# Multiplicative functions in short intervals revisited

#### Kaisa Matomäki (Joint work with M. Radziwiłł)

University of Turku, Finland

#### Quebec-Maine Number Theory Conference 26th September 2020

We say that  $f : \mathbb{N} \to \mathbb{C}$  is multiplicative if

f(mn) = f(m)f(n) whenever gcd(m, n) = 1.

We shall concentrate on  $f : \mathbb{N} \to [-1, 1]$ .

We say that  $f \colon \mathbb{N} \to \mathbb{C}$  is multiplicative if

$$f(mn) = f(m)f(n)$$
 whenever  $gcd(m, n) = 1$ .

We shall concentrate on  $f \colon \mathbb{N} \to [-1, 1]$ . Some examples:

The Möbius function

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n = p_1 \cdots p_k \text{ with } p_j \text{ distinct;} \\ 0 & \text{otherwise.} \end{cases}$$

We say that  $f \colon \mathbb{N} \to \mathbb{C}$  is multiplicative if

$$f(mn) = f(m)f(n)$$
 whenever  $gcd(m, n) = 1$ .

We shall concentrate on  $f \colon \mathbb{N} \to [-1, 1]$ . Some examples:

The Möbius function

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n = p_1 \cdots p_k \text{ with } p_j \text{ distinct;} \\ 0 & \text{otherwise.} \end{cases}$$

 The indicator function for the set N of numbers that can be written as a sum of two squares;

$$\mathbf{1}_{\mathcal{N}}(p^k) = egin{cases} 0 & ext{if } p \equiv 3 \pmod{4} ext{ and } k ext{ is odd}; \ 1 & ext{otherwise}. \end{cases}$$

We say that  $f \colon \mathbb{N} \to \mathbb{C}$  is multiplicative if

$$f(mn) = f(m)f(n)$$
 whenever  $gcd(m, n) = 1$ .

We shall concentrate on  $f \colon \mathbb{N} \to [-1, 1]$ . Some examples:

• The Möbius function

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n = p_1 \cdots p_k \text{ with } p_j \text{ distinct;} \\ 0 & \text{otherwise.} \end{cases}$$

 The indicator function for the set N of numbers that can be written as a sum of two squares;

$$\mathbf{1}_{\mathcal{N}}(p^k) = egin{cases} 0 & ext{if } p \equiv 3 \pmod{4} ext{ and } k ext{ is odd}; \ 1 & ext{otherwise}. \end{cases}$$

The indicator function of the set of *y*-smooth numbers (*n* is *y*-smooth if *p* | *n* ⇒ *p* ≤ *y*).

### Averages over $n \leq x$

- Averages of multiplicative functions  $f : \mathbb{N} \to [-1, 1]$  over  $n \leq x$  are well understood (at least qualitatively):
- The mean value is o(1) if f does not pretend to be 1 and otherwise the mean value is  $\neq 0$  and can be calculated:

#### Averages over $n \leq x$

- Averages of multiplicative functions f: N → [-1,1] over n ≤ x are well understood (at least qualitatively):
- The mean value is o(1) if f does not pretend to be 1 and otherwise the mean value is  $\neq 0$  and can be calculated:
- Delange: If

$$\sum_{p} \frac{1-f(p)}{p} < \infty, \tag{1}$$

then

$$\frac{1}{x}\sum_{n\leq x}f(n) = (1+o(1))\prod_{p\leq x}\left(1-\frac{1}{p}\right)\left(1+\frac{f(p)}{p}+\frac{f(p^2)}{p^2}+\dots\right)$$

### Averages over $n \leq x$

- Averages of multiplicative functions  $f : \mathbb{N} \to [-1, 1]$  over  $n \leq x$  are well understood (at least qualitatively):
- The mean value is o(1) if f does not pretend to be 1 and otherwise the mean value is  $\neq 0$  and can be calculated:
- Delange: If

$$\sum_{p} \frac{1-f(p)}{p} < \infty, \tag{1}$$

then

$$\frac{1}{x}\sum_{n\leq x}f(n) = (1+o(1))\prod_{p\leq x}\left(1-\frac{1}{p}\right)\left(1+\frac{f(p)}{p}+\frac{f(p^2)}{p^2}+\dots\right)$$

• Wirsing: If (1) does not hold, then

$$\frac{1}{x}\sum_{n\leq x} f(n) = o(1);$$
 e.g.  $\frac{1}{x}\sum_{n\leq x} \mu(n) = o(1).$ 

• Halász's theorem gives quantitative results.

i

• Radziwiłł and I have shown that the same story holds in almost all short intervals: For  $f : \mathbb{N} \to [-1, 1]$ , one has

$$\left|\frac{1}{h}\sum_{x < n \le x+h} f(n) - \frac{1}{X}\sum_{X < n \le 2X} f(n)\right| = O\left((\log h)^{-1/200}\right)$$

• Radziwiłł and I have shown that the same story holds in almost all short intervals: For  $f : \mathbb{N} \to [-1, 1]$ , one has

$$\left|\frac{1}{h}\sum_{x < n \le x+h} f(n) - \frac{1}{X}\sum_{X < n \le 2X} f(n)\right| = O\left((\log h)^{-1/200}\right)$$

- This work has led into numerous applications and developments, including Tao's resolution of logarithmically averaged Chowla conjecture and Erdős discrepancy problem.
- But it has shortcomings that we now address:

• Radziwiłł and I have shown that the same story holds in almost all short intervals: For  $f : \mathbb{N} \to [-1, 1]$ , one has

$$\left|\frac{1}{h}\sum_{x < n \le x+h} f(n) - \frac{1}{X}\sum_{X < n \le 2X} f(n)\right| = O\left((\log h)^{-1/200}\right)$$

- This work has led into numerous applications and developments, including Tao's resolution of logarithmically averaged Chowla conjecture and Erdős discrepancy problem.
- But it has shortcomings that we now address:
- The quantitative bounds are pretty weak (in the paper slightly better than in this simplified form)

• Radziwiłł and I have shown that the same story holds in almost all short intervals: For  $f : \mathbb{N} \to [-1, 1]$ , one has

$$\left|\frac{1}{h}\sum_{x < n \le x+h} f(n) - \frac{1}{X}\sum_{X < n \le 2X} f(n)\right| = O\left((\log h)^{-1/200}\right)$$

- This work has led into numerous applications and developments, including Tao's resolution of logarithmically averaged Chowla conjecture and Erdős discrepancy problem.
- But it has shortcomings that we now address:
- The quantitative bounds are pretty weak (in the paper slightly better than in this simplified form)
- The theorem is trivial e.g. for 1<sub>n∈N</sub> (the indicator function of sums of two squares) since the long average is C(log X)<sup>-1/2</sup>.

• Radziwiłł and I have shown that the same story holds in almost all short intervals: For  $f : \mathbb{N} \to [-1, 1]$ , one has

$$\left|\frac{1}{h}\sum_{x < n \le x+h} f(n) - \frac{1}{X}\sum_{X < n \le 2X} f(n)\right| = O\left((\log h)^{-1/200}\right)$$

- This work has led into numerous applications and developments, including Tao's resolution of logarithmically averaged Chowla conjecture and Erdős discrepancy problem.
- But it has shortcomings that we now address:
- The quantitative bounds are pretty weak (in the paper slightly better than in this simplified form)
- The theorem is trivial e.g. for 1<sub>n∈N</sub> (the indicator function of sums of two squares) since the long average is C(log X)<sup>-1/2</sup>.
- For many applications, one needs a result for complex f.

 $\bullet$  Recall  ${\cal N}$  is the set of numbers that can be written as a sum of two squares. Then

$$\frac{1}{x}\sum_{n\leq x} 1_{\mathcal{N}}(n) = (C+o(1))\frac{1}{(\log x)^{1/2}}.$$

• This means that the average gap between consecutive  $m, n \in \mathcal{N} \cap [X, 2X]$  is  $\asymp (\log X)^{1/2} =: h_1$ .

 $\bullet$  Recall  ${\cal N}$  is the set of numbers that can be written as a sum of two squares. Then

$$\frac{1}{x}\sum_{n\leq x} 1_{\mathcal{N}}(n) = (C+o(1))\frac{1}{(\log x)^{1/2}}.$$

- This means that the average gap between consecutive  $m, n \in \mathcal{N} \cap [X, 2X]$  is  $\asymp (\log X)^{1/2} =: h_1$ .
- In particular in intervals of length  $y = o(h_1)$ , typically  $\sum_{x < n \le x+y} 1_{\mathcal{N}}(n) = 0.$

 $\bullet~\mbox{Recall}~\mathcal{N}$  is the set of numbers that can be written as a sum of two squares. Then

$$\frac{1}{x}\sum_{n\leq x} 1_{\mathcal{N}}(n) = (C+o(1))\frac{1}{(\log x)^{1/2}}.$$

- This means that the average gap between consecutive  $m, n \in \mathcal{N} \cap [X, 2X]$  is  $\asymp (\log X)^{1/2} =: h_1$ .
- In particular in intervals of length  $y = o(h_1)$ , typically  $\sum_{x < n \le x+y} 1_{\mathcal{N}}(n) = 0.$
- But for longer intervals one would expect regular behaviour, i.e

$$\Big|\frac{1}{h_0 h_1} \sum_{x < n \le x + h_0 h_1} 1_{\mathcal{N}}(n) - \frac{C}{(\log X)^{1/2}}\Big| = o\left(\frac{1}{(\log X)^{1/2}}\right)$$

for almost all  $x \in (X, 2X]$  as soon as  $h_0 o \infty$  with  $x o \infty$ 

#### Theorem (M-Radziwiłł (202?))

For any 
$$\delta > 0$$
, and any  $h_0 \ge 1$ ,  
 $\Big| \frac{1}{h_0 (\log X)^{1/2}} \sum_{x < n \le x + h_0 (\log X)^{1/2}} 1_{\mathcal{N}}(n) - \frac{C}{(\log X)^{1/2}} \Big| \le \frac{\delta}{(\log X)^{1/2}}.$ 

for all but at most

$$O_{\delta}(Xh_0^{-c\delta^{12}})$$

integers  $x \in (X, 2X]$ , for some c > 0.

# Theorem (M-Radziwiłł (202?)) For any $\delta > 0$ , and any $h_0 \ge 1$ , $\left|\frac{1}{h_0(\log X)^{1/2}} \sum_{x < n \le x + h_0(\log X)^{1/2}} 1_{\mathcal{N}}(n) - \frac{C}{(\log X)^{1/2}}\right| \le \frac{\delta}{(\log X)^{1/2}}.$ for all but at most $O_{\delta}(Xh_0^{-c\delta^{12}})$ integers $x \in (X, 2X]$ , for some c > 0.

- Note that the exceptional set bound saves polynomially in  $h_0$ .
- Previously Hooley (1994) and Plaksin (1987, 1992) showed that, for almost all  $x \in (X, 2X]$  one has

$$\frac{1}{h_0(\log X)^{1/2}} \sum_{x < n \le x + h_0(\log X)^{1/2}} 1_{\mathcal{N}}(n) \asymp \frac{1}{(\log X)^{1/2}}.$$

### General vanishing case

• When |f(p)| has average value  $\alpha \in (0,1)$ , it is known that

$$\frac{1}{x}\sum_{n\leq x}|f(n)|\asymp\prod_{p\leq x}\left(1+\frac{|f(p)|-1}{p}\right)\asymp(\log x)^{\alpha-1}.$$

• Write 
$$h_1 := \prod_{p \leq x} \left(1 + \frac{|f(p)| - 1}{p}\right)^{-1} \asymp (\log x)^{1 - \alpha}.$$

### General vanishing case

• When |f(p)| has average value  $lpha \in (0,1)$ , it is known that

$$\frac{1}{x}\sum_{n\leq x}|f(n)|\asymp\prod_{p\leq x}\left(1+\frac{|f(p)|-1}{p}\right)\asymp(\log x)^{\alpha-1}.$$

• Write 
$$h_1 := \prod_{p \leq x} \left(1 + \frac{|f(p)| - 1}{p}\right)^{-1} \asymp (\log x)^{1 - \alpha}.$$

In case |f(n)| takes values in {0,1}, the average of f(n) in intervals of length o(h<sub>1</sub>) is typically 0.

### General vanishing case

• When |f(p)| has average value  $lpha \in (0,1)$ , it is known that

$$\frac{1}{x}\sum_{n\leq x}|f(n)|\asymp\prod_{p\leq x}\left(1+\frac{|f(p)|-1}{p}\right)\asymp(\log x)^{\alpha-1}.$$

• Write 
$$h_1 := \prod_{p \leq x} \left(1 + \frac{|f(p)| - 1}{p}\right)^{-1} \asymp (\log x)^{1 - \alpha}.$$

- In case |f(n)| takes values in {0,1}, the average of f(n) in intervals of length o(h<sub>1</sub>) is typically 0.
- But for longer intervals we expect regular behaviour, i.e

$$\left|\frac{1}{h_0h_1}\sum_{x< n\leq x+h_0h_1}f(n)-\frac{1}{X}\sum_{X< n\leq 2X}f(n)\right|$$
$$=o\left(\prod_{p\leq X}\left(1+\frac{|f(p)|-1}{p}\right)\right)$$

for almost all x as soon as  $h_0 o \infty$  with  $x o \infty$ 

Let  $\varepsilon > 0$ . Let  $f : \mathbb{N} \to [-1, 1]$  be a multiplicative function s.t.

$$\sum_{w$$

for all  $2 \le w < z < x^{\varepsilon}$ . Set  $h_1 := \prod_{p \le X} \left( 1 + \frac{1 - |f(p)|}{p} \right)$ .

Let  $\varepsilon > 0$ . Let  $f : \mathbb{N} \to [-1, 1]$  be a multiplicative function s.t.

$$\sum_{w$$

for all  $2 \le w < z < x^{\varepsilon}$ . Set  $h_1 := \prod_{p \le X} \left(1 + \frac{1 - |f(p)|}{p}\right)$ . For any  $\delta > 0$ , and any  $h_0 \ge 1$ ,

$$\left|\frac{1}{h_0 h_1} \sum_{x < n \le x + h_0 h_1} f(n) - \frac{1}{X} \sum_{X < n \le 2X} f(n)\right| \le \delta \prod_{p \le X} \left(1 + \frac{|f(p)| - 1}{p}\right)$$

for all but at most  $O(Xh_0^{-c\delta^{\kappa}})$  integers  $x \in (X, 2X]$ , for some  $c = c(\varepsilon)$  and  $\kappa = \kappa(\varepsilon) > 0$ .

Let  $\varepsilon > 0$ . Let  $f : \mathbb{N} \to [-1, 1]$  be a multiplicative function s.t.

$$\sum_{w$$

for all  $2 \le w < z < x^{\varepsilon}$ . Set  $h_1 := \prod_{p \le X} \left(1 + \frac{1 - |f(p)|}{p}\right)$ . For any  $\delta > 0$ , and any  $h_0 \ge 1$ ,

$$\left|\frac{1}{h_0 h_1} \sum_{x < n \le x + h_0 h_1} f(n) - \frac{1}{X} \sum_{X < n \le 2X} f(n)\right| \le \delta \prod_{p \le X} \left(1 + \frac{|f(p)| - 1}{p}\right)$$

for all but at most  $O(Xh_0^{-c\delta^{\kappa}})$  integers  $x \in (X, 2X]$ , for some  $c = c(\varepsilon)$  and  $\kappa = \kappa(\varepsilon) > 0$ .

If f is complex-valued, a twist in main term.

## Limitations of Hooley's and Plaksin's methods

• Recall Hooley's and Plaksin's works giving that for almost all x one has

$$\frac{1}{h_0(\log X)^{1/2}} \sum_{x < n \le x + h_0(\log X)^{1/2}} 1_{\mathcal{N}}(n) \asymp \frac{1}{(\log X)^{1/2}}.$$

### Limitations of Hooley's and Plaksin's methods

• Recall Hooley's and Plaksin's works giving that for almost all x one has

$$\frac{1}{h_0(\log X)^{1/2}} \sum_{x < n \le x + h_0(\log X)^{1/2}} 1_{\mathcal{N}}(n) \asymp \frac{1}{(\log X)^{1/2}}.$$

• The main arithmetic information they used was the solution to the shifted convolution problem

$$\sum_{n \le x} r_{\mathcal{K}}(n) r_{\mathcal{K}}(n+h) \tag{2}$$

with  $r_{\mathcal{K}}(n)$  the coefficients of the Dedekind zeta function of  $\mathcal{K} = \mathbb{Q}(i)$ .

# Limitations of Hooley's and Plaksin's methods

• Recall Hooley's and Plaksin's works giving that for almost all x one has

$$\frac{1}{h_0(\log X)^{1/2}} \sum_{x < n \le x + h_0(\log X)^{1/2}} 1_{\mathcal{N}}(n) \asymp \frac{1}{(\log X)^{1/2}}.$$

• The main arithmetic information they used was the solution to the shifted convolution problem

$$\sum_{n\leq x} r_{\mathcal{K}}(n) r_{\mathcal{K}}(n+h) \tag{2}$$

with  $r_{\mathcal{K}}(n)$  the coefficients of the Dedekind zeta function of  $\mathcal{K} = \mathbb{Q}(i)$ .

- (2) is completely open when degree of *K* exceeds two. Hence the previous approaches completely fail for generalisations.
- We only use multiplicativity, so we have chances to generalise!

We say an integer n is norm-form of a number field K if n equals a norm of an algebraic integer in K. Write g<sub>K</sub>(n) for the indicator function. In particular g<sub>Q(i)</sub>(n) = 1<sub>n∈N</sub>(n)

- We say an integer n is norm-form of a number field K if n equals a norm of an algebraic integer in K. Write g<sub>K</sub>(n) for the indicator function. In particular g<sub>Q(i)</sub>(n) = 1<sub>n∈N</sub>(n)
- By Odoni's work the density in [1, x] of norm forms of K is

$$\delta_{\mathcal{K}}(x) := \prod_{\substack{p \leq x, \, p \neq N \mathfrak{a} \\ \mathfrak{a} \text{ integral ideal}}} \left(1 - rac{1}{p}
ight)$$

If *K* is a normal extension of  $\mathbb{Q}$  of degree *k*, then  $\delta_{K}(x) \asymp (\log x)^{-1+1/k}$ .

- We say an integer n is norm-form of a number field K if n equals a norm of an algebraic integer in K. Write g<sub>K</sub>(n) for the indicator function. In particular g<sub>Q(i)</sub>(n) = 1<sub>n∈N</sub>(n)
- By Odoni's work the density in [1, x] of norm forms of K is

$$\delta_{\mathcal{K}}(x) := \prod_{\substack{p \leq x, \, p \neq N \mathfrak{a} \\ \mathfrak{a} \text{ integral ideal}}} \left(1 - rac{1}{p}
ight)$$

If *K* is a normal extension of  $\mathbb{Q}$  of degree *k*, then  $\delta_{K}(x) \asymp (\log x)^{-1+1/k}$ .

 Issue: When the class number of K exceeds 1, g<sub>K</sub>(n) is not multiplicative.

- We say an integer n is norm-form of a number field K if n equals a norm of an algebraic integer in K. Write g<sub>K</sub>(n) for the indicator function. In particular g<sub>Q(i)</sub>(n) = 1<sub>n∈N</sub>(n)
- By Odoni's work the density in [1, x] of norm forms of K is

$$\delta_{\mathcal{K}}(x) := \prod_{\substack{p \leq x, \, p \neq N \mathfrak{a} \\ \mathfrak{a} \text{ integral ideal}}} \left(1 - rac{1}{p}
ight)$$

If *K* is a normal extension of  $\mathbb{Q}$  of degree *k*, then  $\delta_{K}(x) \asymp (\log x)^{-1+1/k}$ .

- Issue: When the class number of K exceeds 1, g<sub>K</sub>(n) is not multiplicative.
- However, following work of Odoni, we show that g<sub>K</sub>(n) is a linear combination of (complex-valued) multiplicative functions.

- We say an integer n is norm-form of a number field K if n equals a norm of an algebraic integer in K. Write g<sub>K</sub>(n) for the indicator function. In particular g<sub>Q(i)</sub>(n) = 1<sub>n∈N</sub>(n)
- By Odoni's work the density in [1, x] of norm forms of K is

$$\delta_{\mathcal{K}}(x) := \prod_{\substack{p \leq x, \, p \neq N \mathfrak{a} \\ \mathfrak{a} \text{ integral ideal}}} \left(1 - rac{1}{p}
ight)$$

If *K* is a normal extension of  $\mathbb{Q}$  of degree *k*, then  $\delta_{K}(x) \asymp (\log x)^{-1+1/k}$ .

- Issue: When the class number of K exceeds 1, g<sub>K</sub>(n) is not multiplicative.
- However, following work of Odoni, we show that g<sub>K</sub>(n) is a linear combination of (complex-valued) multiplicative functions.
- Applying our results to each function in the linear combination, we get a theorem in arbitrary number fields.

Let K be a number field over  $\mathbb{Q}$ , and let

$$\delta_{\mathcal{K}}(x) := \prod_{\substack{p \leq x, \ p \neq N\mathfrak{a} \\ \mathfrak{a} \ integral \ ideal}} \left(1 - \frac{1}{p}\right).$$

Then, as  $X \to \infty$ , uniformly in  $2 \le h \le X$  one has

$$\left|\frac{1}{h\delta_{\mathcal{K}}(X)^{-1}}\sum_{x\leq n\leq x+h\delta_{\mathcal{K}}(X)^{-1}}g_{\mathcal{K}}(n)-C_{\mathcal{K}}\delta_{\mathcal{K}}(X)\right|\leq \varepsilon\delta_{\mathcal{K}}(X)$$

for all  $x \in (X, 2X]$  with at most  $O_{\varepsilon}(Xh^{-c\varepsilon^{\kappa}})$  exceptions where  $c, \kappa, C_{K} > 0$  depend solely on K.

Let K be a number field over  $\mathbb{Q}$ , and let

$$\delta_{\mathcal{K}}(x) := \prod_{\substack{p \leq x, \ p \neq N\mathfrak{a} \\ \mathfrak{a} \ integral \ ideal}} \left(1 - \frac{1}{p}\right).$$

Then, as  $X \to \infty$ , uniformly in  $2 \le h \le X$  one has

$$\left|\frac{1}{h\delta_{\mathcal{K}}(X)^{-1}}\sum_{x\leq n\leq x+h\delta_{\mathcal{K}}(X)^{-1}}g_{\mathcal{K}}(n)-C_{\mathcal{K}}\delta_{\mathcal{K}}(X)\right|\leq \varepsilon\delta_{\mathcal{K}}(X)$$

for all  $x \in (X, 2X]$  with at most  $O_{\varepsilon}(Xh^{-c\varepsilon^{\kappa}})$  exceptions where  $c, \kappa, C_{\kappa} > 0$  depend solely on K.

This is a vast extension of Hooley's and Plaksin's works for  $\mathbb{Q}(i)$ , with an asymptotic formula.

- Hooley (1971) and Plaksin (1987, 1992) have also studied gaps between sums of two squares.
- Writing  $1 = s_1 < s_2 < \ldots$  for the sequence of integers in  $\mathcal{N}$ . Plaksin showed that, for any  $\gamma \in [1, 2)$  (Hooley:  $\gamma \in [1, 5/3)$ ),

$$\sum_{s_n \leq x} (s_{n+1} - s_n)^{\gamma} \asymp x (\log x)^{\frac{1}{2}(\gamma - 1)}$$

- Hooley (1971) and Plaksin (1987, 1992) have also studied gaps between sums of two squares.
- Writing  $1 = s_1 < s_2 < \ldots$  for the sequence of integers in  $\mathcal{N}$ . Plaksin showed that, for any  $\gamma \in [1, 2)$  (Hooley:  $\gamma \in [1, 5/3)$ ),

$$\sum_{s_n \leq x} (s_{n+1} - s_n)^{\gamma} \asymp x (\log x)^{\frac{1}{2}(\gamma-1)}$$

• That is to say, for any  $h \ge 1$ , the number of  $x \in [X, 2X]$  for which

$$(x, x + h(\log X)^{1/2}] \cap \mathcal{N} = \emptyset$$

is at most  $O(Xh^{-1+\varepsilon})$ .

- Hooley (1971) and Plaksin (1987, 1992) have also studied gaps between sums of two squares.
- Writing  $1 = s_1 < s_2 < \ldots$  for the sequence of integers in  $\mathcal{N}$ . Plaksin showed that, for any  $\gamma \in [1, 2)$  (Hooley:  $\gamma \in [1, 5/3)$ ),

$$\sum_{s_n \leq x} (s_{n+1} - s_n)^{\gamma} \asymp x (\log x)^{\frac{1}{2}(\gamma - 1)}$$

• That is to say, for any  $h \ge 1$ , the number of  $x \in [X, 2X]$  for which

$$(x, x + h(\log X)^{1/2}] \cap \mathcal{N} = \emptyset$$

is at most  $O(Xh^{-1+\varepsilon})$ .

• Again this is based on the shifted convolution problem for  $r_{\kappa}(n)$  and does not extend beyond quadratic number fields.

- Hooley (1971) and Plaksin (1987, 1992) have also studied gaps between sums of two squares.
- Writing  $1 = s_1 < s_2 < \ldots$  for the sequence of integers in  $\mathcal{N}$ . Plaksin showed that, for any  $\gamma \in [1, 2)$  (Hooley:  $\gamma \in [1, 5/3)$ ),

$$\sum_{s_n \leq x} (s_{n+1} - s_n)^{\gamma} \asymp x (\log x)^{\frac{1}{2}(\gamma - 1)}$$

• That is to say, for any  $h \ge 1$ , the number of  $x \in [X, 2X]$  for which

$$(x, x + h(\log X)^{1/2}] \cap \mathcal{N} = \emptyset$$

is at most  $O(Xh^{-1+\varepsilon})$ .

- Again this is based on the shifted convolution problem for  $r_{\kappa}(n)$  and does not extend beyond quadratic number fields.
- If, like Hooley and Plaksin, we do not request asymptotic formula, we get an improved bound for our exceptional set for any *K*.

Let K be a number field and let  $\delta_{K}(x)$  be the density of norm-forms. Then, for any  $\varepsilon > 0$ , there exists a constant  $c = c(K, \varepsilon)$  such that, for any  $h \ge 1$ , one has

$$rac{1}{h\delta_{K}(x)^{-1}}\sum_{x < n \leq x+h\delta_{K}(x)^{-1}}g_{K}(n) \geq c\delta_{K}(x)$$

for all but  $O_{\varepsilon,K}(Xh^{-1/2+\varepsilon})$  of  $x \in (X, 2X]$ .

Let K be a number field and let  $\delta_K(x)$  be the density of norm-forms. Then, for any  $\varepsilon > 0$ , there exists a constant  $c = c(K, \varepsilon)$  such that, for any  $h \ge 1$ , one has

$$\frac{1}{h\delta_{\mathcal{K}}(x)^{-1}}\sum_{x < n \leq x + h\delta_{\mathcal{K}}(x)^{-1}} g_{\mathcal{K}}(n) \geq c\delta_{\mathcal{K}}(x)$$

for all but  $O_{\varepsilon,K}(Xh^{-1/2+\varepsilon})$  of  $x \in (X, 2X]$ . Consequently, letting  $1 \le n_1 < n_2 < \ldots$  denote the sequence of positive norm-forms of K, one has for any  $\gamma \in [1, 3/2)$ ,

$$\sum_{n_i\leq x}(n_{i+1}-n_i)^{\gamma}\asymp_{\gamma,K}x\delta_K(x)^{\gamma-1}.$$

Let K be a number field and let  $\delta_K(x)$  be the density of norm-forms. Then, for any  $\varepsilon > 0$ , there exists a constant  $c = c(K, \varepsilon)$  such that, for any  $h \ge 1$ , one has

$$\frac{1}{h\delta_{\mathcal{K}}(x)^{-1}}\sum_{x < n \leq x + h\delta_{\mathcal{K}}(x)^{-1}} g_{\mathcal{K}}(n) \geq c\delta_{\mathcal{K}}(x)$$

for all but  $O_{\varepsilon,K}(Xh^{-1/2+\varepsilon})$  of  $x \in (X, 2X]$ . Consequently, letting  $1 \le n_1 < n_2 < \ldots$  denote the sequence of positive norm-forms of K, one has for any  $\gamma \in [1, 3/2)$ ,

$$\sum_{n_i\leq x}(n_{i+1}-n_i)^{\gamma}\asymp_{\gamma,K}x\delta_K(x)^{\gamma-1}.$$

This vastly extends Hooley's and Plaksin's results (case  $K = \mathbb{Q}(i)$  with  $\gamma \in [1, 5/3)$  and  $\gamma \in [1, 2)$  repsectively).

Also these results work more generally. E.g. we get

#### Corollary

Let  $\varepsilon > 0$  be given and  $h \ge 1$ . Then the number of intervals [x, x + h] with  $x \in [X, 2X]$  that do not contain an  $x^{\varepsilon}$ -smooth number is  $\ll_{\eta,\varepsilon} Xh^{-1/2+\eta}$  for all  $\eta > 0$ .

Consequently, letting  $1 \le n_1 < n_2 < \ldots$  denote the sequence of integers n such that all prime factors of n are  $\le n^{\varepsilon}$ , one has, for any  $\gamma \in [1, 3/2)$ ,

$$\sum_{n_i \leq x} (n_{i+1} - n_i)^{\gamma} \asymp_{\varepsilon, \gamma} x \tag{3}$$

Also these results work more generally. E.g. we get

#### Corollary

Let  $\varepsilon > 0$  be given and  $h \ge 1$ . Then the number of intervals [x, x + h] with  $x \in [X, 2X]$  that do not contain an  $x^{\varepsilon}$ -smooth number is  $\ll_{\eta,\varepsilon} Xh^{-1/2+\eta}$  for all  $\eta > 0$ .

Consequently, letting  $1 \le n_1 < n_2 < \ldots$  denote the sequence of integers n such that all prime factors of n are  $\le n^{\varepsilon}$ , one has, for any  $\gamma \in [1, 3/2)$ ,

$$\sum_{n_i \leq x} (n_{i+1} - n_i)^{\gamma} \asymp_{\varepsilon, \gamma} x \tag{3}$$

This improves on a recent result of Heath-Brown who got for (3) the upper bound  $\ll x^{1+\eta}$  for any  $\eta > 0$ .

Let  $\varepsilon > 0$ . Let  $f : \mathbb{N} \to [-1, 1]$  be a multiplicative function s.t.

$$\sum_{w$$

for all  $2 \le w < z < x^{\varepsilon}$ . Set  $h_1 := \prod_{p \le X} \left( 1 + \frac{1 - |f(p)|}{p} \right)$ .

Let  $\varepsilon > 0$ . Let  $f : \mathbb{N} \to [-1, 1]$  be a multiplicative function s.t.

$$\sum_{w$$

for all  $2 \le w < z < x^{\varepsilon}$ . Set  $h_1 := \prod_{p \le X} \left(1 + \frac{1 - |f(p)|}{p}\right)$ . For any  $\delta > 0$ , and any  $h_0 \ge 1$ ,

$$\left|\frac{1}{h_0 h_1} \sum_{x < n \le x + h_0 h_1} f(n) - \frac{1}{X} \sum_{X < n \le 2X} f(n)\right| \le \delta \prod_{p \le X} \left(1 + \frac{|f(p)| - 1}{p}\right)$$

for all but at most  $O(Xh_0^{-c\delta^{\kappa}})$  integers  $x \in (X, 2X]$ , for some  $c = c(\varepsilon)$  and  $\kappa = \kappa(\varepsilon) > 0$ .

# Proof ideas

- For simplicity concentrate on case when the average of f is 0.
- Our starting point is Perron's formula, giving

$$\sum_{x < n \le x+H} f(n) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} \cdot \frac{(x+H)^{1+it} - x^{1+it}}{1+it}$$
$$\approx \frac{H}{2\pi i} \int_{-X/H}^{X/H} \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} x^{it} dt.$$

# Proof ideas

- For simplicity concentrate on case when the average of f is 0.
- Our starting point is Perron's formula, giving

$$\sum_{x < n \le x+H} f(n) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} \cdot \frac{(x+H)^{1+it} - x^{1+it}}{1+it}$$
$$\approx \frac{H}{2\pi i} \int_{-X/H}^{X/H} \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} x^{it} dt.$$

If we just put absolute values here, we have an impossible task

 with squareroot cancellation RHS would be of size X<sup>1/2</sup>,
 i.e. much larger than the trivial bound H.

# Proof ideas

- For simplicity concentrate on case when the average of f is 0.
- Our starting point is Perron's formula, giving

$$\sum_{x < n \le x+H} f(n) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} \cdot \frac{(x+H)^{1+it} - x^{1+it}}{1+it}$$
$$\approx \frac{H}{2\pi i} \int_{-X/H}^{X/H} \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} x^{it} dt.$$

- If we just put absolute values here, we have an impossible task

   with squareroot cancellation RHS would be of size X<sup>1/2</sup>,
   i.e. much larger than the trivial bound H.
- Normally one would go on and average over x, getting

$$\frac{1}{X}\int_X^{2X}\left|\sum_{x< n\leq x+H}f(n)\right|^2dx\ll \int_{-X/H}^{X/H}\left|\sum_{X< n\leq 3X}\frac{f(n)}{n^{1+it}}\right|^2dt.$$

#### The mean square

• Normally one would go on and average over x, getting

$$\frac{1}{X}\int_X^{2X}\left|\sum_{x< n\leq x+H}f(n)\right|^2 dx \ll \int_{-X/H}^{X/H}\left|\sum_{X< n\leq 3X}\frac{f(n)}{n^{1+it}}\right|^2 dt.$$

#### The mean square

• Normally one would go on and average over x, getting

$$\frac{1}{X}\int_X^{2X}\left|\sum_{x< n\leq x+H}f(n)\right|^2 dx \ll \int_{-X/H}^{X/H}\left|\sum_{X< n\leq 3X}\frac{f(n)}{n^{1+it}}\right|^2 dt.$$

• Recall  $H = h_0 h_1$  with  $h_1 = \prod_{p \le X} (1 + \frac{1 - |f(p)|}{p})$ . Now to show that

$$\left|\frac{1}{h_0h_1}\sum_{x$$

with  $O(Xh_0^{-c\delta^{\kappa}})$  exceptions, we would need the bound

$$\int_{-X/H}^{X/H} \left| \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} \right|^2 dt \ll \delta^2 h_0^{-c\delta^{\kappa}} \prod_{p \le X} \left( 1 + \frac{|f(p)| - 1}{p} \right)^2$$

# Studying the mean square

• We would need the bound

$$\int_{-X/H}^{X/H} \left| \sum_{X < n \leq 3X} \frac{f(n)}{n^{1+it}} \right|^2 dt \ll \delta^2 h_0^{-c\delta^{\kappa}} \prod_{p \leq X} \left( 1 + \frac{|f(p)| - 1}{p} \right)^2.$$

• Using Shiu's and Henriot's bounds for averages and correlations of multiplicative functions, one can tweak the usual MVT to show that, for any  $a_n \leq |f(n)|$ ,

$$\int_{-X/H}^{X/H} \left| \sum_{X < n \leq 3X} \frac{a_n}{n^{1+it}} \right|^2 dt \ll \prod_{p \leq X} \left( 1 + \frac{|f(p)| - 1}{p} \right)^2.$$

# Studying the mean square

• We would need the bound

$$\int_{-X/H}^{X/H} \left| \sum_{X < n \leq 3X} \frac{f(n)}{n^{1+it}} \right|^2 dt \ll \delta^2 h_0^{-c\delta^{\kappa}} \prod_{p \leq X} \left( 1 + \frac{|f(p)| - 1}{p} \right)^2.$$

 Using Shiu's and Henriot's bounds for averages and correlations of multiplicative functions, one can tweak the usual MVT to show that, for any a<sub>n</sub> ≤ |f(n)|,

$$\int_{-X/H}^{X/H} \left| \sum_{X < n \leq 3X} \frac{a_n}{n^{1+it}} \right|^2 dt \ll \prod_{p \leq X} \left( 1 + \frac{|f(p)| - 1}{p} \right)^2.$$

- Same situation as in our previous work need to save something compared to the MVT bound.
- After reproving Halasz and Lipschitz type estimates in the sparse setting, we can repeat those arguments.

# Studying the mean square

• We would need the bound

$$\int_{-X/H}^{X/H} \left| \sum_{X < n \leq 3X} \frac{f(n)}{n^{1+it}} \right|^2 dt \ll \delta^2 h_0^{-c\delta^{\kappa}} \prod_{p \leq X} \left( 1 + \frac{|f(p)| - 1}{p} \right)^2.$$

 Using Shiu's and Henriot's bounds for averages and correlations of multiplicative functions, one can tweak the usual MVT to show that, for any a<sub>n</sub> ≤ |f(n)|,

$$\int_{-X/H}^{X/H} \left| \sum_{X < n \leq 3X} \frac{a_n}{n^{1+it}} \right|^2 dt \ll \prod_{p \leq X} \left( 1 + \frac{|f(p)| - 1}{p} \right)^2.$$

- Same situation as in our previous work need to save something compared to the MVT bound.
- After reproving Halasz and Lipschitz type estimates in the sparse setting, we can repeat those arguments.
- But this gives about  $h_0^{-c\delta^{\kappa}} + (\log X)^{-\kappa}$  where we want  $h_0^{-c\delta^{\kappa}}$

#### An issue with mean square

Actually showing the bound

$$\int_{-X/H}^{X/H} \left| \sum_{X < n \leq 3X} \frac{f(n)}{n^{1+it}} \right|^2 dt \ll \delta^2 h_0^{-c\delta^{\kappa}} \prod_{p \leq X} \left( 1 + \frac{|f(p)| - 1}{p} \right)^2.$$

in general is not possible — there might be some points twhere  $\sum f(n)n^{-1+it}$  is  $\asymp \prod_{p \le X} (1 + \frac{|f(p)| - 1}{p})(\log X)^{-\kappa}$ .

#### An issue with mean square

• Actually showing the bound

$$\int_{-X/H}^{X/H} \left| \sum_{X < n \leq 3X} \frac{f(n)}{n^{1+it}} \right|^2 dt \ll \delta^2 h_0^{-c\delta^{\kappa}} \prod_{p \leq X} \left( 1 + \frac{|f(p)| - 1}{p} \right)^2.$$

in general is not possible — there might be some points twhere  $\sum_{i=1}^{\infty} f(n) n^{-1+it}$  is  $\asymp \prod_{p \le X} (1 + \frac{|f(p)| - 1}{p}) (\log X)^{-\kappa}$ .

But if we had something like

$$\sum_{P$$

for all  $P = 2^j \in [X^{\varepsilon^3}, X^{\varepsilon^2}]$ , then that method would give the desired bound. (also we need to construct a good sieve majorant for f(n) to handle " $n \notin S$ ")

#### An issue with mean square

• Actually showing the bound

$$\int_{-X/H}^{X/H} \left| \sum_{X < n \leq 3X} \frac{f(n)}{n^{1+it}} \right|^2 dt \ll \delta^2 h_0^{-c\delta^{\kappa}} \prod_{p \leq X} \left( 1 + \frac{|f(p)| - 1}{p} \right)^2.$$

in general is not possible — there might be some points twhere  $\sum_{i=1}^{n} f(n) n^{-1+it}$  is  $\asymp \prod_{p \le X} (1 + \frac{|f(p)| - 1}{p}) (\log X)^{-\kappa}$ .

But if we had something like

$$\sum_{P$$

for all  $P = 2^{j} \in [X^{\varepsilon^{3}}, X^{\varepsilon^{2}}]$ , then that method would give the desired bound. (also we need to construct a good sieve majorant for f(n) to handle " $n \notin S$ ")

• Key new idea: Handle "exceptional" *t* before taking the mean square over *x*.

# Splitting into ${\mathcal T}$ and ${\mathcal U}$

Recall

$$\sum_{x < n \le x+H} f(n) \approx \frac{H}{2\pi i} \int_{-X/H}^{X/H} \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} x^{it} dt.$$

# Splitting into ${\mathcal T}$ and ${\mathcal U}$

Recall

$$\sum_{x < n \le x+H} f(n) \approx \frac{H}{2\pi i} \int_{-X/H}^{X/H} \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} x^{it} dt.$$

• Split  $[-X/H, X/H] = \mathcal{T} \cup \mathcal{U}$  with  $t \in \mathcal{T}$  iff

$$\left|\sum_{P$$

# Splitting into ${\mathcal T}$ and ${\mathcal U}$

Recall

$$\sum_{x < n \leq x+H} f(n) \approx \frac{H}{2\pi i} \int_{-X/H}^{X/H} \sum_{X < n \leq 3X} \frac{f(n)}{n^{1+it}} x^{it} dt.$$

• Split  $[-X/H, X/H] = \mathcal{T} \cup \mathcal{U}$  with  $t \in \mathcal{T}$  iff

$$\left|\sum_{P$$

• By MVT  $|\mathcal{U}| \leq (X/H)^{1/2-arepsilon}$ , and by previous discussion,

$$\frac{1}{X} \int_{X}^{2X} \left| \frac{H}{2\pi i} \int_{\mathcal{T}} \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} x^{it} dt \right|^{2} dx$$
$$\ll \int_{\mathcal{T}} \left| \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} \right|^{2} dt \ll \delta^{2} h_{0}^{-c\delta^{\kappa}} \prod_{p \le X} \left( 1 + \frac{|f(p)| - 1}{p} \right)^{2}.$$

# $\mathsf{Handling}\; \mathcal{U}$

• We are left with studying, for certain  $|\mathcal{U}| \leq (X/H)^{1/2-arepsilon}$ ,

$$\frac{H}{2\pi i} \int_{\mathcal{U}} \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} x^{it} dt.$$
(4)

# Handling $\mathcal{U}$

• We are left with studying, for certain  $|\mathcal{U}| \leq (X/H)^{1/2-arepsilon}$ ,

$$\frac{H}{2\pi i} \int_{\mathcal{U}} \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} x^{it} dt.$$
(4)

• Since most integers have at least two prime factors from  $(X^{\varepsilon^2}, X^{\varepsilon}]$ , we can at least morally replace  $\sum \frac{f(n)}{n^{1+it}}$  by

$$\sum_{P_1, P_2 \in (X^{\varepsilon^2}, X^{\varepsilon}]} \sum_{P_1$$

# Handling $\mathcal{U}$

• We are left with studying, for certain  $|\mathcal{U}| \leq (X/H)^{1/2-arepsilon}$  ,

$$\frac{H}{2\pi i} \int_{\mathcal{U}} \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} x^{it} dt.$$
(4)

• Since most integers have at least two prime factors from  $(X^{\varepsilon^2}, X^{\varepsilon}]$ , we can at least morally replace  $\sum \frac{f(n)}{n^{1+it}}$  by

$$\sum_{P_1, P_2 \in (X^{\varepsilon^2}, X^{\varepsilon}]} \sum_{P_1$$

 $\bullet\,$  Now, by Huxley's large value theorem, those  $t\in\mathcal{U}$  for which

$$\Big|\sum_{P$$

give an acceptable contribution to square mean of (4).

# Handling $\mathcal{U}$

• We are left with studying, for certain  $|\mathcal{U}| \leq (X/H)^{1/2-arepsilon}$  ,

$$\frac{H}{2\pi i} \int_{\mathcal{U}} \sum_{X < n \le 3X} \frac{f(n)}{n^{1+it}} x^{it} dt.$$
(4)

• Since most integers have at least two prime factors from  $(X^{\varepsilon^2}, X^{\varepsilon}]$ , we can at least morally replace  $\sum \frac{f(n)}{n^{1+it}}$  by

$$\sum_{P_1, P_2 \in (X^{\varepsilon^2}, X^{\varepsilon}]} \sum_{P_1$$

 $\bullet\,$  Now, by Huxley's large value theorem, those  $t\in\mathcal{U}$  for which

$$\Big|\sum_{P$$

give an acceptable contribution to square mean of (4).

• The complement is tiny and there (4) is o(H) by Halász + tailored Halász-Montgomery type large value results.

Kaisa Matomäki

Multiplicative functions in short intervals revisited

# The results with positive proportion lower bound

When one only wants, for f: N → [0, 1], with a good exceptional set,

$$\frac{1}{H}\sum_{x < n < x+H} f(n) \ge \delta \prod_{p \le x} \left(1 + \frac{f(p) - 1}{p}\right),$$

# The results with positive proportion lower bound

• When one only wants, for  $f : \mathbb{N} \to [0, 1]$ , with a good exceptional set,

$$\frac{1}{H}\sum_{x < n < x+H} f(n) \ge \delta \prod_{p \le x} \left(1 + \frac{f(p) - 1}{p}\right),$$

it suffices to show that, for  $K = \lfloor 1/\varepsilon^{10} \rfloor$ ,

$$\frac{1}{H}\sum_{\substack{x < p_1 \cdots p_{K-1}m \leq x+H\\ p_j \in [X^{(1-\varepsilon^{10})/K}, X^{(1+\varepsilon^{10})/K}]}} f(p_1) \cdots f(p_{K-1})f(m)$$
$$\gg \prod_{p \leq x} \left(1 + \frac{f(p) - 1}{p}\right).$$

- The resulting Dirichlet polynomial is a product of short factors. This gives a lot more flexibility with applying mean and large value theorems
- This way we get the desired result.

# Thank you!