E_2 numbers in almost all short intervals

Kaisa Matomäki (joint with Joni Teräväinen)

University of Turku, Finland

ELAZ, August 22nd, 2022

Contents

Outline

1. Background and results
   - Primes in short intervals
   - Primes in almost all short intervals
   - Almost primes in (almost all) short intervals
2. Methods
   - Harman's sieve
   - Reductions
   - Type I and II estimates
3. Summary and more theorems

Primes in short intervals

- By the prime number theorem, the number of primes up to x is about $x/\log x$.
- One wants to know about primes in short intervals: If we look at a "short" segment $(x, x+H]$ around x, is the density of primes in that segment still $1/\log x$?

Primes in short intervals

- By the prime number theorem, the number of primes up to x is about $x/\log x$.
- One wants to know about primes in short intervals: If we look at a "short" segment $(x, x+H]$ around x, is the density of primes in that segment still $1/\log x$?
- The smaller the H, the more difficult the problem.

Primes in short intervals

- By the prime number theorem, the number of primes up to x is about $x/\log x$.
- One wants to know about primes in short intervals: If we look at a "short" segment $(x, x+H]$ around x, is the density of primes in that segment still $1/\log x$?
- The smaller the H, the more difficult the problem.
- Huxley's prime number theorem from 1972 gives

$$\sum_{x<p\leq x+H} 1 = (1+o(1))\frac{H}{\log x}, \quad H \geq x^{7/12+\varepsilon}.$$

Primes in short intervals

- By the prime number theorem, the number of primes up to x is about $x/\log x$.
- One wants to know about primes in short intervals: If we look at a "short" segment $(x, x+H]$ around x, is the density of primes in that segment still $1/\log x$?
- The smaller the H, the more difficult the problem.
- Huxley's prime number theorem from 1972 gives

$$\sum_{x<p\leq x+H} 1 = (1+o(1))\frac{H}{\log x}, \quad H \geq x^{7/12+\varepsilon}.$$

- This is based on Huxley's zero-density estimate for the zeta function and has resisted improvements, except Heath-Brown (1988) has shown this for $H \geq x^{7/12-o(1)}$.

Primes in short intervals

- Baker-Harman-Pintz (2001) showed with a sieve method

$$\sum_{x<p\leq x+H} 1 \geq \varepsilon \frac{H}{\log X}, \quad H \geq x^{0.525}$$

for some $\varepsilon > 0$.

Primes in short intervals

- Baker-Harman-Pintz (2001) showed with a sieve method

$$\sum_{x<p\leq x+H} 1 \geq \varepsilon \frac{H}{\log X}, \quad H \geq x^{0.525}$$

  for some $\varepsilon > 0$.
- For shorter intervals one does not even know existence of primes!
- Assuming RH one knows that $[x, x + x^{1/2} \log x]$ always contains primes.

Primes in short intervals

- Baker-Harman-Pintz (2001) showed with a sieve method

$$\sum_{x<p\leq x+H} 1 \geq \varepsilon \frac{H}{\log X}, \quad H \geq x^{0.525}$$

  for some $\varepsilon > 0$.
- For shorter intervals one does not even know existence of primes!
- Assuming RH one knows that $[x, x + x^{1/2} \log x]$ always contains primes.
- Cramer made a probabilistic model based on "probability of n being prime is $1/\log n$". Based on this, one expects that intervals $[x, x + (\log x)^{2+\varepsilon}]$ contain primes for all large x.

Primes in short intervals

- Baker-Harman-Pintz (2001) showed with a sieve method

$$\sum_{x<p\leq x+H} 1 \geq \varepsilon \frac{H}{\log X}, \quad H \geq x^{0.525}$$

  for some $\varepsilon > 0$.
- For shorter intervals one does not even know existence of primes!
- Assuming RH one knows that $[x, x + x^{1/2} \log x]$ always contains primes.
- Cramer made a probabilistic model based on "probability of n being prime is $1/\log n$". Based on this, one expects that intervals $[x, x + (\log x)^{2+\varepsilon}]$ contain primes for all large x.
- Huge gap between what's known and what's expected!

Primes in almost all short intervals

- Even under RH it is not known that $[x, x + x^{1/2}]$ always contains primes.
- What if one only requires that almost all intervals contain primes?

Primes in almost all short intervals

- Even under RH it is not known that $[x, x + x^{1/2}]$ always contains primes.
- What if one only requires that almost all intervals contain primes?
- A variant of Huxley's prime number theorem says that, for almost all $x \in [X, 2X]$ (i.e. with $o(X)$ exceptions),

$$\sum_{x < p \leq x+H} 1 = (1 + o(1)) \frac{H}{\log X}, \quad H \geq x^{1/6+\varepsilon}.$$

- This can be proved using the same zero-density estimates and has also resisted improvements.

Primes in almost all short intervals

- Even under RH it is not known that $[x, x + x^{1/2}]$ always contains primes.
- What if one only requires that almost all intervals contain primes?
- A variant of Huxley's prime number theorem says that, for almost all $x \in [X, 2X]$ (i.e. with $o(X)$ exceptions),

$$\sum_{x < p \leq x+H} 1 = (1 + o(1)) \frac{H}{\log X}, \quad H \geq x^{1/6+\varepsilon}.$$

- This can be proved using the same zero-density estimates and has also resisted improvements.
- A lower bound has been shown for $H \geq X^{1/20}$ by Jia.

Primes in almost all short intervals

- Even under RH it is not known that $[x, x + x^{1/2}]$ always contains primes.
- What if one only requires that almost all intervals contain primes?
- A variant of Huxley's prime number theorem says that, for almost all $x \in [X, 2X]$ (i.e. with $o(X)$ exceptions),

$$\sum_{x<p\leq x+H} 1 = (1+o(1))\frac{H}{\log X}, \quad H \geq x^{1/6+\varepsilon}.$$

- This can be proved using the same zero-density estimates and has also resisted improvements.
- A lower bound has been shown for $H \geq X^{1/20}$ by Jia.
- One expects that, for any $h \to \infty$ with $X \to \infty$, the interval $(x, x + h \log x]$ contains primes for almost all $x \in [X, 2X]$.

Almost primes

- One expects that, for any $h \to \infty$ with $X \to \infty$, the interval $(x - h \log X, x]$ contains primes for almost all $x \in [X/2, X]$.
- One can ask similar questions about almost-primes, i.e. P_k numbers that have at most k prime factors or E_k numbers that have exactly k prime factors.

Almost primes

- One expects that, for any $h \to \infty$ with $X \to \infty$, the interval $(x - h\log X, x]$ contains primes for almost all $x \in [X/2, X]$.
- One can ask similar questions about almost-primes, i.e. P_k numbers that have at most k prime factors or E_k numbers that have exactly k prime factors.
- I have recently shown that, as soon as $h \to \infty$ with $X \to \infty$, the interval $(x - h\log X, x]$ contains P_2-numbers for almost all $x \in [X/2, X]$.

Almost primes

- One expects that, for any $h \to \infty$ with $X \to \infty$, the interval $(x - h \log X, x]$ contains primes for almost all $x \in [X/2, X]$.
- One can ask similar questions about almost-primes, i.e. P_k numbers that have at most k prime factors or E_k numbers that have exactly k prime factors.
- I have recently shown that, as soon as $h \to \infty$ with $X \to \infty$, the interval $(x - h \log X, x]$ contains P_2-numbers for almost all $x \in [X/2, X]$.
- Wu has shown that the interval $(x - x^{101/232}, x]$ contains P_2 numbers for all sufficiently large x.
- Teräväinen (2016) has shown that almost all intervals of length $(\log \log X)^{6+\varepsilon} \log X$ contain E_3 numers.

- From now on we will concentrate on E_2 numbers in almost all short intervals.

E_2 numbers in almost all short intervals

- From now on we will concentrate on E_2 numbers in almost all short intervals.
- More difficult than P_2 due to parity barrier

E_2 numbers in almost all short intervals

- From now on we will concentrate on E_2 numbers in almost all short intervals.
- More difficult than P_2 due to parity barrier
- Previous results saying that, for almost all $x \in [X, 2X]$, the interval $(x, x+H]$ contains E_2 numbers:
  - Motohashi (1979): $H = X^{\varepsilon}$
  - Wolke (1979): $H = (\log X)^{5 \cdot 10^6}$
  - Harman (1982): $H = (\log X)^{7+\varepsilon}$
  - Teräväinen (2016): $H = (\log X)^{3.51}$
  - Riemann hypothesis: $H = (\log X)^{2+\varepsilon}$

E_2 numbers in almost all very short intervals

Theorem (M.-Teräväinen (202?))

The interval $(x - (\log X)^{2.1}, x]$ *contains* E_2*-numbers for almost all* $x \in [X/2, X]$.

E_2 numbers in almost all very short intervals

Theorem (M.-Teräväinen (202?))

The interval $(x - (\log X)^{2.1}, x]$ *contains* E_2*-numbers for almost all* $x \in [X/2, X]$.

- Recall that the previous record was exponent 3.51 and even Riemann hypothesis cannot get below 2.

E_2 numbers in almost all very short intervals

Theorem (M.-Teräväinen (202?))

The interval $(x - (\log X)^{2.1}, x]$ *contains* E_2*-numbers for almost all* $x \in [X/2, X]$.

- Recall that the previous record was exponent 3.51 and even Riemann hypothesis cannot get below 2.
- Assuming a slight variant of the density hypothesis, Harman's method would have yielded $3 + \varepsilon$. This was also the limit of Teräväinen's previous result.

E_2 numbers in almost all very short intervals

Theorem (M.-Teräväinen (202?))

The interval $(x - (\log X)^{2.1}, x]$ *contains* E_2*-numbers for almost all* $x \in [X/2, X]$.

- Recall that the previous record was exponent 3.51 and even Riemann hypothesis cannot get below 2.
- Assuming a slight variant of the density hypothesis, Harman's method would have yielded $3 + \varepsilon$. This was also the limit of Teräväinen's previous result.
- We manage to overcome this limitation.
- In this talk I will actually prove this for slightly larger exponent 2.11 and cheat at some places as this simplifies the argument.

Outline

Sketch of the proof

- We study $p_1 p_2 \in (x - (\log X)^{2.11}, x]$ with $p_1 \sim P_1 := (\log X)^{1.11}$, i.e. one of the primes is very small.

Sketch of the proof

- We study $p_1 p_2 \in (x - (\log X)^{2.11}, x]$ with $p_1 \sim P_1 := (\log X)^{1.11}$, i.e. one of the primes is very small.
- For p_2, use Harman's sieve to find a suitable minorant $\rho^-(n) \leq 1_{n \in \mathbb{P}}$ and reduce to studying, for $H = (\log X)^{2.11}$,

$$\int_{X^{1/1000}}^{X/H} \left| \sum_{p_1 \sim P_1} \frac{1}{p_1^{1+it}} \sum_{n \sim X/P_1} \frac{\rho^-(n)}{n^{1+it}} \right|^2 dt.$$

Sketch of the proof

- We study $p_1 p_2 \in (x - (\log X)^{2.11}, x]$ with $p_1 \sim P_1 := (\log X)^{1.11}$, i.e. one of the primes is very small.
- For p_2, use Harman's sieve to find a suitable minorant $\rho^-(n) \leq 1_{n \in \mathbb{P}}$ and reduce to studying, for $H = (\log X)^{2.11}$,

$$\int_{X^{1/1000}}^{X/H} \left| \sum_{p_1 \sim P_1} \frac{1}{p_1^{1+it}} \sum_{n \sim X/P_1} \frac{\rho^-(n)}{n^{1+it}} \right|^2 dt.$$

- If $|\sum_{p_1 \sim P_1} p_1^{-1-it}| \leq P^{-\varepsilon}$, easy.
- Otherwise, decompose $\rho^-(n)$ as appropriate type I and type II sums.

Sketch of the proof

- We study $p_1 p_2 \in (x - (\log X)^{2.11}, x]$ with $p_1 \sim P_1 := (\log X)^{1.11}$, i.e. one of the primes is very small.
- For p_2, use Harman's sieve to find a suitable minorant $\rho^-(n) \leq 1_{n \in \mathbb{P}}$ and reduce to studying, for $H = (\log X)^{2.11}$,

$$\int_{X^{1/1000}}^{X/H} \left| \sum_{p_1 \sim P_1} \frac{1}{p_1^{1+it}} \sum_{n \sim X/P_1} \frac{\rho^-(n)}{n^{1+it}} \right|^2 dt.$$

- If $|\sum_{p_1 \sim P_1} p_1^{-1-it}| \leq P^{-\varepsilon}$, easy.
- Otherwise, decompose $\rho^-(n)$ as appropriate type I and type II sums.
- In each case we amplify by $P^{k\varepsilon} |\sum_{p_1 \sim P_1} p_1^{-1-it}|^k \geq 1$.

Sketch of the proof

- We study $p_1 p_2 \in (x - (\log X)^{2.11}, x]$ with $p_1 \sim P_1 := (\log X)^{1.11}$, i.e. one of the primes is very small.
- For p_2, use Harman's sieve to find a suitable minorant $\rho^-(n) \leq 1_{n \in \mathbb{P}}$ and reduce to studying, for $H = (\log X)^{2.11}$,

$$\int_{X^{1/1000}}^{X/H} \left| \sum_{p_1 \sim P_1} \frac{1}{p_1^{1+it}} \sum_{n \sim X/P_1} \frac{\rho^-(n)}{n^{1+it}} \right|^2 dt.$$

- If $|\sum_{p_1 \sim P_1} p_1^{-1-it}| \leq P^{-\varepsilon}$, easy.
- Otherwise, decompose $\rho^-(n)$ as appropriate type I and type II sums.
- In each case we amplify by $P^{k\varepsilon} |\sum_{p_1 \sim P_1} p_1^{-1-it}|^k \geq 1$.
- We use mean value theorem of Deshouillers-Iwaniec for type I sums.
- For type II sums we use large value theorems and Heath-Brown's recent sparse mean value theorem.

Harman's sieve

- Use Harman's sieve to find a suitable $\rho^-(n) \leq 1_{n \in \mathbb{P}}$:

Harman's sieve

- Use Harman's sieve to find a suitable $\rho^-(n) \leq 1_{n \in \mathbb{P}}$:
- Write $\rho(n, z) = 1_{p | n \implies p > z}$.

Harman's sieve

- Use Harman's sieve to find a suitable $\rho^-(n) \leq 1_{n \in \mathbb{P}}$:
- Write $\rho(n, z) = 1_{p | n \implies p > z}$.
- Buchstab's identity states that, for any $z \geq w \geq 2$,

$$\rho(n, z) = \rho(n, w) - \sum_{\substack{n = pm \\ w \leq p < z}} \rho(m, p).$$

Harman's sieve

- Use Harman's sieve to find a suitable $\rho^{-}(n) \leq 1_{n \in \mathbb{P}}$:
- Write $\rho(n, z) = 1_{p | n \implies p > z}$.
- Buchstab's identity states that, for any $z \geq w \geq 2$,

$$\rho(n, z) = \rho(n, w) - \sum_{\substack{n = pm \\ w \leq p < z}} \rho(m, p).$$

- Harman's sieve is based on applying, for $n \in [X, 2X]$, Buchstab's identity several times to $1_{n \in \mathbb{P}} = \rho(n, 2X^{1/2})$.

Harman's sieve

- Use Harman's sieve to find a suitable $\rho^{-}(n) \leq 1_{n \in \mathbb{P}}$:
- Write $\rho(n, z) = 1_{p|n \implies p > z}$.
- Buchstab's identity states that, for any $z \geq w \geq 2$,

$$\rho(n, z) = \rho(n, w) - \sum_{\substack{n = pm \\ w \leq p < z}} \rho(m, p).$$

- Harman's sieve is based on applying, for $n \in [X, 2X]$, Buchstab's identity several times to $1_{n \in \mathbb{P}} = \rho(n, 2X^{1/2})$.
- This leads to some type I and type II sums, and some sums we cannot deal with. But if they have positive sign, we can discard them when looking for a minorant.

Harman's sieve

- Use Harman's sieve to find a suitable $\rho^-(n) \leq 1_{n \in \mathbb{P}}$:
- Write $\rho(n, z) = 1_{p | n \implies p > z}$.
- Buchstab's identity states that, for any $z \geq w \geq 2$,

$$\rho(n, z) = \rho(n, w) - \sum_{\substack{n = pm \\ w \leq p < z}} \rho(m, p).$$

- Harman's sieve is based on applying, for $n \in [X, 2X]$, Buchstab's identity several times to $1_{n \in \mathbb{P}} = \rho(n, 2X^{1/2})$.
- This leads to some type I and type II sums, and some sums we cannot deal with. But if they have positive sign, we can discard them when looking for a minorant.
- We are not allowed to discard too much, we need $\rho^-(n)$ to have average $\gg X / \log X$.

Recall $\rho(n, z) = 1_{p|n \implies p>z}$. Let $n \in [X, 2X]$, $z = X^{0.185}$. By Buchstab

Harman's sieve

Recall $\rho(n,z) = 1_{p|n \implies p>z}$. Let $n \in [X, 2X]$, $z = X^{0.185}$. By Buchstab

$$1_{n \in \mathbb{P}} = \rho(n, 2X^{1/2}) = \rho(n,z) - \sum_{\substack{n=qm \\ z \leq q < 2X^{1/2}}} \rho(m,q)$$

Recall $\rho(n,z) = 1_{p|n \implies p>z}$. Let $n \in [X, 2X]$, $z = X^{0.185}$. By Buchstab

$$1_{n\in\mathbb{P}} = \rho(n, 2X^{1/2}) = \rho(n,z) - \sum_{\substack{n=qm \\ z\leq q<2X^{1/2}}} \rho(m,q)$$

$$= \rho(n,z) - \sum_{\substack{n=qm \\ z\leq q<2X^{1/2}}} \rho(m,z) + \sum_{\substack{n=q_1q_2m \\ z\leq q_2<q_1<2X^{1/2}}} \rho(m,q_2)$$

Harman's sieve

Recall $\rho(n,z) = 1_{p|n \implies p>z}$. Let $n \in [X, 2X]$, $z = X^{0.185}$. By Buchstab

$$\begin{aligned} 1_{n\in\mathbb{P}} = \rho(n, 2X^{1/2}) &= \rho(n,z) - \sum_{\substack{n=qm \\ z\le q<2X^{1/2}}} \rho(m,q) \\ &= \rho(n,z) - \sum_{\substack{n=qm \\ z\le q<2X^{1/2}}} \rho(m,z) + \sum_{\substack{n=q_1q_2m \\ z\le q_2<q_1<2X^{1/2}}} \rho(m,q_2) \\ &\ge \rho(n,z) - \sum_{\substack{n=qm \\ z\le q<2X^{1/2}}} \rho(m,z) =: \rho^-(n). \end{aligned}$$

Harman's sieve

Recall $\rho(n,z) = 1_{p|n \implies p>z}$. Let $n \in [X, 2X]$, $z = X^{0.185}$. By Buchstab

$$\begin{aligned} 1_{n \in \mathbb{P}} = \rho(n, 2X^{1/2}) &= \rho(n,z) - \sum_{\substack{n=qm \\ z \le q < 2X^{1/2}}} \rho(m,q) \\ &= \rho(n,z) - \sum_{\substack{n=qm \\ z \le q < 2X^{1/2}}} \rho(m,z) + \sum_{\substack{n=q_1 q_2 m \\ z \le q_2 < q_1 < 2X^{1/2}}} \rho(m,q_2) \\ &\ge \rho(n,z) - \sum_{\substack{n=qm \\ z \le q < 2X^{1/2}}} \rho(m,z) =: \rho^-(n). \end{aligned}$$

Now z has been chosen in such a way that by PNT

$$\sum_{\substack{X < q_1 q_2 m \le 2X \\ z \le q_2 < q_1 < 2X^{1/2}}} \rho(m, q_2) \le 0.99 \frac{X}{\log X}$$

Harman's sieve

Recall $\rho(n,z) = 1_{p|n \implies p>z}$. Let $n \in [X, 2X]$, $z = X^{0.185}$. By Buchstab

$$\begin{aligned} 1_{n\in\mathbb{P}} = \rho(n, 2X^{1/2}) &= \rho(n,z) - \sum_{\substack{n=qm \\ z\le q<2X^{1/2}}} \rho(m,q) \\ &= \rho(n,z) - \sum_{\substack{n=qm \\ z\le q<2X^{1/2}}} \rho(m,z) + \sum_{\substack{n=q_1q_2m \\ z\le q_2<q_1<2X^{1/2}}} \rho(m,q_2) \\ &\ge \rho(n,z) - \sum_{\substack{n=qm \\ z\le q<2X^{1/2}}} \rho(m,z) =: \rho^-(n). \end{aligned}$$

Now z has been chosen in such a way that by PNT

$$\sum_{\substack{X<q_1q_2m\le 2X \\ z\le q_2<q_1<2X^{1/2}}} \rho(m,q_2) \le 0.99\frac{X}{\log X} \implies \sum_{X<n\le 2X} \rho^-(n) \ge 0.01\frac{X}{\log X}.$$

$$\rho^-(n) = \rho(n,z) - \sum_{\substack{n=qm \\ z \le q < 2X^{1/2}}} \rho(m,z)$$

$$\rho^-(n) = \rho(n,z) - \sum_{\substack{n=qm \\ z \le q < 2X^{1/2}}} \rho(m,z)$$

$$= \rho(n, X^{\varepsilon^2}) - \sum_{\substack{n=mp \\ X^{\varepsilon^2} \le q < z}} \rho(m,q) - \sum_{\substack{n=qm \\ z \le q < 2X^{1/2}}} \rho(m, X^{\varepsilon^2})$$

$$+ \sum_{\substack{n=q_1 q_2 m \\ X^{\varepsilon^2} \le q_1 < z \le q_2 < 2X^{1/2}}} \rho(m, q_1)$$

$$\begin{aligned}\rho^-(n) &= \rho(n,z) - \sum_{\substack{n=qm \\ z\leq q<2X^{1/2}}} \rho(m,z) \\ &= \rho(n,X^{\varepsilon^2}) - \sum_{\substack{n=mp \\ X^{\varepsilon^2}\leq q<z}} \rho(m,q) - \sum_{\substack{n=qm \\ z\leq q<2X^{1/2}}} \rho(m,X^{\varepsilon^2}) \\ &+ \sum_{\substack{n=q_1q_2m \\ X^{\varepsilon^2}\leq q_1<z\leq q_2<2X^{1/2}}} \rho(m,q_1)\end{aligned}$$

By the fundamental lemma of the sieve, the first and third terms are more-or-less type I sums

$$\sum_{n=rm, r\leq X^{1/2+\varepsilon}} a_r$$

$$
\begin{aligned}
\rho^-(n) &= \rho(n,z) - \sum_{\substack{n=qm \\ z \le q < 2X^{1/2}}} \rho(m,z) \\
&= \rho(n, X^{\varepsilon^2}) - \sum_{\substack{n=mp \\ X^{\varepsilon^2} \le q < z}} \rho(m,q) - \sum_{\substack{n=qm \\ z \le q < 2X^{1/2}}} \rho(m, X^{\varepsilon^2}) \\
&+ \sum_{\substack{n=q_1q_2m \\ X^{\varepsilon^2} \le q_1 < z \le q_2 < 2X^{1/2}}} \rho(m, q_1)
\end{aligned}
$$

By the fundamental lemma of the sieve, the first and third terms are more-or-less type I sums

$$
\sum_{n=rm, r \le X^{1/2+\varepsilon}} a_r
$$

and the second and fourth terms type II sums

$$
\sum_{\substack{n=qr \\ X^{\varepsilon^2} \le q < z}} 1_{q \in \mathbb{P}} a_r.
$$

Reduction to Dirichlet polynomials

- By Perron's formula

$$\sum_{\substack{x<p_1 n\leq x+H \\ p_1\sim P_1}} \rho^-(n) \approx \int_{-X/H}^{X/H} P_1(1+it)P(1+it)\frac{(x+H)^{1+it}-x^{it}}{1+it}dt,$$

where

$$P_1(s) := \sum_{p_1\sim P_1} \frac{1}{p_1^s}, \quad P(s) := \sum_{X/(2P_1)\leq n\leq 4X/P_1} \frac{\rho^-(n)}{n^s},$$

Reduction to Dirichlet polynomials

- By Perron's formula

$$\sum_{\substack{x < p_1 n \leq x+H \\ p_1 \sim P_1}} \rho^-(n) \approx \int_{-X/H}^{X/H} P_1(1+it) P(1+it) \frac{(x+H)^{1+it} - x^{it}}{1+it} dt,$$

where

$$P_1(s) := \sum_{p_1 \sim P_1} \frac{1}{p_1^s}, \quad P(s) := \sum_{X/(2P_1) \leq n \leq 4X/P_1} \frac{\rho^-(n)}{n^s},$$

- Studying the mean square over $x \in [X, 2X]$, it more-or-less suffices to show that

$$\int_{X^{1/1000}}^{X/H} |P_1(1+it)|^2 |P(1+it)|^2 dt \ll \frac{1}{(\log X)^{2+\varepsilon}}.$$

Dirichlet polynomials

- Mean value theorem:

$$\int_{-T}^{T} \left| \sum_{n \sim N} a_n n^{it} \right|^2 dt = (2T + O(N)) \sum_{n \leq N} |a_n|^2$$

Dirichlet polynomials

- Mean value theorem:

$$\int_{-T}^{T} \left| \sum_{n \sim N} a_n n^{it} \right|^2 dt = (2T + O(N)) \sum_{n \leq N} |a_n|^2$$

- Deshouillers-Iwaniec mean value theorem: Let $N \leq T$ and $A \leq T^{1/5}$. Then

$$\int_{T/2}^{T} \left| \sum_{n \sim N} n^{it} \right|^4 \left| \sum_{m \sim A} a_m m^{it} \right|^2 dt \ll_\eta T^{1+\eta} \sum_{m \sim A} |a_m|^2.$$

Dirichlet polynomials

- Mean value theorem:

$$\int_{-T}^{T} \left| \sum_{n \sim N} a_n n^{it} \right|^2 dt = (2T + O(N)) \sum_{n \leq N} |a_n|^2$$

- Deshouillers-Iwaniec mean value theorem: Let $N \leq T$ and $A \leq T^{1/5}$. Then

$$\int_{T/2}^{T} \left| \sum_{n \sim N} n^{it} \right|^4 \left| \sum_{m \sim A} a_m m^{it} \right|^2 dt \ll_\eta T^{1+\eta} \sum_{m \sim A} |a_m|^2.$$

- Heath-Brown's sparse mean value theorem. Let $\mathcal{M} \subseteq [1, T], N \geq T^{2/3}$, and $|\varepsilon_m|, |a_n| \leq 1$. Then

$$\int_{-T}^{T} |\sum_{m \in \mathcal{M}} \varepsilon_m m^{it}|^2 |\sum_{n \sim N} a_n n^{it}|^2 dt \ll_\eta |\mathcal{M}|^2 N^2 + N^\eta |\mathcal{M}| NT.$$

Dirichlet polynomials

- We wish to show that

$$\int_{X^{1/1000}}^{X/H} |P_1(1+it)|^2 |P(1+it)|^2 dt \ll \frac{1}{(\log X)^{2+\varepsilon}}.$$

Dirichlet polynomials

- We wish to show that

$$\int_{X^{1/1000}}^{X/H} |P_1(1+it)|^2 |P(1+it)|^2 dt \ll \frac{1}{(\log X)^{2+\varepsilon}}.$$

- We partition $[X^{1/1000}, X/H] = \mathcal{T} \cup \mathcal{U}$ with

$$\mathcal{T} := \{t \in [X^{1/1000}, X/H] \colon |P_1(1+it)| \leq P_1^{-\varepsilon}\}.$$

Dirichlet polynomials

- We wish to show that

$$\int_{X^{1/1000}}^{X/H} |P_1(1+it)|^2 |P(1+it)|^2 dt \ll \frac{1}{(\log X)^{2+\varepsilon}}.$$

- We partition $[X^{1/1000}, X/H] = \mathcal{T} \cup \mathcal{U}$ with

$$\mathcal{T} := \{t \in [X^{1/1000}, X/H] \colon |P_1(1+it)| \leq P_1^{-\varepsilon}\}.$$

- For $t \in \mathcal{T}$ use the pointwise bound $|P_1(1+it)| \leq P_1^{-\varepsilon}$ and estimate the mean square of $P(1+it)$ using (an improved) mean value theorem.

The required type I and type II estimates

.

- Recall that $\rho^-(n)$ can be decomposed type I and II sums

$$\sum_{n=rm, r \leq X^{1/2+\varepsilon}} a_r \quad \text{and} \quad \sum_{\substack{n=qr \\ X^{\varepsilon^2} \leq q < z}} 1_{q \in \mathbb{P}} a_r.$$

- Suffices to show, for every $R \leq X^{1/2+\varepsilon}$, type I estimate

$$\int_{\mathcal{U}} \left| \sum_{r \sim R} \frac{a_r}{r^{1+it}} \right|^2 \left| \sum_{m \sim X/(P_1 R)} \frac{1}{m^{1+it}} \right|^2 dt \ll \frac{1}{(\log X)^{100}}$$

and, for every $Q \in [X^{\varepsilon}, z]$, type II estimate

$$\int_{\mathcal{U}} \left| \sum_{r \sim X/(P_1 Q)} \frac{a_r}{r^{1+it}} \right|^2 \left| \sum_{q \sim Q} \frac{1_{q \in \mathbb{P}}}{q^{1+it}} \right|^2 dt \ll \frac{1}{(\log X)^{100}}$$

Type I estimate

- Recall $\mathcal{U} := \{t \in [X^{1/1000}, X/H] \colon |P_1(1+it)| > P_1^{-\varepsilon}\}$.

Type I estimate

- Recall $\mathcal{U} := \{t \in [X^{1/1000}, X/H] \colon |P_1(1+it)| > P_1^{-\varepsilon}\}$.
- We wish to show that, for every $R \leq X^{1/2+\varepsilon}$.

$$\int_{\mathcal{U}} \left| \sum_{r \sim R} \frac{a_r}{r^{1+it}} \right|^2 \left| \sum_{m \sim X/(P_1 R)} \frac{1}{m^{1+it}} \right|^2 dt \ll \frac{1}{(\log X)^{100}}$$

Type I estimate

- Recall $\mathcal{U} := \{t \in [X^{1/1000}, X/H] \colon |P_1(1+it)| > P_1^{-\varepsilon}\}$.
- We wish to show that, for every $R \leq X^{1/2+\varepsilon}$,

$$\int_{\mathcal{U}} \left|\sum_{r \sim R} \frac{a_r}{r^{1+it}}\right|^2 \left|\sum_{m \sim X/(P_1 R)} \frac{1}{m^{1+it}}\right|^2 dt \ll \frac{1}{(\log X)^{100}}$$

- Choose k so that $P_1^k \asymp T^{1/10}$. For $t \in \mathcal{U}$, $1 \leq |P_1(1+it)|^k P_1^{k\varepsilon}$. Hence by C-S the above integral is

$$\ll P_1^{k\varepsilon} \left(\int_{X^{1/1000}}^{X/H} \left|\sum_{r \sim R} \frac{a_r}{r^{1+it}}\right|^4\right)^{1/2}$$

$$\cdot \left(\int_{X^{1/1000}}^{X/H} \left|\sum_{m \sim X/(P_1 R)} \frac{1}{m^{1+it}}\right|^4 |P_1(1+it)|^{2k}\, dt\right)^{1/2}.$$

Type I estimate

- Recall $\mathcal{U} := \{t \in [X^{1/1000}, X/H] \colon |P_1(1+it)| > P_1^{-\varepsilon}\}$.
- We wish to show that, for every $R \leq X^{1/2+\varepsilon}$,

$$\int_{\mathcal{U}} \left| \sum_{r \sim R} \frac{a_r}{r^{1+it}} \right|^2 \left| \sum_{m \sim X/(P_1 R)} \frac{1}{m^{1+it}} \right|^2 dt \ll \frac{1}{(\log X)^{100}}$$

- Choose k so that $P_1^k \asymp T^{1/10}$. For $t \in \mathcal{U}$, $1 \leq |P_1(1+it)|^k P_1^{k\varepsilon}$. Hence by C-S the above integral is

$$\ll P_1^{k\varepsilon} \left(\int_{X^{1/1000}}^{X/H} \left| \sum_{r \sim R} \frac{a_r}{r^{1+it}} \right|^4 \right)^{1/2}$$

$$\cdot \left(\int_{X^{1/1000}}^{X/H} \left| \sum_{m \sim X/(P_1 R)} \frac{1}{m^{1+it}} \right|^4 |P_1(1+it)|^{2k} \, dt \right)^{1/2}.$$

- The claim follows from MVT and Deshouillers-Iwaniec MVT.

Type II estimate

- Recall $\mathcal{U} := \{t \in [X^{1/1000}, X/H] \colon |P_1(1+it)| > P_1^{-\varepsilon}\}$. Need to show, for every $Q \in [X^{\varepsilon^2}, z]$,

$$\int_{\mathcal{U}} \left| \sum_{r \sim X/(P_1 Q)} \frac{a_r}{r^{1+it}} \right|^2 \left| \sum_{q \sim Q} \frac{1_{q \in \mathbb{P}}}{q^{1+it}} \right|^2 dt \ll \frac{1}{(\log X)^{100}}$$

Type II estimate

- Recall $\mathcal{U} := \{t \in [X^{1/1000}, X/H] \colon |P_1(1+it)| > P_1^{-\varepsilon}\}$. Need to show, for every $Q \in [X^{\varepsilon^2}, z]$,

$$\int_{\mathcal{U}} \left| \sum_{r \sim X/(P_1 Q)} \frac{a_r}{r^{1+it}} \right|^2 \left| \sum_{q \sim Q} \frac{1_{q \in \mathbb{P}}}{q^{1+it}} \right|^2 dt \ll \frac{1}{(\log X)^{100}}$$

- Further split $\mathcal{U}$ into $\ll \log X$ sets $\mathcal{U}_\sigma$, where

$$\mathcal{U}_\sigma := \left\{ t \in \mathcal{U} \colon \left| \sum_{q \sim Q} \frac{1_{q \in \mathbb{P}}}{q^{1+it}} \right| \in (Q^{-\sigma}, 2Q^{-\sigma}] \right\}.$$

Type II estimate

- Recall $\mathcal{U} := \{t \in [X^{1/1000}, X/H] \colon |P_1(1+it)| > P_1^{-\varepsilon}\}$. Need to show, for every $Q \in [X^{\varepsilon^2}, z]$,

$$\int_{\mathcal{U}} \left| \sum_{r \sim X/(P_1 Q)} \frac{a_r}{r^{1+it}} \right|^2 \left| \sum_{q \sim Q} \frac{1_{q \in \mathbb{P}}}{q^{1+it}} \right|^2 dt \ll \frac{1}{(\log X)^{100}}$$

- Further split $\mathcal{U}$ into $\ll \log X$ sets $\mathcal{U}_\sigma$, where

$$\mathcal{U}_\sigma := \left\{ t \in \mathcal{U} \colon \left| \sum_{q \sim Q} \frac{1_{q \in \mathbb{P}}}{q^{1+it}} \right| \in (Q^{-\sigma}, 2Q^{-\sigma}] \right\}.$$

- For $\sigma \leq 0.237$, apply Jutila's large value estimate.

Type II estimate

- Recall $\mathcal{U} := \{t \in [X^{1/1000}, X/H] \colon |P_1(1+it)| > P_1^{-\varepsilon}\}$. Need to show, for every $Q \in [X^{\varepsilon^2}, z]$,

$$\int_{\mathcal{U}} \left| \sum_{r \sim X/(P_1 Q)} \frac{a_r}{r^{1+it}} \right|^2 \left| \sum_{q \sim Q} \frac{1_{q \in \mathbb{P}}}{q^{1+it}} \right|^2 dt \ll \frac{1}{(\log X)^{100}}$$

- Further split $\mathcal{U}$ into $\ll \log X$ sets $\mathcal{U}_\sigma$, where

$$\mathcal{U}_\sigma := \left\{ t \in \mathcal{U} \colon \left| \sum_{q \sim Q} \frac{1_{q \in \mathbb{P}}}{q^{1+it}} \right| \in (Q^{-\sigma}, 2Q^{-\sigma}] \right\}.$$

- For $\sigma \leq 0.237$, apply Jutila's large value estimate.
- Can assume $\sigma > 0.237$ and the claim reduces to

$$\int_{\mathcal{U}} \left| \sum_{r \sim X/(P_1 Q)} \frac{a_r}{r^{1+it}} \right|^2 \ll \frac{Q^{2 \cdot 0.237}}{(\log X)^{101}}.$$

Type II estimate

- Suffices to show that

$$\int_{\mathcal{U}} \left| \sum_{r \sim X/(P_1 Q)} \frac{a_r}{r^{1+it}} \right|^2 \ll \frac{Q^{0.474}}{(\log X)^{101}}.$$

Type II estimate

- Suffices to show that

$$\int_{\mathcal{U}} \left| \sum_{r \sim X/(P_1 Q)} \frac{a_r}{r^{1+it}} \right|^2 \ll \frac{Q^{0.474}}{(\log X)^{101}}.$$

- Choose k so that $P_1^k = X^{1-o(1)}$. By $1_{t \in \mathcal{U}} \leq |P(1+it)|^{2k} P^{2\varepsilon k}$, the integral above is

$$\ll P_1^{2\varepsilon k} \int_{\mathcal{U}} \left| \sum_{r \sim X/(P_1 Q)} \frac{a_r}{r^{1+it}} \right|^2 |P_1(1+it)|^{2k} dt.$$

Type II estimate

- Suffices to show that
$$\int_{\mathcal{U}} \left| \sum_{r \sim X/(P_1 Q)} \frac{a_r}{r^{1+it}} \right|^2 \ll \frac{Q^{0.474}}{(\log X)^{101}}.$$
- Choose k so that $P_1^k = X^{1-o(1)}$. By $1_{t \in \mathcal{U}} \leq |P(1+it)|^{2k} P^{2\varepsilon k}$, the integral above is
$$\ll P_1^{2\varepsilon k} \int_{\mathcal{U}} \left| \sum_{r \sim X/(P_1 Q)} \frac{a_r}{r^{1+it}} \right|^2 |P_1(1+it)|^{2k} dt.$$
- Coefficients of $P_1(s)^k$ are supported on $P_1 = (\log X)^{1.11}$-smooth numbers, so they have a very sparse support (of size $X^{1-1/1.1+o(1)}$).

Type II estimate

- Suffices to show that

$$\int_{\mathcal{U}} \left| \sum_{r \sim X/(P_1 Q)} \frac{a_r}{r^{1+it}} \right|^2 \ll \frac{Q^{0.474}}{(\log X)^{101}}.$$

- Choose k so that $P_1^k = X^{1-o(1)}$. By $1_{t \in \mathcal{U}} \leq |P(1+it)|^{2k} P^{2\varepsilon k}$, the integral above is

$$\ll P_1^{2\varepsilon k} \int_{\mathcal{U}} \left| \sum_{r \sim X/(P_1 Q)} \frac{a_r}{r^{1+it}} \right|^2 |P_1(1+it)|^{2k} dt.$$

- Coefficients of $P_1(s)^k$ are supported on $P_1 = (\log X)^{1.11}$-smooth numbers, so they have a very sparse support (of size $X^{1-1/1.1+o(1)}$).
- Invoke Heath-Brown's mean value theorem for sparse Dirichlet polynomials

Outline

1. Background and results
   - Primes in short intervals
   - Primes in almost all short intervals
   - Almost primes in (almost all) short intervals
2. Methods
   - Harman's sieve
   - Reductions
   - Type I and II estimates
3. Summary and more theorems

Theorem (M.-Teräväinen (202?))

The interval $(x-(\log X)^{2.1}, x]$ *contains* E_2*-numbers for almost all* $x \in [X/2, X]$.

E_2 numbers in almost all very short intervals

Theorem (M.-Teräväinen (202?))

The interval $(x - (\log X)^{2.1}, x]$ *contains* E_2*-numbers for almost all* $x \in [X/2, X]$.

- It is not difficult to adapt the argument to show that under Lindelöf one gets down to $2 + \varepsilon$. 2 is also the limit under RH.

E_2 numbers in almost all very short intervals

Theorem (M.-Teräväinen (202?))

The interval $(x - (\log X)^{2.1}, x]$ *contains* E_2*-numbers for almost all* $x \in [X/2, X]$.

- It is not difficult to adapt the argument to show that under Lindelöf one gets down to $2 + \varepsilon$. 2 is also the limit under RH.
- Actually, with some work, density hypothesis seems to suffice for $2 + \varepsilon$. But this is another story.

Sketch of the proof

- We study $p_1 p_2 \in (x - (\log X)^{2.11}, x]$ with $p_1 \sim P_1 := (\log X)^{1.11}$, i.e. one of the primes is very small.

Sketch of the proof

- We study $p_1 p_2 \in (x - (\log X)^{2.11}, x]$ with $p_1 \sim P_1 := (\log X)^{1.11}$, i.e. one of the primes is very small.
- For p_2, use Harman's sieve to find a suitable minorant $\rho^-(n) \leq 1_{n \in \mathbb{P}}$ and reduce to studying, for $H = (\log X)^{2.11}$,

$$\int_{X^{1/1000}}^{X/H} \left| \sum_{p_1 \sim P_1} \frac{1}{p_1^{1+it}} \sum_{n \sim X/P_1} \frac{\rho^-(n)}{n^{1+it}} \right|^2 dt.$$

Sketch of the proof

- We study $p_1 p_2 \in (x - (\log X)^{2.11}, x]$ with $p_1 \sim P_1 := (\log X)^{1.11}$, i.e. one of the primes is very small.
- For p_2, use Harman's sieve to find a suitable minorant $\rho^-(n) \leq 1_{n \in \mathbb{P}}$ and reduce to studying, for $H = (\log X)^{2.11}$,

$$\int_{X^{1/1000}}^{X/H} \left| \sum_{p_1 \sim P_1} \frac{1}{p_1^{1+it}} \sum_{n \sim X/P_1} \frac{\rho^-(n)}{n^{1+it}} \right|^2 dt.$$

- If $|\sum_{p_1 \sim P_1} p_1^{-1-it}| \leq P^{-\varepsilon}$, easy.
- Otherwise, decompose $\rho^-(n)$ as appropriate type I and type II sums.

Sketch of the proof

- We study $p_1 p_2 \in (x - (\log X)^{2.11}, x]$ with $p_1 \sim P_1 := (\log X)^{1.11}$, i.e. one of the primes is very small.
- For p_2, use Harman's sieve to find a suitable minorant $\rho^-(n) \leq 1_{n \in \mathbb{P}}$ and reduce to studying, for $H = (\log X)^{2.11}$,

$$\int_{X^{1/1000}}^{X/H} \left| \sum_{p_1 \sim P_1} \frac{1}{p_1^{1+it}} \sum_{n \sim X/P_1} \frac{\rho^-(n)}{n^{1+it}} \right|^2 dt.$$

- If $|\sum_{p_1 \sim P_1} p_1^{-1-it}| \leq P^{-\varepsilon}$, easy.
- Otherwise, decompose $\rho^-(n)$ as appropriate type I and type II sums.
- In each case we amplify by $P^{k\varepsilon} |\sum_{p_1 \sim P_1} p_1^{-1-it}|^k \geq 1$.

Sketch of the proof

- We study $p_1 p_2 \in (x - (\log X)^{2.11}, x]$ with $p_1 \sim P_1 := (\log X)^{1.11}$, i.e. one of the primes is very small.
- For p_2, use Harman's sieve to find a suitable minorant $\rho^-(n) \leq 1_{n \in \mathbb{P}}$ and reduce to studying, for $H = (\log X)^{2.11}$,

$$\int_{X^{1/1000}}^{X/H} \left| \sum_{p_1 \sim P_1} \frac{1}{p_1^{1+it}} \sum_{n \sim X/P_1} \frac{\rho^-(n)}{n^{1+it}} \right|^2 dt.$$

- If $|\sum_{p_1 \sim P_1} p_1^{-1-it}| \leq P^{-\varepsilon}$, easy.
- Otherwise, decompose $\rho^-(n)$ as appropriate type I and type II sums.
- In each case we amplify by $P^{k\varepsilon} |\sum_{p_1 \sim P_1} p_1^{-1-it}|^k \geq 1$.
- We use mean value theorem of Deshouillers-Iwaniec for type I sums.
- For type II sums we use large value theorems and Heath-Brown's recent sparse mean value theorem.

Some more theorems

The Dirichlet polynomial estimate concerning E_2 numbers in almost all short intervals also gives

Theorem (M.-Teräväinen (202?))

The interval $(x - \sqrt{x}(\log x)^{1.55}, x]$ *contains* E_3 *numbers for every large* x*.*

Some more theorems

The Dirichlet polynomial estimate concerning E_2 numbers in almost all short intervals also gives

Theorem (M.-Teräväinen (202?))

The interval $(x - \sqrt{x}(\log x)^{1.55}, x]$ *contains* E_3 *numbers for every large* x.

Our earlier work gave an asymptotic formula for E_2 numbers in all intervals:

Theorem (M.-Teräväinen (202?))

$$\sum_{\substack{x < p_1 p_2 \leq x+H \\ p_j \in \mathbb{P}}} 1 = H \frac{\log \log x}{\log x} + O\left(H \frac{\log \log \log x}{\log x}\right), \quad H \geq x^{0.55+\varepsilon}.$$

Thank you!