The partial sum of a random multiplicative function on integers with a large prime factor

SSANT Paris

Daniele Mastrostefano

University of Warwick

02/07/2021

Random multiplicative functions

A Rademacher random multiplicative function $f(n)$ is defined as

$$f(n) := \prod_{p|n} f(p)$$

for all squarefree n and $f(n) = 0$ otherwise, where $\{f(p)\}_p$ are independent random variables taking values ± 1 with probability $1/2$ each. Note that the values of $f(n)$ are not all independent.

Random multiplicative functions

A Rademacher random multiplicative function $f(n)$ is defined as

$$f(n) := \prod_{p|n} f(p)$$

for all squarefree n and $f(n) = 0$ otherwise, where $\{f(p)\}_p$ are independent random variables taking values ± 1 with probability $1/2$ each. Note that the values of $f(n)$ are not all independent.
Wintner, in 1944, proved that

$$\sum_{n \leq x} f(n) = O(x^{1/2+\varepsilon})$$

$$\sum_{n \leq x} f(n) \neq O(x^{1/2-\varepsilon})$$

for all fixed $\varepsilon > 0$, almost surely.

Improvements

Wintner's results were refined by Erdős and later by Halász, and the best to date upper bound is due to Lau, Tenenbaum and Wu (independently, Basquin), who showed that

$$\sum_{n \le x} f(n) \ll \sqrt{x}(\log\log x)^{2+\varepsilon}$$

for any $\varepsilon > 0$ almost surely, whereas the best known lower bound has been given by Harper, who established that almost surely

$$\left| \sum_{n \le x} f(n) \right| \ge \sqrt{x}(\log\log x)^{1/4+o(1)},$$

for arbitrarily large values of x.

Improvements

Wintner's results were refined by Erdős and later by Halász, and the best to date upper bound is due to Lau, Tenenbaum and Wu (independently, Basquin), who showed that

$$\sum_{n\le x} f(n) \ll \sqrt{x}(\log\log x)^{2+\varepsilon}$$

for any $\varepsilon > 0$ almost surely, whereas the best known lower bound has been given by Harper, who established that almost surely

$$\left|\sum_{n\le x} f(n)\right| \ge \sqrt{x}(\log\log x)^{1/4+o(1)},$$

for arbitrarily large values of x. Motivated by the Law of the Iterated Logarithm, Harper conjectured that

$$\left|\sum_{n\le x} f(n)\right| \asymp \frac{\sqrt{x}}{(\log\log x)^{1/4}} \times \sqrt{\log\log x} = \sqrt{x}(\log\log x)^{1/4}.$$

A new result

In the direction of the previous conjecture, we prove the following theorem.

Theorem (M.)

Let $f(n)$ be a Rademacher random multiplicative function. Then, for any $\varepsilon > 0$ and as $x \to +\infty$, we almost surely have

$$\sum_{\substack{n \le x \\ P(n) > \sqrt{x}}} f(n) \ll \sqrt{x}(\log\log x)^{1/4+\varepsilon}.$$

Proof setup

We would like to show that the event

$$\Big|\sum_{\substack{n\le x\\ P(n)>\sqrt{x}}} f(n)\Big| > \sqrt{x}(\log\log x)^{1/4+\varepsilon}, \text{ for infinitely many } x,$$

holds with null probability.

Proof setup

We would like to show that the event

$$\Bigg|\sum_{\substack{n\le x\\ P(n)>\sqrt{x}}} f(n)\Bigg| > \sqrt{x}(\log\log x)^{1/4+\varepsilon}, \text{ for infinitely many } x,$$

holds with null probability. To this aim:

- We study the event on a sequence of 'test points' x_i;
- We control the increments of f between two consecutive test points;
- We collect together inside 'test intervals' $[X_{\ell-1}, X_\ell)$ the information we gather, and use the first Borel–Cantelli's lemma.

The value on test points

We let $x_i := \lfloor e^{i^{\varepsilon}} \rfloor$. We can write

$$\sum_{\substack{n \le x_i \\ P(n) > \sqrt{x_i}}} f(n) = \sum_{\sqrt{x_i} < p \le x_i} f(p) \sum_{m \le x_i/p} f(m).$$

The value on test points

We let $x_i := \lfloor e^{i^{\varepsilon}} \rfloor$. We can write

$$\sum_{\substack{n \le x_i \\ P(n) > \sqrt{x_i}}} f(n) = \sum_{\sqrt{x_i} < p \le x_i} f(p) \sum_{m \le x_i/p} f(m).$$

By Hoeffding's inequality

$$\mathbb{P}\Bigg(\Bigg| \sum_{\substack{n \le x_i \\ P(n) > \sqrt{x_i}}} f(n) \Bigg| \ge \sqrt{x_i} (\log\log x_i)^{1/4+\varepsilon} \; |\{ f(q) : q \le \sqrt{x_i} \} \Bigg)$$

$$\ll \exp\left(- \frac{x_i (\log\log x_i)^{1/2+2\varepsilon}}{V(x_i)} \right),$$

where $V(x_i) := \sum_{\sqrt{x_i} < p \le x_i} \left| \sum_{m \le x_i/p} f(m) \right|^2$.

The value on test points

We let $x_i := \lfloor e^{i^{\varepsilon}} \rfloor$. We can write

$$\sum_{\substack{n \le x_i \\ P(n) > \sqrt{x_i}}} f(n) = \sum_{\sqrt{x_i} < p \le x_i} f(p) \sum_{m \le x_i/p} f(m).$$

By Hoeffding's inequality

$$\mathbb{P}\Bigg(\Bigg| \sum_{\substack{n \le x_i \\ P(n) > \sqrt{x_i}}} f(n) \Bigg| \ge \sqrt{x_i} (\log\log x_i)^{1/4+\varepsilon} \, |\{f(q) : q \le \sqrt{x_i}\} \Bigg)$$

$$\ll \exp\left(- \frac{x_i (\log\log x_i)^{1/2+2\varepsilon}}{V(x_i)} \right),$$

where $V(x_i) := \sum_{\sqrt{x_i} < p \le x_i} \left| \sum_{m \le x_i/p} f(m) \right|^2$. In this way, we gain a $\log\log x$ factor compared to Basquin and Lau–Tenenbaum–Wu, by replacing $\log\log x$ high moments bounds with just one.

The main conditioning

We would like to show that $V(x_i) \ll x_i/\sqrt{\log\log x_i}$, uniformly on $x_i \in [X_{\ell-1}, X_\ell]$.

The main conditioning

We would like to show that $V(x_i) \ll x_i/\sqrt{\log\log x_i}$, uniformly on $x_i \in [X_{\ell-1}, X_\ell]$. To this aim, we rewrite $V(x_i)$ as roughly

$$\frac{x_i}{\log x_i} \int_{-\infty}^{+\infty} \left| \frac{\mathcal{S}_{x_i}(1/2+it)}{1/2+it} \right|^2 dt,$$

where

$$\mathcal{S}_{x_i}(1/2+it) := \prod_{p \le x_i} \left(1 + \frac{f(p)}{p^{1/2+it}}\right).$$

The integral turns out to be a submartingale with respect to the filtration $\mathcal{F}_i := \sigma(\{f(p) : \ p \le x_i\})$.

The main conditioning

We would like to show that $V(x_i) \ll x_i/\sqrt{\log\log x_i}$, uniformly on $x_i \in [X_{\ell-1}, X_\ell]$. To this aim, we rewrite $V(x_i)$ as roughly

$$\frac{x_i}{\log x_i} \int_{-\infty}^{+\infty} \left| \frac{\mathcal{S}_{x_i}(1/2+it)}{1/2+it} \right|^2 dt,$$

where

$$\mathcal{S}_{x_i}(1/2+it) := \prod_{p \le x_i} \left(1 + \frac{f(p)}{p^{1/2+it}}\right).$$

The integral turns out to be a submartingale with respect to the filtration $\mathcal{F}_i := \sigma(\{f(p) : \ p \le x_i\})$.

We then condition on the event that

$$\int_{-\infty}^{+\infty} \frac{|\mathcal{S}_{X_{\ell-1}}(1/2+it)|^2}{|1/2+it|^2} dt \le \frac{\sqrt{T} 2^{(\ell-1)^K}}{\sqrt{(\ell-1)^K}}.$$

The main conditioning

We would like to show that $V(x_i) \ll x_i/\sqrt{\log\log x_i}$, uniformly on $x_i \in [X_{\ell-1}, X_\ell]$. To this aim, we rewrite $V(x_i)$ as roughly

$$\frac{x_i}{\log x_i}\int_{-\infty}^{+\infty}\left|\frac{\mathcal{S}_{x_i}(1/2+it)}{1/2+it}\right|^2 dt,$$

where

$$\mathcal{S}_{x_i}(1/2+it) := \prod_{p\le x_i}\left(1+\frac{f(p)}{p^{1/2+it}}\right).$$

The integral turns out to be a submartingale with respect to the filtration $\mathcal{F}_i := \sigma(\{f(p) :\ p \le x_i\})$.
We then condition on the event that

$$\int_{-\infty}^{+\infty}\frac{|\mathcal{S}_{X_{\ell-1}}(1/2+it)|^2}{|1/2+it|^2}dt \le \frac{\sqrt{T}2^{(\ell-1)^K}}{\sqrt{(\ell-1)^K}}.$$

The impact of the denominator normalization in our theorem is a further gain of a $(\log\log x)^{1/4}$ factor compared to past works.

Managing low moments

To improve on the trivial bound on the conditional variance, we need to consider it over a very large number of test points. For this reason, we let $X_\ell := e^{2^{\ell^K}}$, with $K := 1/(4\varepsilon)$.

Managing low moments

To improve on the trivial bound on the conditional variance, we need to consider it over a very large number of test points. For this reason, we let $X_\ell := e^{2^{\ell^K}}$, with $K := 1/(4\varepsilon)$. This will force us to suitably renormalize the submartingale sequence:

$$\frac{1}{\log x_i}\left(\frac{\log x_i}{\log X_{\ell-1}}\right)^{1/(\ell-1)^K}\int_{-\infty}^{+\infty}\left|\frac{\mathcal{S}_{x_i}(1/2+it)}{1/2+it}\right|^2 dt.$$

Managing low moments

To improve on the trivial bound on the conditional variance, we need to consider it over a very large number of test points. For this reason, we let $X_\ell := e^{2^{\ell^K}}$, with $K := 1/(4\varepsilon)$. This will force us to suitably renormalize the submartingale sequence:

$$\frac{1}{\log x_i}\left(\frac{\log x_i}{\log X_{\ell-1}}\right)^{1/(\ell-1)^K}\int_{-\infty}^{+\infty}\left|\frac{\mathcal{S}_{x_i}(1/2+it)}{1/2+it}\right|^2 dt.$$

By means of Doob's maximal inequality we can show that with high probability it is uniformly small on test points.

Managing low moments

To improve on the trivial bound on the conditional variance, we need to consider it over a very large number of test points. For this reason, we let $X_\ell := e^{2^{\ell^K}}$, with $K := 1/(4\varepsilon)$. This will force us to suitably renormalize the submartingale sequence:

$$\frac{1}{\log x_i}\left(\frac{\log x_i}{\log X_{\ell-1}}\right)^{1/(\ell-1)^K}\int_{-\infty}^{+\infty}\left|\frac{\mathcal{S}_{x_i}(1/2+it)}{1/2+it}\right|^2 dt.$$

By means of Doob's maximal inequality we can show that with high probability it is uniformly small on test points.
The renormalization allows to take $T \approx (\log\log x)^\varepsilon$, compared to Basquin and Lau–Tenenbaum–Wu choice $T \approx (\log\log x)^{1+\varepsilon}$. As a consequence, we gain a last $\sqrt{\log\log x}$ factor in our bound compared to previous works, overall passing from $(\log\log x)^{2+\varepsilon}$ to $(\log\log x)^{1/4+\varepsilon}$.

Thank you for your attention!