. 
Everything somehow works out

- I want to prove that this horrible pentadiagonal symmetric matrix is negative semidefinite for \( \nu \) large.
- I computed the Cholesky decomposition for numerical examples to get a hint.
- Eventually I found a (somewhat) clean proof that that it is negative semidefinite for \( \nu \geq (\sqrt{d} + 1)^2 \).

- **Theorem**
  
  For \( k = 2d + 1 \), we have \( \nu_k \leq d + 2\sqrt{d} + 1 \).
Everything somehow works out

- I want to prove that this horrible pentadiagonal symmetric matrix is negative semidefinite for $\nu$ large.
- I computed the Cholesky decomposition for numerical examples to get a hint.
- Eventually I found a (somewhat) clean proof that it is negative semidefinite for $\nu \geq (\sqrt{d} + 1)^2$.

- **Theorem**
  For $k = 2d + 1$, we have $\nu_k \leq d + 2\sqrt{d} + 1$.

- This result is not best-possible: numerical calculations indicate it can be improved to $\nu_k \leq d + \frac{\sqrt{d}}{2} + O(1)$. 
Can we really get a square-root improvement?

- In our relaxed setting, it is possible to construct a polynomial $f(x)$ of degree $d$ such that

$$\sum_{n\geq 0}(1 - n)f(n)f(n + 1)\frac{\nu^n}{n!} > 0$$

with $\nu \geq d + \Omega(\sqrt{d})$. 
Can we really get a square-root improvement?

- In our relaxed setting, it is possible to construct a polynomial $f(x)$ of degree $d$ such that

$$\sum_{n \geq 0} (1 - n)f(n)f(n + 1)\frac{\nu^n}{n!} > 0$$

with $\nu \geq d + \Omega(\sqrt{d})$.

- Does this mean that $\nu_{2d+1} \geq d + \Omega(\sqrt{d})$?
Can we really get a square-root improvement?

- In our relaxed setting, it is possible to construct a polynomial \( f(x) \) of degree \( d \) such that

\[
\sum_{n \geq 0} (1 - n)f(n)f(n + 1) \frac{\nu^n}{n!} > 0
\]

with \( \nu \geq d + \Omega(\sqrt{d}) \).

- Does this mean that \( \nu_{2d+1} \geq d + \Omega(\sqrt{d}) \)?

- The first few roots of such an \( f \) (for \( d \sim 500 \)) are

\[
1, \{2.53, 3.53\}, \{5.19, 6.19\}, \{7.43, 8.43\}, ...
\]
Can we really get a square-root improvement?

- In our relaxed setting, it is possible to construct a polynomial $f(x)$ of degree $d$ such that

$$
\sum_{n \geq 0} (1 - n)f(n)f(n + 1)\frac{\nu^n}{n!} > 0
$$

with $\nu \geq d + \Omega(\sqrt{d})$.

- Does this mean that $\nu_{2d+1} \geq d + \Omega(\sqrt{d})$?

- The first few roots of such an $f$ (for $d \sim 500$) are

  1, {2.53, 3.53}, {5.19, 6.19}, {7.43, 8.43}, ...

- Most of the improvement can be traced back to allowing the second and third roots to be at 2.5 and 3.5.
How much of an improvement can we really get?

- I don’t believe in a square-root improvement, but I want to show there is a real, definite improvement we can make.
How much of an improvement can we really get?

- I don’t believe in a square-root improvement, but I want to show there is a real, definite improvement we can make.

- **Idea:** Take the roots from Selberg’s construction, and round each multiplicity-two root up and down.
How much of an improvement can we really get?

- I don’t believe in a square-root improvement, but I want to show there is a real, definite improvement we can make.

- **Idea:** Take the roots from Selberg’s construction, and round each multiplicity-two root up and down.

- Numerically, this seems to give us a (small) improvement.
How much of an improvement can we really get?

- I don’t believe in a square-root improvement, but I want to show there is a real, definite improvement we can make.

- **Idea:** Take the roots from Selberg’s construction, and round each multiplicity-two root up and down.

- Numerically, this seems to give us a (small) improvement.

- **Problem:** we can’t guarantee that doing this rounding won’t make things worse.
How to make an improvement safely

- Recall our objective function (up to scale):

\[ \sum_n \theta(n) \frac{\nu^n}{n!}. \]
How to make an improvement safely

- Recall our objective function (up to scale):
  \[ \sum_{n} \theta(n) \frac{\nu^n}{n!}. \]
- Every single summand, other than \( \theta(0) \), is negative (or 0).
How to make an improvement safely

- Recall our objective function (up to scale):
  \[ \sum_n \theta(n) \frac{\nu^n}{n!}. \]

- Every single summand, other than \( \theta(0) \), is negative (or 0).

- **Idea:** To guarantee that the objective increases, we try to **decrease** the absolute value \( |\theta(n)| \) for all \( n \in \mathbb{N}^+ \).
Safer rounding

- Write Selberg’s $\theta(x)$ as a product:

\[ \theta(x) = (1 - x) \left(1 - \frac{x}{r_1}\right)^2 \cdots \left(1 - \frac{x}{r_d}\right)^2. \]
Safer rounding

- Write Selberg's $\theta(x)$ as a product:

$$\theta(x) = (1 - x) \left(1 - \frac{x}{r_1}\right)^2 \cdots \left(1 - \frac{x}{r_d}\right)^2.$$ 

- Replace each factor $(1 - x/r_i)^2$ by a quadratic $q_i(x)$ such that:
  - $q_i(0) = 1$,
  - $q_i(x) \geq 0$ for $x \in \mathbb{N}^+$,
  - $q_i(x) \leq \left(1 - \frac{x}{r_i}\right)^2$ for $x \in \mathbb{N}^+$, and
  - at least one of $\lfloor r_i \rfloor$, $\lceil r_i \rceil$ is a root of $q_i(x)$. 

This definitely doesn't hurt us. Does it help?

We can now guarantee that at least one of $\theta(\lfloor r_i \rfloor)$, $\theta(\lceil r_i \rceil)$ has been replaced with 0!
Safer rounding

- Write Selberg’s $\theta(x)$ as a product:

\[ \theta(x) = (1 - x) \left(1 - \frac{x}{r_1}\right)^2 \cdots \left(1 - \frac{x}{r_d}\right)^2. \]

- Replace each factor $(1 - x/r_i)^2$ by a quadratic $q_i(x)$ such that:
  - $q_i(0) = 1$, 
  - $q_i(x) \geq 0$ for $x \in \mathbb{N}^+$, 
  - $q_i(x) \leq (1 - x/r_i)^2$ for $x \in \mathbb{N}^+$, 
  - and at least one of $\lfloor r_i \rfloor, \lceil r_i \rceil$ is a root of $q_i(x)$.

This definitely doesn’t hurt us. Does it help?

- We can now guarantee that at least one of $\theta(\lfloor r_i \rfloor), \theta(\lceil r_i \rceil)$ has been replaced with 0!
Safer rounding

- Write Selberg’s $\theta(x)$ as a product:

$$\theta(x) = (1 - x)(1 - \frac{x}{r_1})^2 \cdots (1 - \frac{x}{r_d})^2.$$  

- Replace each factor $(1 - x/r_i)^2$ by a quadratic $q_i(x)$ such that:
  - $q_i(0) = 1$,
  - $q_i(x) \geq 0$ for $x \in \mathbb{N}^+$,
  - at least one of $\lfloor r_i \rfloor$, $\lceil r_i \rceil$ is a root of $q_i(x)$.

This definitely doesn't hurt us. Does it help?

We can now guarantee that at least one of $\theta(\lfloor r_i \rfloor)$, $\theta(\lceil r_i \rceil)$ has been replaced with 0!
Safer rounding

- Write Selberg’s $\theta(x)$ as a product:

$$\theta(x) = (1 - x)\left(1 - \frac{x}{r_1}\right)^2 \cdots \left(1 - \frac{x}{r_d}\right)^2.$$ 

- Replace each factor $(1 - x/r_i)^2$ by a quadratic $q_i(x)$ such that:
  - $q_i(0) = 1$,
  - $q_i(x) \geq 0$ for $x \in \mathbb{N}^+$,
  - $q_i(x) \leq (1 - x/r_i)^2$ for $x \in \mathbb{N}^+$, and
  - at least one of $\lfloor r_i \rfloor, \lceil r_i \rceil$ is a root of $q_i(x)$. 

This definitely doesn’t hurt us. Does it help?

We can now guarantee that at least one of $\theta(\lfloor r_i \rfloor), \theta(\lceil r_i \rceil)$ has been replaced with 0!
Safer rounding

- Write Selberg’s $\theta(x)$ as a product:

$$\theta(x) = (1 - x)(1 - \frac{x}{r_1})^2 \cdots (1 - \frac{x}{r_d})^2.$$ 

- Replace each factor $(1 - x/r_i)^2$ by a quadratic $q_i(x)$ such that:
  - $q_i(0) = 1$,
  - $q_i(x) \geq 0$ for $x \in \mathbb{N}^+$,
  - $q_i(x) \leq (1 - x/r_i)^2$ for $x \in \mathbb{N}^+$, and
  - at least one of $\lfloor r_i \rfloor$, $\lceil r_i \rceil$ is a root of $q_i(x)$. 

This definitely doesn't hurt us. Does it help?

We can now guarantee that at least one of $\theta(\lfloor r_i \rfloor)$, $\theta(\lceil r_i \rceil)$ has been replaced with 0!
Safer rounding

- Write Selberg’s $\theta(x)$ as a product:

$$\theta(x) = (1 - x) \left(1 - \frac{x}{r_1}\right)^2 \cdots \left(1 - \frac{x}{r_d}\right)^2.$$ 

- Replace each factor $(1 - x/r_i)^2$ by a quadratic $q_i(x)$ such that:
  - $q_i(0) = 1$,
  - $q_i(x) \geq 0$ for $x \in \mathbb{N}^+$,
  - $q_i(x) \leq (1 - x/r_i)^2$ for $x \in \mathbb{N}^+$, and
  - at least one of $\lfloor r_i \rfloor$, $\lceil r_i \rceil$ is a root of $q_i(x)$.

- This definitely doesn’t hurt us. Does it help?
Safer rounding

- Write Selberg’s $\theta(x)$ as a product:

$$\theta(x) = (1 - x) \left(1 - \frac{x}{r_1}\right)^2 \cdots \left(1 - \frac{x}{r_d}\right)^2.$$ 

- Replace each factor $(1 - x/r_i)^2$ by a quadratic $q_i(x)$ such that:
  - $q_i(0) = 1$,
  - $q_i(x) \geq 0$ for $x \in \mathbb{N}^+$,
  - $q_i(x) \leq (1 - x/r_i)^2$ for $x \in \mathbb{N}^+$, and
  - at least one of $\lfloor r_i \rfloor$, $\lceil r_i \rceil$ is a root of $q_i(x)$.

- This definitely doesn’t hurt us. Does it help?
- We can now guarantee that at least one of $\theta(\lfloor r_i \rfloor), \theta(\lceil r_i \rceil)$ has been replaced with 0!
An understandable improvement

- If we perform the safer rounding, we guarantee improving our (rescaled) objective function by at least

\[
\sum_{r_i} \min\left( |\theta([r_i])| \frac{\nu^{[r_i]}}{[r_i]!}, \ |\theta([r_i])| \frac{\nu^{[r_i]}}{[r_i]!} \right).
\]

- So now we need to understand two things:
  - Where are the roots of Selberg's function \( \theta \)?
  - How big is \( \theta \) at the nearby integers?

- We have exact, combinatorial formulas for the coefficients of Selberg's function.

- Slight wrinkle: Selberg's function is optimized for \( \nu = d + 1 \). So we modify it for larger \( \nu \), before rounding.
An understandable improvement

- If we perform the safer rounding, we guarantee improving our (rescaled) objective function by at least

\[
\sum_{r_i} \min \left( |\theta([r_i])| \frac{\nu^{\lfloor r_i \rfloor}}{[r_i]!}, \ |\theta([r_i])| \frac{\nu^{\lfloor r_i \rfloor}}{[r_i]!} \right).
\]

- So now we need to understand two things:
An understandable improvement

- If we perform the safer rounding, we guarantee improving our (rescaled) objective function by at least

\[ \sum_{r_i} \min \left( |\theta(\lfloor r_i \rfloor)| \frac{\nu^{|r_i|}}{|r_i|!}, |\theta(\lceil r_i \rceil)| \frac{\nu^{|r_i|}}{|r_i|!} \right). \]

- So now we need to understand two things:
  - Where are the roots of Selberg’s function \( \theta \)?
An understandable improvement

- If we perform the safer rounding, we guarantee improving our (rescaled) objective function by at least

\[
\sum_{r_i} \min \left( \left| \theta([r_i]) \right| \frac{\nu[r_i]}{[r_i]!}, \left| \theta([r_i]) \right| \frac{\nu[r_i]}{[r_i]!} \right).
\]

- So now we need to understand two things:
  - Where are the roots of Selberg’s function \( \theta \)?
  - How big is \( \theta \) at the nearby integers?
An understandable improvement

- If we perform the safer rounding, we guarantee improving our (rescaled) objective function by at least

\[
\sum_{r_i} \min \left( |\theta(\lfloor r_i \rfloor)| \frac{\nu^{|r_i|}}{\lfloor r_i \rfloor !}, |\theta(\lceil r_i \rceil)| \frac{\nu^{|r_i|}}{\lceil r_i \rceil !} \right).
\]

- So now we need to understand two things:
  - Where are the roots of Selberg’s function \( \theta \)?
  - How big is \( \theta \) at the nearby integers?

- We have exact, combinatorial formulas for the coefficients of Selberg’s function.
An understandable improvement

- If we perform the safer rounding, we guarantee improving our (rescaled) objective function by at least

\[ \sum_{r_i} \min \left( |\theta(\lfloor r_i \rfloor)| \frac{\nu^{\lfloor r_i \rfloor}}{[r_i]!}, |\theta(\lceil r_i \rceil)| \frac{\nu^{\lceil r_i \rceil}}{[r_i]!} \right). \]

- So now we need to understand two things:
  - Where are the roots of Selberg’s function \( \theta \)?
  - How big is \( \theta \) at the nearby integers?

- We have exact, combinatorial formulas for the coefficients of Selberg’s function.

- Slight wrinkle: Selberg’s function is optimized for \( \nu = d + 1 \). So we modify it for larger \( \nu \), before rounding.
Explicit formula for Selberg's function

- Selberg's function is \( \theta(x) = (1 - x)f(x)^2 \), where \( f \) is given by

\[
f(n + 2) = \frac{1}{(d + 1)^{n+1}} \sum_i (-1)^i a(n, i)d^i.
\]
Explicit formula for Selberg’s function

Selberg’s function is \( \theta(x) = (1 - x)f(x)^2 \), where \( f \) is given by

\[
f(n + 2) = \frac{1}{(d + 1)^{n+1}} \sum_i (-1)^i a(n, i)d^i.
\]

Here \( a(n, i) \) is the number of permutations of an \( n \)-set having exactly \( i \) cycles of size greater than 1.
Explicit formula for Selberg’s function

- Selberg’s function is \( \theta(x) = (1 - x)f(x)^2 \), where \( f \) is given by

\[
f(n + 2) = \frac{1}{(d + 1)^{n+1}} \sum_i (-1)^i a(n, i) d^i.
\]

- Here \( a(n, i) \) is the number of permutations of an \( n \)-set having exactly \( i \) cycles of size greater than 1.

- For \( \nu > d + 1 \), we use the function \( f_\nu \) given by

\[
f_\nu(n + 2) = \frac{1}{\nu^{n+1}} \sum_i (-1)^i a_q(n, i) d^i,
\]

where \( q = \nu - d \) and

\[
a_q(n, i) = \sum_{\sigma \in S_n, i \text{ nontrivial cycles}} q^{\# \text{Fix}(\sigma)}.
\]
Let’s at least understand $f(3)$ and $f(4)$

- To understand the contribution from rounding at the smallest root, we compute $f_\nu(3)$ and $f_\nu(4)$.
Let’s at least understand $f(3)$ and $f(4)$

- To understand the contribution from rounding at the smallest root, we compute $f_\nu(3)$ and $f_\nu(4)$.
- We have

$$f_\nu(1 + 2) = \frac{1}{\nu^{1+1}} (a_q(1, 0)d^0) = \frac{q}{\nu^2},$$

and

$$f_\nu(2 + 2) = \frac{1}{\nu^{2+1}} (a_q(2, 0)d^0 - a_q(2, 1)d^1) = -\frac{d - q^2}{\nu^3}.$$
Let’s at least understand $f(3)$ and $f(4)$

- To understand the contribution from rounding at the smallest root, we compute $f_\nu(3)$ and $f_\nu(4)$.

- We have

$$f_\nu(1 + 2) = \frac{1}{\nu^{1+1}}(a_q(1, 0)d^0) = \frac{q}{\nu^2},$$

and

$$f_\nu(2 + 2) = \frac{1}{\nu^{2+1}}(a_q(2, 0)d^0 - a_q(2, 1)d^1) = -\frac{d - q^2}{\nu^3}.$$

- These have opposite sign, so $f_\nu$ has a root between 3 and 4, and both $|f_\nu(3)|, |f_\nu(4)|$ are $\gg \frac{1}{d^2}$. 

Most of the contribution to $f_\nu(n)$ comes from permutations which are almost entirely 2-cycles, so the result depends heavily on whether $n$ is even or odd.
Let’s at least understand $f(3)$ and $f(4)$

- To understand the contribution from rounding at the smallest root, we compute $f_\nu(3)$ and $f_\nu(4)$.

- We have

\[
  f_\nu(1 + 2) = \frac{1}{\nu^{1+1}}(a_q(1, 0)d^0) = \frac{q}{\nu^2},
\]

and

\[
  f_\nu(2 + 2) = \frac{1}{\nu^{2+1}}(a_q(2, 0)d^0 - a_q(2, 1)d^1) = -\frac{d - q^2}{\nu^3}.
\]

- These have opposite sign, so $f_\nu$ has a root between 3 and 4, and both $|f_\nu(3)|, |f_\nu(4)|$ are $\gg \frac{1}{d^2}$.

- Most of the contribution to $f_\nu(n)$ comes from permutations which are almost entirely 2-cycles, so the result depends heavily on whether $n$ is even or odd.
I continued with the combinatorial analysis, eventually proving that $a_q(n, i)$ is log-concave in $i$ in order to get strong enough approximations...
Saddle point method

- I continued with the combinatorial analysis, eventually proving that $a_q(n, i)$ is log-concave in $i$ in order to get strong enough approximations...

- My advisor (Sound) suggested a different approach.
Saddle point method

I continued with the combinatorial analysis, eventually proving that \( a_q(n, i) \) is log-concave in \( i \) in order to get strong enough approximations...

My advisor (Sound) suggested a different approach.

We can compute \( f_\nu \) via a contour integral:

\[
f_\nu(n + 2) = \frac{n!}{2\pi i} \int_C e^{\nu z} (1 - z)^d \frac{dz}{z^{n+1}}.
\]
Saddle point method

- I continued with the combinatorial analysis, eventually proving that $a_q(n, i)$ is log-concave in $i$ in order to get strong enough approximations...

- My advisor (Sound) suggested a different approach.

- We can compute $f_\nu$ via a contour integral:

$$f_\nu(n + 2) = \frac{n!}{2\pi i} \int_C e^{\nu z}(1 - z)^d \frac{dz}{z^{n+1}}.$$

- The integrand has saddle points at $z_0, \bar{z}_0$ solving the quadratic

$$\nu z_0^2 - (n + q)z_0 + n = 0.$$
Saddle point method

- I continued with the combinatorial analysis, eventually proving that $a_q(n, i)$ is log-concave in $i$ in order to get strong enough approximations...

- My advisor (Sound) suggested a different approach.
- We can compute $f_\nu$ via a contour integral:

  $$f_\nu(n + 2) = \frac{n!}{2\pi i} \int_C e^{\nu z} (1 - z)^d \frac{dz}{z^{n+1}}.$$

- The integrand has saddle points at $z_0, \bar{z}_0$ solving the quadratic

  $$\nu z_0^2 - (n + q)z_0 + n = 0.$$

- Either way, we get a somewhat complicated sinusoidal expression for $f_\nu$. 
The dust settles

Theorem

If \( k = 2d + 1 \) then

\[ \nu_k - d \geq (c + o(1)) \sqrt[3]{d}, \]

where \( c \approx \frac{1}{12.14} \) is the greatest positive solution of the inequality

\[
\int_0^{\infty} \frac{1}{x^{3/2}} \min \left( \sin^2 \left( (\frac{x^3}{3} + c) \sqrt{x} \right), \cos^2 \left( (\frac{x^3}{3} + c) \sqrt{x} \right) \right) \, dx \geq 2\pi c.
\]
Thank you for your attention.