Rendiconti Sem. Mat. Univ. Pol. Torino
Vol. 78, 1 (2020), 125 – 141

Alessandro Gambini, Remis Tonon, Alessandro Zaccagnini,
with an addendum by Jacques Benatar and Alon Nishry

SIGNED HARMONIC SUMS OF INTEGERS WITH k DISTINCT PRIME FACTORS

Abstract. We give some theoretical and computational results on "random" harmonic sums with prime numbers, and more generally, for integers with a fixed number of prime factors.

Keywords: Egyptian fractions; harmonic numbers; harmonic sums.
2010 Mathematics Subject Classification: Primary 11D75, Secondary 11B99.

1. Introduction and general setting

It is well known that the harmonic series restricted to prime numbers diverges, as the harmonic series itself. This was first proved by Leonhard Euler in 1737 [7], and it is considered as a landmark in number theory. The proof relies on the fact that

$$\sum_{n=1}^{N} \frac{1}{n} = \log N + \gamma + O(1/N),$$

where $\gamma \simeq 0.577215\ldots$ is the Euler–Mascheroni constant. The corresponding result for primes is one of the formulae proved by Mertens, namely

$$\sum_{p \le N} \frac{1}{p} = \log\log N + A + O\left(\frac{1}{\log N}\right),$$

where $A \simeq 0.2614972\ldots$ is the Meissel–Mertens constant. It is also referred to as Hadamard–de la Vallée-Poussin constant that appears in Mertens' second theorem.

Recently, Bettin, Molteni and Sanna [2] studied the random harmonic series

$$X := \sum_{n=1}^{\infty} \frac{s_n}{n}, \tag{1}$$

where $s_1, s_2, \ldots$ are independent uniformly distributed random variables in $\{-1,+1\}$. Based on the previous work by Morrison [9, 10] and Schmuland [12], they proved the almost sure convergence of (1) to a density function g, getting lower and upper bounds of the minimum of the distance of a number $\tau \in \mathbb{R}$ to a partial sum $\sum_{n=1}^{N} s_n/n$. In 1976 Worley studied the same problem in terms of upper bound of (1) both in the case $\tau = 0$ (see [13]) and for a generic $\tau \in \mathbb{R}$ (see [14]); his approach is not probabilistic but he has achieved an upper bound comparable to that of [2]. For further references, see also Bleicher and Erdős [3, 4], where the authors treated the number of distinct subsums of $\sum_1^N 1/n$, which corresponds to taking s_i independent uniformly distributed random variables in $\{0,1\}$. A more complete list of references can be found in [2].

The purpose of this paper is firstly to show that basically the same results hold for a general sequence of integers under some suitable, and not too restrictive, conditions; moreover, that a stronger result can be reached if we restrict to integers with exactly k distinct prime factors.

Although Bettin, Molteni and Sanna [2] treat both the lower bound and the upper bound, we are mainly interested in the upper bound using a probabilistic approach. As we will see, in the cases that we treat, we will not be able to say anything about the lower bound, except in terms of numerical computations.

We will use a consistent notation with the previous works by Bettin, Molteni and Sanna [1], [2], Crandall [6] and Schmuland [12].

1.1. General setting of the problem

We denote by $\mathbb{N}$ the set of positive integers. Let $(a_n)_{n\in\mathbb{N}}$ be a strictly decreasing sequence of positive real numbers such that

$$\lim_{n\to+\infty} a_n = 0 \qquad \text{and} \qquad \sum_{n\geq 1} a_n = +\infty. \tag{2}$$

Notice that

$$\sum_{n\geq 1} (-1)^n a_n$$

converges (not absolutely) by Leibniz's rule. Hence, by Riemann's theorem, given λ, $\Lambda \in [-\infty,+\infty]$ with $\lambda \leq \Lambda$, we can arrange the choice of the signs $s_n = s_n(\lambda,\Lambda) \in \{-1,1\}$, in such a way that

$$\liminf_{N\to+\infty} \sum_{n\leq N} s_n a_n = \lambda \qquad \text{and} \qquad \limsup_{N\to+\infty} \sum_{n\leq N} s_n a_n = \Lambda.$$

As we said above, we are mainly interested in prime numbers, so we introduce some further reasonable hypotheses on the sequence a_n: we assume that $b_n = a_n^{-1} \in \mathbb{N}$, so that b_n is strictly increasing, and that

$$n \leq b_n \leq nB(n), \tag{3}$$

where $B(n) = n^{\beta(n)}$, with β a real-valued decreasing function such that $\beta(n) = o(1)$. In order to prove Proposition 20 below, we will assume a more restrictive condition on β, that is

$$\beta(n) \leq \frac{1}{8\log\log n} \qquad \text{for sufficiently large } n. \tag{4}$$

Actually, this assumption is not strictly necessary and we will discuss this in Remark 25. Nevertheless, since the series $\sum a_n$ must diverge, this condition is not too restrictive, and besides it is satisfied by most of the interesting sequences, like arithmetic progressions, the one of primes, and primes in arithmetic progressions.

Let us introduce some more notation: we consider the set

$$\mathfrak{S}(N) = \Big\{ \sum_{n \le N} s_n a_n \colon s_n \in \{-1,1\} \text{ for } n \in \{1,\dots,N\} \Big\}, \tag{5}$$

and, for a given $\tau \in \mathbb{R}$, we set

$$\mathfrak{m}_N(\tau) = \min\big\{ |S_N - \tau| \colon S_N \in \mathfrak{S}(N) \big\}.$$

In other words, for a given $N \in \mathbb{N}$, the goal is to find the choice of signs such that $|S_N - \tau|$ attains its minimum value. Finally, we define the random variable

$$X_N := \sum_{n=1}^{N} s_n a_n,$$

where the signs s_n are taken uniformly and independently at random in $\{-1,1\}$. We will study its small scale distribution. With a slight abuse of notation, we denote by s_n both the signs in the definition (5) and the random variables in the definition above.

1.2. Results

For ease of comparison with the results in Bettin, Molteni and Sanna [2], we now state our main results in the following form, even though more precise versions of them are to be found within the paper.

Theorem 12. *Let β satisfy (4). Then there exists $C > 0$ such that for every $\tau \in \mathbb{R}$ we have*

$$\mathfrak{m}_N(\tau) < \exp(-C \log^2 N)$$

for all sufficiently large N depending on τ.

Theorem 13. *Let $(b_n)_{n\in\mathbb{N}}$ be the sequence of integers having exactly k distinct prime factors. Then, for every $\tau \in \mathbb{R}$ and for all sufficiently large N depending on τ, we have*

$$\mathfrak{m}_N(\tau) < \exp(-f(N)),$$

where f is any function satisfying

$$f(N) = o\left(N^{1/(2k+1)-\varepsilon}\right).$$

Remark 14. We emphasize the fact that the estimate obtained in Theorem 13 holds uniformly for every $\tau \in \mathbb{R}$ in any fixed compact set.

Corollary 15 (J. Benatar and A. Nishry)**.** *For any fixed $\tau \in \mathbb{R}$, $\varepsilon > 0$ and any sufficiently large N there exists a choice of signs $(s_n)_{n\le N} \in \{-1,1\}^N$, such that*

$$\left| \sum_{n \le N} \frac{s_n}{n} - \tau \right| \ll_{\tau,\varepsilon} \exp\left(-N^{1/3-\varepsilon}\right).$$

We collect some numerical results for $k = 1$ in Tables 1, 2 and 3. The sequence of Tables 1 and 2 appears in OEIS A332399: see [5].

Acknowledgements. We thank Sandro Bettin and Giuseppe Molteni for many conversations on the subject, and Mattia Cafferata for his help in computing the tables at the end of the present paper. We also warmly thank Jacques Benatar and Alon Nishry for their fruitful suggestions which improved our paper, for providing us references and for letting us include their proof of Corollary 15 in this paper. R. Tonon and A. Zaccagnini are members of the INdAM group GNSAGA, which partially funded their participation to the Second Symposium on Analytic Number Theory in Cetraro, where some of this work was done.

2. Lemmas

In this section we study some properties of the general sequence defined in (2), using the classical notation: $\mathbb{E}[X]$ denotes the expected value of a random variable X, $\mathbb{P}(E)$ the probability of an event E. For each continuous function with compact support $\Phi \in \mathcal{C}_c(\mathbb{R})$ we denote by $\widehat{\Phi}$ its Fourier transform defined as follows:

$$\widehat{\Phi}(x) := \int_{\mathbb{R}} \Phi(y)\, \mathrm{e}^{-2\pi i x y}\, \mathrm{d}y.$$

We are actually interested in smooth functions, because the smoothness of the density of any random variable X is related to the decay at infinity of its characteristic function, defined precisely by its Fourier transform.

For each $N \in \mathbb{N} \cup \{\infty\}$, for any $x \in \mathbb{R}$ and for any sequence satisfying (2), we also define the product

$$\rho_N(x) := \prod_{n=1}^{N} \cos(\pi x a_n) \qquad \text{and} \qquad \rho(x) := \rho_\infty(x).$$

We begin with the following lemma, which is a more general version of Lemma 2.4 from [2].

Lemma 16. *We have*

$$\mathbb{E}[\Phi(X_N)] = \int_{\mathbb{R}} \widehat{\Phi}(x) \rho_N(2x)\, \mathrm{d}x$$

for all $\Phi \in \mathcal{C}_c^1(\mathbb{R})$.

Proof. By the definition of expected value we have

$$\mathbb{E}[\Phi(X_N)] = \frac{1}{2^N} \sum_{s_1,\dots,s_N \in \{-1,1\}} \Phi\left(\sum_{n=1}^{N} s_n a_n \right).$$

Using the inverse Fourier transform we get

$$\begin{aligned}\mathbb{E}[\Phi(X_N)] &= \frac{1}{2^N}\sum_{s_1,\ldots,s_N\in\{-1,1\}}\int_{\mathbb{R}}\widehat{\Phi}(x)\exp\left(2\pi i x\sum_{n=1}^N s_n a_n\right)\mathrm{d}x \\ &= \int_{\mathbb{R}}\widehat{\Phi}(x)\frac{1}{2^N}\sum_{s_1,\ldots,s_N\in\{-1,1\}}\exp\left(2\pi i x\sum_{n=1}^N s_n a_n\right)\mathrm{d}x.\end{aligned}$$

Exploiting the fact that $e^{i\alpha}+e^{-i\alpha}=2\cos(\alpha)$, we have

$$\sum_{s_1,\ldots,s_N\in\{-1,1\}}\exp\left(2\pi i x\sum_{n=1}^N s_n a_n\right)=\frac{1}{2}\sum_{s_1,\ldots,s_N\in\{-1,1\}}2\cos\left(2\pi x\sum_{n=1}^N s_n a_n\right).$$

Finally, taking advantage of Werner's trigonometric identities, we obtain

$$\mathbb{E}[\Phi(X_N)]=\int_{\mathbb{R}}\widehat{\Phi}(x)\rho_N(2x)\,\mathrm{d}x. \qquad \square$$

We will need also a generalisation of Lemma 2.5 from [2], which is the following

Lemma 17. *For all $N\in\mathbb{N}$ and $x\in[0,\sqrt{N}]$ we have*

$$\rho_N(x)=\rho(x)\left(1+O\left(x^2/N\right)\right).$$

Proof. We recall that a_n is defined as in (2) and satisfies (3). In particular $a_n=O(1/n)$, so that the same argument in the proof of Lemma 2.5 of [2] holds. □

Let us now define, for every positive integer N and any real δ and x the set

$$\mathcal{S}\big(N,\delta,x,(a_n)_{n\geq 1}\big):=\{n\in\{1,\ldots,N\}:\ \|xa_n\|\geq\delta\},$$

where $\|\cdot\|$ denotes the distance from the nearest integer. For brevity, we sometimes drop the dependence on the sequence $(a_n)_{n\geq 1}$.

Lemma 18. *For all $N\in\mathbb{N}$ and for all $x,\delta\geq 0$ we have*

$$|\rho_N(x)|\leq\exp\left(-\frac{\pi^2\delta^2}{2}\cdot\#\mathcal{S}(N,\delta,x)\right).$$

Proof. It is a straightforward consequence of the inequality

$$|\cos(\pi x)|\leq\exp\left(-\frac{\pi^2\|x\|^2}{2}\right). \qquad \square$$

Lemma 19. *For any $N\in\mathbb{N}$, $x\in\mathbb{R}$ and $0<\delta<1/2$ we have*

$$\frac{N}{2}-D(N,y(\delta),x)<\#\mathcal{S}(N,\delta,x)<N-D(N,y(\delta)/2,x),$$

where

$$D(N,y,x) = D(N,y,x,(b_n)_{n\geq 1}) := \sum_{x-y<m<x+y} \sum_{\substack{b_n \mid m \\ N/2\leq n\leq N}} 1$$

and $y(\delta) := \delta N B(N)$.

Proof. As in Lemma 3.3 of [2], we observe that

$$\frac{N}{2} - T(N,\delta,x) < \#\mathcal{S}(N,\delta,x) < N - T(N,\delta,x),$$

where

$$T(N,\delta,x) := \#\{n \in \mathbb{N} \cap [N/2,N] \colon \|x a_n\| < \delta\}.$$

Now, recalling that $a_n = 1/b_n$, we have

$$\begin{aligned} T(N,\delta,x) &= \#\{n \in \mathbb{N} \cap [N/2,N] : \exists \ell \in \mathbb{N},\ \ell - \delta < x a_n < \ell + \delta\} \\ &= \#\{n \in \mathbb{N} \cap [N/2,N] : \exists \ell \in \mathbb{N},\ x - \delta b_n < \ell b_n < x + \delta b_n\}. \end{aligned}$$

From our hypothesis (3) we know that $b_n \leq N B(N)$; then

$$\begin{aligned} T(N,\delta,x) &< \#\{n \in \mathbb{N} \cap [N/2,N] \colon \exists \ell \in \mathbb{Z},\ x - y(\delta) < \ell b_n < x + y(\delta)\} \\ &= D(N,y(\delta),x). \end{aligned}$$

This proves the lower bound; the upper bound follows with the same argument. □

Proposition 20. *Let A be a fixed positive constant and, for N sufficiently large,*

$$\beta(N) \leq \frac{1}{8 \log\log N}.$$

Then there exists $C' > 0$ *such that* $|\rho_N(x)| < x^{-A}$ *for all sufficiently large positive integers N and for all* $x \in [N, \exp(C'(\log N)^2)]$.

Proof. The proof follows along the same lines as Proposition 3.2 of [2]: we take

$$\overline{\delta} = \frac{2\sqrt{2A\log x}}{\pi} N^{-1/2} \quad \text{and} \quad x \in \left[N, \exp\left(\frac{\pi^2 N}{32A}\right)\right),$$

so that $0 < \overline{\delta} < 1/2$ and $y(\overline{\delta}) = \overline{\delta} N B(N) < x$.

By Lemmas 18 and 19, if we show that $D(N,y(\overline{\delta}),x) < N/4$, then we get $|\rho_N(x)| < 1/x^A$. Considering that b_n is a sequence of positive integers, we use Rankin's trick with $w \in (1/4, 1/2)$ and Ramanujan's result on $\sigma_{-s}(n)$ [11] to obtain

$$
\begin{aligned}
D(N, y(\bar{\delta}), x) &< \frac{4}{\pi}\sqrt{2AN\log x}\,B(N)\cdot\max_{m\le 2x}\sum_{\substack{b_n\mid m\\ N/2\le n\le N}} 1 \\
&< \frac{4}{\pi}\sqrt{2AN\log x}\,B(N)\cdot\max_{m\le 2x}\sum_{\substack{k\mid m\\ N/2\le k\le NB(N)}} 1 \\
&\le \frac{4}{\pi}\sqrt{2AN\log x}\,B(N)\cdot\max_{m\le 2x}\sum_{\substack{k\mid m\\ N/2\le k\le NB(N)}} \left(\frac{NB(N)}{k}\right)^{w} \\
&= \frac{4}{\pi}N^{\frac{1}{2}+w}B(N)^{1+w}\sqrt{2A\log x}\cdot\max_{m\le 2x}\sum_{\substack{k\mid m\\ N/2\le k\le NB(N)}} k^{-w} \\
&\le \frac{4}{\pi}N^{\frac{1}{2}+w}B(N)^{1+w}\sqrt{2A\log x}\cdot\max_{m\le 2x}\sigma_{-w}(m) \\
&< \frac{4}{\pi}N^{\frac{1}{2}+w}B(N)^{1+w}\sqrt{2A\log x}\cdot\exp\left(C_1\frac{(\log 2x)^{1-w}}{\log\log 2x}\right),
\end{aligned}
$$

where C_1 is the constant of Ramanujan's theorem, as it is stated in Lemma 3.4 of [2].

Let $w = w(x) := 1/2 - \varphi(x)$, where φ is a positive decreasing function that we will choose later. Then we have

$$
B(N)^{1+w} = \exp\left(\left(\frac{3}{2} - \varphi(x)\right)\beta(N)\log N\right),
$$

and so we would be done if we showed that

$$
N^{1-\varphi(x)+(3/2-\varphi(x))\beta(N)}\sqrt{\log x}\cdot\exp\left(C_1\frac{(\log 2x)^{1/2+\varphi(x)}}{\log\log 2x}\right) = o(N),
$$

that is

$$
\sqrt{\log x}\cdot\exp\left(C_1\frac{(\log 2x)^{1/2+\varphi(x)}}{\log\log 2x}\right) = o\big(N^{\varphi(x)+(\varphi(x)-3/2)\beta(N)}\big).
$$

Hence we must have

$$
\varphi(x) + (\varphi(x) - 3/2)\beta(N) > 0,
$$

that is

$$
\beta(N) < \frac{\varphi(x)}{3/2 - \varphi(x)} \approx \frac{2}{3}\varphi(x).
$$

Since φ is decreasing and we want to maintain the same range for x as in [2], that is $x \in \left[N, \exp\left(C'(\log N)^2\right)\right]$, we need to have

$$
\beta(N) \lesssim \frac{2}{3}\varphi\left(\exp\left(C'(\log N)^2\right)\right).
$$

Let us take $\varphi(x) = (\log\log 2x)^{-1}$ and $\beta(N)$ such that for $x \in \left[N, \exp\left(C'(\log N)^2\right)\right]$ it holds

$$\beta(N) \le \frac{2}{3J}\varphi(x) = \frac{2}{3J}\frac{1}{\log\log 2x}, \tag{6}$$

where $J \in \mathbb{R}$, $J > 1$. Then we would achieve our goal if we showed that

$$\sqrt{\log x} \cdot \exp\left(C_1 e \frac{(\log 2x)^{1/2}}{\log\log 2x}\right) = o\left(\exp\left(\left(1 - \frac{1}{J} + o(1)\right)\frac{\log N}{\log\log 2x}\right)\right),$$

that is

$$\exp\left(C_1 e \frac{(\log 2x)^{1/2}}{\log\log 2x} - \left(1 - \frac{1}{J} + o(1)\right)\frac{\log N}{\log\log 2x} + \frac{1}{2}\log\log x\right) = o(1).$$

This condition is equivalent to

$$C_1 e \frac{(\log 2x)^{1/2}}{\log\log 2x} - \left(1 - \frac{1}{J} + o(1)\right)\frac{\log N}{\log\log 2x} + \frac{1}{2}\log\log x \to -\infty.$$

Taking into account the ranges for x, we see that it is sufficient to have

$$\frac{1}{\log\log N}\left[C_1\sqrt{C'}\, e \log N(1 + o(1)) - \left(1 - \frac{1}{J}\right)\log N + O\left((\log\log N)^2\right)\right] \to -\infty.$$

We recall that, by our choice of x and N, we have $\log\log x \asymp \log\log N$. Hence, we just need to take C' sufficiently small, in a way that

$$C' < \left(\frac{J-1}{C_1 e J}\right)^2, \tag{7}$$

to guarantee that $D(N, y(\overline{\delta}), x) < N/4$ for large N. For the sake of simplicity, we take $J = 2$ and the proposition is proved as stated. □

Remark 21. We remark here that condition (4) on β, which we assumed to prove the proposition, was necessary to ensure the existence of the function φ satisfying all the properties we needed, and in particular (6).

Corollary 22. *Let A be a fixed positive constant and β satisfy (4). Then $|\rho(x)| < x^{-A}$ for all sufficiently large $x \in \mathbb{R}$.*

Proof. It holds

$$|\rho(x)| = \left|\rho_{\lfloor x\rfloor+1}(x) \prod_{n > \lfloor x\rfloor+1} \cos(\pi x a_n)\right| < x^{-A}. \square$$

Theorem 23. *Let $C' > 0$ satisfy (7) and β satisfy (4). Then for all intervals $I \subseteq \mathbb{R}$ of length $|I| > \exp(-C'(\log N)^2)$ one has*

$$\mathbb{P}[X_N \in I] = \int_I g(x)\,\mathrm{d}x + o(|I|),$$

as $N \to \infty$, where

$$g(x) := 2\int_0^\infty \cos(2\pi ux)\prod_{n=1}^{\infty}\cos\left(\frac{2\pi u}{b_n}\right)\mathrm{d}u = 2\int_0^\infty \cos(2\pi ux)\rho(2u)\,\mathrm{d}u.$$

The proof follows along the same lines as Theorem 2.1 in [2] and we omit the details for brevity.

Corollary 24. *Let β satisfy (4). For all $\tau \in \mathbb{R}$ and $C' > 0$ satisfying (7), we have*

$$\#\left\{ (s_1,\dots,s_N) \in \{-1,+1\}^N : \left|\tau - \sum_{n=1}^{N}\frac{s_n}{b_n}\right| < \delta \right\} \sim 2^{N+1}g(\tau)\delta(1+o_{C',\tau}(1))$$

as $N \to \infty$ and $\delta \to 0$, uniformly in $\delta \geq \exp(-C'(\log N)^2)$. In particular, for large enough N, one has $\mathfrak{m}_N(\tau) < \exp(-C'(\log N)^2)$.

Remark 25. We have imposed condition (4) for β to keep the same range of validity for x as in [2]. We remark that the hypotheses on β could be relaxed at the price of restricting this range: for example, we could take

$$\beta(N) = \frac{\log\log\log N}{\log\log N},$$

and obtain the result of Proposition 20 for $x \in [N, \exp(\log^a N)]$, where $a \in (1,2)$ is a suitable constant. In fact, this would weaken directly the estimates that we have just found in Theorem 23 and Corollary 24, where $\exp(-C'(\log N)^2)$ would be replaced by $\exp(-\log^a N)$.

3. Products of k primes

We now leave the general case and concentrate on primes and products of k distinct primes. Hence, we define

$$\mathcal{P}_k := \{\, n \in \mathbb{N} \mid n \text{ is the product of } k \text{ distinct primes} \,\};$$

we will denote by $b_n^{(k)}$ the n-th element of the ordered set $\mathcal{P}_k$. Let us recall the definition of $\mathcal{S}(N,\delta,x)$ in the case $a_n = 1/b_n^{(k)}$:

$$\mathcal{S}(N,\delta,x) := \{\, n \in \{1,\dots,N\} : \|x/b_n^{(k)}\| \geq \delta \,\}.$$

We remark that, since we left the general case, we can now take $B(n) = b_n^{(k)}/n$, and denote it by $B_k(n)$. In 1900, Landau [8] proved that

$$\pi_k(t) := |\mathcal{P}_k \cap \{ n \in \mathbb{N} \mid n \le t \}| = \frac{t}{\log t} \frac{(\log\log t)^{k-1}}{(k-1)!} + O\left(\frac{t(\log\log t)^{k-2}}{\log t} \right),$$

which implies that

$$B_k(n) \sim \log n \frac{(k-1)!}{(\log\log n)^{k-1}}. \tag{8}$$

We can now start with a refinement of Proposition 20, where we extend the interval of validity for x in the case $b_n = b_n^{(k)}$.

Proposition 26. *Let A be a fixed positive constant, $k \in \mathbb{N}$ be fixed and $a_n = 1/b_n^{(k)}$, where $b_n^{(k)}$ is the n-th element of the ordered set $\mathcal{P}_k$. Then $|\rho_N(x)| < x^{-A}$ for all sufficiently large positive integers N and for all $x \in [U, \exp(f(N))]$, where $\log N = o(f(N))$ and*

$$f(N) = o\left(\left(\frac{N}{B_k^2(N)} \right)^{1/(2k+1)} \right),$$

and $U > 1$ is a constant depending on f.

Proof. Let $x \in [N, \exp(f(N))]$. As in the proof of Proposition 20, we need to show that $D(N, y(\overline{\delta}), x) < N/4$, where $\overline{\delta}$ is chosen in the same way and $y(\overline{\delta}) = \overline{\delta} N B_k(N)$. Since now we are considering $x \ge N$, it is easy to see that for sufficiently large N we have $y(\overline{\delta}) \le x$. We recall here that the prime omega function $\omega(n)$ is defined as the number of different prime factors of n, and that

$$\omega(n) \ll \frac{\log n}{\log\log n},$$

as a consequence of the prime number theorem. In this case, we have

$$\begin{aligned} D(N, y(\overline{\delta}), x) &:= \sum_{x - y(\overline{\delta}) < m < x + y(\overline{\delta})} \sum_{\substack{b_n^{(k)} \mid m \\ N/2 \le n \le N}} 1 \le \sum_{x - y(\overline{\delta}) < m < x + y(\overline{\delta})} \sum_{\substack{p_1 \cdots p_k \mid m \\ p_i \text{ distinct primes}}} 1 \\ &\le \sum_{x - y(\overline{\delta}) < m < x + y(\overline{\delta})} \omega(m)^k \le (2y(\overline{\delta}) + 1) \max_{m < x + y(\overline{\delta})} \omega(m)^k \\ &\ll (N \log x)^{1/2} B_k(N) \left(\frac{\log 2x}{\log\log 2x} \right)^k \ll N^{1/2} B_k(N) (\log x)^{k+1/2}, \end{aligned}$$

where we used the trivial bound for the prime omega function. If we show that this quantity is $o(N)$, we are done. So we need

$$\log x = o\left(\left(\frac{N}{B_k^2(N)} \right)^{1/(2k+1)} \right).$$

Hence we can take any f that satisfies

$$f(N) = o\left(\left(\frac{N}{B_k^2(N)}\right)^{1/(2k+1)}\right),$$

where we recall that B_k satisfies (8). The theorem is then proved for $x \in [N, \exp(f(N))]$. If $x < N$, it holds

$$|\rho_N(x)| \le |\rho_{\lfloor x \rfloor}(x)|,$$

hence the result we have just proved holds also whenever $x \le \exp\big(f(\lfloor x \rfloor)\big)$. But there must exist $U > 0$ such that this holds for any $x > U$, since $\log x = o(f(x))$. □

We are now ready to prove a more general version of Theorem 2.1 of [2] for the sequence $\big(b_n^{(k)}\big)_{n\in\mathbb{N}}$.

Theorem 27. *Let f and a_n be defined as in Proposition 26. Then for all intervals $I \subseteq \mathbb{R}$ of length $|I| > \exp(-f(N))$ one has*

$$\mathbb{P}[X_N \in I] = \int_I g(x)\,\mathrm{d}x + o(|I|),$$

as $N \to \infty$, where

$$g(x) := 2\int_0^\infty \cos(2\pi ux) \prod_{n=1}^\infty \cos\left(\frac{2\pi u}{b_n^{(k)}}\right) \mathrm{d}u = 2\int_0^\infty \cos(2\pi ux)\rho(2u)\,\mathrm{d}u.$$

Proof. The proof follows the one of Theorem 2.1 of [2]. Let $\varepsilon > 0$ be fixed. We define

$$\begin{aligned} \xi = \xi_{N,-\varepsilon} &:= \exp(-(1-\varepsilon)f(N)),\\ \xi_+ = \xi_{N,+\varepsilon} &:= \exp(-(1+\varepsilon)f(N)),\\ \xi_0 := \xi_{N,0} &= \exp(-f(N)), \end{aligned}$$

so that $\xi^{-1} < \xi_0^{-1}$ and Proposition 26 holds for $x \in [N, \xi_0^{-1}]$. For an interval $I = [a,b]$ with $b - a > 2\xi_0$, let us define $I^+ := [a-\xi, b+\xi]$ and $I^- := [a+\xi_+, b-\xi_+]$. Then one can construct two smooth functions $\Phi^{\pm}_{N,\varepsilon,I}(x) : \mathbb{R} \to [0,1]$ (from now on, we will drop the subscripts when they are clear by the context) such that

$$\begin{cases} \operatorname{supp}\Phi^+ \subseteq I^+ & \\ \Phi^+(x) = 1 & \text{for } x \in I,\\ \operatorname{supp}\Phi^- \subseteq I & \\ \Phi^-(x) = 1 & \text{for } x \in I^-,\\ (\Phi^{\pm})^{(j)}(x) \ll_j \xi^{-j} & \text{for all } j \ge 0. \end{cases}$$

By the last equation, we know that the Fourier transforms of $\Phi^{\pm}$ satisfy

$$\widehat{\Phi^{\pm}}(x) \ll_B (1+|x|\xi)^{-B} \quad \text{for any } B > 0 \text{ and } x \in \mathbb{R}. \tag{9}$$

Since
$$\mathbb{E}[\Phi^-(X_N)] \le \mathbb{P}[X_N \in I] \le \mathbb{E}[\Phi^+(X_N)],$$
we just need to show that
$$\mathbb{E}[\Phi^\pm(X_N)] = \int_{\mathbb{R}} \Phi^\pm(x) g(x)\,\mathrm{d}x + o_\varepsilon(|I|).$$
From now on, Φ will indicate either Φ^+ or Φ^-. By Lemma 16 we have
$$\mathbb{E}[\Phi(X_N)] = \frac{1}{2}\int_{\mathbb{R}} \widehat{\Phi}(x/2)\rho_N(x)\,\mathrm{d}x = I_1 + I_2 + I_3,$$
where I_1, I_2 and I_3 are the integrals supported respectively in $|x| < N^\varepsilon$, $|x| \in [N^\varepsilon, \xi^{-(1+\varepsilon)}]$ and $|x| > \xi^{-(1+\varepsilon)}$. Note that $\xi^{-(1+\varepsilon)} = \exp((1-\varepsilon^2)f(N)) > \exp(\varepsilon \log N) = N^\varepsilon$, that $\xi^{-(1+\varepsilon)} = \xi_0^{-(1-\varepsilon^2)} < \xi_0^{-1}$, and that $\xi^{-(1+\varepsilon)} \cdot \xi = \xi^{-\varepsilon} = \xi_0^{-\varepsilon(1-\varepsilon)} \to +\infty$ as $N \to +\infty$. By Lemma 17 and Corollary 22, we have
$$\begin{aligned} I_1 &= \frac{1}{2}\int_{-N^\varepsilon}^{N^\varepsilon} \widehat{\Phi}(x/2)\rho_N(x)\,\mathrm{d}x = \frac{1}{2}\int_{-N^\varepsilon}^{N^\varepsilon} \widehat{\Phi}(x/2)\rho(x)\,\mathrm{d}x + O\left(\|\widehat{\Phi}\|_\infty N^{-1+3\varepsilon}\right) \\ &= \frac{1}{2}\int_{\mathbb{R}} \widehat{\Phi}(x/2)\rho(x)\,\mathrm{d}x + O_A\left(\|\widehat{\Phi}\|_\infty N^{-(A-1)\varepsilon}\right) + O\left(\|\widehat{\Phi}\|_\infty N^{-1+3\varepsilon}\right) \\ &= \int_{\mathbb{R}} \widehat{\Phi}(x)\rho(2x)\,\mathrm{d}x + O_\varepsilon\left(\|\Phi\|_1 N^{-1+3\varepsilon}\right), \end{aligned}$$
where to conclude we chose $A = A(\varepsilon)$ sufficiently large. For the second integral, we use Proposition 26 and obtain
$$\begin{aligned} |I_2| &\le \|\widehat{\Phi}\|_\infty \int_{N^\varepsilon}^{\xi^{-(1+\varepsilon)}} |\rho_N(x)|\,\mathrm{d}x \le \|\Phi\|_1 \int_{N^\varepsilon}^{\xi^{-(1+\varepsilon)}} x^{-A}\,\mathrm{d}x \le \|\Phi\|_1 \int_{N^\varepsilon}^{+\infty} x^{-A}\,\mathrm{d}x \\ &\ll_\varepsilon \|\Phi\|_1 N^{-A\varepsilon+\varepsilon} \ll_\varepsilon \|\Phi\|_1 N^{-1}, \end{aligned}$$
where, as before, to conclude we took $A = A(\varepsilon)$ sufficiently large. For the last integral, we recall that trivially $|\rho_N(x)| \le 1$; using the bound (9), we obtain
$$\begin{aligned} |I_3| &\le \int_{|x|>\xi^{-(1+\varepsilon)}} |\widehat{\Phi}(x/2)|\,\mathrm{d}x \ll_B \int_{\xi^{-(1+\varepsilon)}}^{+\infty} (1+x\xi)^{-B}\,\mathrm{d}x = (B-1)(\xi^{-1}+\xi^{-(1+\varepsilon)})^{1-B} \\ &\ll_B \xi_0^{B-1} = o_\varepsilon(\xi_0) = o_\varepsilon(|I|), \end{aligned}$$
where to conclude we chose $B = B(\varepsilon)$ sufficiently large. We can now put these results together: using Parseval's theorem and the fact that $\|\Phi\|_1 = O_\varepsilon(|I|)$, we get
$$\mathbb{E}[\Phi(X_N)] = \int_{\mathbb{R}} \widehat{\Phi}(x)\rho(2x)\,\mathrm{d}x + O_\varepsilon\left(\|\Phi\|_1 N^{-1+3\varepsilon}\right) + o_\varepsilon(|I|) = \int_{\mathbb{R}} \Phi(x)g(x)\,\mathrm{d}x + o_\varepsilon(|I|)$$
and the theorem is then proved. □

Remark 28. By Corollary 22, for any $n \in \mathbb{N}$ it holds

$$\int_{-\infty}^{+\infty} |t^n \rho(t)| \, \mathrm{d}t < \infty,$$

which implies by standard arguments (see e.g. §5 of [12]) that the density g is a smooth strictly positive function. Besides, by the same corollary, $g(x) \ll_D x^{-D}$ for any $D > 0$.

Corollary 29. *For all* $\tau \in \mathbb{R}$, *we have*

$$\#\left\{ (s_1, \ldots, s_N) \in \{-1, 1\}^N : \left| \tau - \sum_{n=1}^{N} \frac{s_n}{b_n^{(k)}} \right| < \delta \right\} \sim 2^{N+1} g(\tau) \delta (1 + o_\tau(1))$$

as $N \to \infty$ *and* $\delta \to 0$, *uniformly in* $\delta \geq \exp(-f(N))$, *where* f *is defined as in Proposition 26. In particular, for* N *large enough, one has* $\mathfrak{m}_N(\tau) < \exp\big(-f(N)\big)$.

4. Addendum (by J. Benatar and A. Nishry): proof of Corollary 15

Proof. Let c_m denote the m-th non-prime integer, so that $c_1 = 1$, $c_2 = 4$, $c_3 = 6, \ldots$ We first approximate τ with a restricted harmonic sum of the form $\sum_{m \leq M} s_m c_m$, where $M = M(N) = N - \pi(N)$. Since $C_m := c_m/m \sim 1$, we may apply Theorem 12 to obtain a sequence of signs $(s_n)_{n \leq M} \in \{-1, 1\}^M$ such that

$$-1 \leq \tau' := \sum_{m \leq M} s_m c_m - \tau \leq 1.$$

Moreover, taking $(p_n)_{n \in \mathbb{N}}$ to be the sequence of primes, we have that $B(n) \sim \log n$ and hence we may apply Theorem 13 to get a choice of signs $(\sigma_n)_{n \leq \pi(N)} \in \{-1, 1\}^{\pi(N)}$ such that

$$\left| \tau' - \sum_{n \leq \pi(N)} \frac{\sigma_n}{p_n} \right| \ll_{\tau, \varepsilon} \exp\left(-N^{1/3 - \varepsilon} \right). \square$$

References

[1] S. Bettin, G. Molteni, and C. Sanna, *Greedy approximations by signed harmonic sums and the Thue–Morse sequence*, Adv. Math. 366, no. 3 (2020), 1–42.

[2] S. Bettin, G. Molteni, and C. Sanna, *Small values of signed harmonic sums*, C. R. Math. Acad. Sci. Paris **356** (2018), no. 11-12, 1062–1074.

[3] M. N. Bleicher and P. Erdős, *The number of distinct subsums of* $\sum_1^N 1/i$, Math. Comp. **29** (1975), 29–42.

[4] M. N. Bleicher and P. Erdős, *Denominators of Egyptian fractions. II*, Illinois J. Math. **20** (1976), no. 4, 598–613.

[5] M. Cafferata, A. Gambini, R. Tonon, and A. Zaccagnini, Sequence A332399 in The On-Line Encyclopedia of Integer Sequences (2020), published electronically at `https://oeis.org/A332399`.

[6] R. E. Crandall, *Theory of ROOF walks*, Unpublished. Available at `http://www.reed.edu/physics/faculty/crandall/papers/ROOF11.pdf`, 2008.

[7] L. Euler, *Variae observationes circa series infinitas*, Commentarii academiae scientiarum imperialis Petropolitanae **9** (1737), 160–188.

[8] E. Landau, *Sur quelques problèmes relatifs à la distribution des nombres premiers*, Bull. Soc. Math. France **28** (1900), 25–38.

[9] K. E. Morrison, *Cosine products, Fourier transforms, and random sums*, Amer. Math. Monthly **102** (1995), no. 8, 716–724.

[10] K. E. Morrison, *Random walks with decreasing steps*, Unpublished manuscript, California Polytechnic State University, 1998.

[11] S. Ramanujan, *Highly composite numbers*, Proc. London Math. Soc. **14** (1915), 347–409.

[12] B. Schmuland, *Random harmonic series*, Amer. Math. Monthly **110** (2003), no. 5, 407–416.

[13] R. T. Worley, *Signed sums of reciprocals. I*, J. Austral. Math. Soc. Ser. A **21** (1976), no. 4, 410–413.

[14] R. T. Worley, *Signed sums of reciprocals. II*, J. Austral. Math. Soc. Ser. A **21** (1976), no. 4, 414–417.

Alessandro Gambini
Dipartimento di Matematica Guido Castelnuovo
Sapienza Università di Roma
Piazzale Aldo Moro, 5
00185 Roma, Italia
email (AG): `alessandro.gambini@uniroma1.it`

Remis Tonon, Alessandro Zaccagnini
Dipartimento di Scienze, Matematiche, Fisiche e Informatiche
Università di Parma
Parco Area delle Scienze, 53/a
43124 Parma, Italia
email (RT): `remis.tonon@unimore.it`
email (AZ): `alessandro.zaccagnini@unipr.it`

Lavoro pervenuto in redazione il 02.10.2019.

1. Numerical data

N	$\mathfrak{m}_N(0) \cdot p_1 \cdots p_N$
1	1
2	1
3	1
4	23
5	43
6	251
7	263
8	21013
9	1407079
10	4919311
11	818778281
12	2402234557
13	379757743297
14	3325743954311
15	54237719914087
16	903944329576111
17	46919460458733911
18	367421942920402841
19	17148430651130576323
20	1236225057834436760243
21	4190310920096832376289
22	535482916756698482410061
23	29119155169912957197310753
24	443284248908491516288671253
25	28438781483496930396689638231
26	10196503226925713726754541885481
27	137512198125317766267968137765087
28	5572821202475305606211985553786081
29	77833992457426020006787481021085581
30	24244850423688161715955346535954790877
31	2030349334778419995324119439659994086131
32	76860130392109667765387079377871685276909
33	5191970624445760882844533168270184721318637
34	329643209271348431895096550792159132283920307
35	19171590315567357340242017182966253037383120953
36	58192378490977430486851365332352874578233287403
37	837477642920747839191618216897250374978659503996169
38	130665466261033919414441892800025408642432364448372023
39	7541550169407232608689149525984967898398947805296216009
40	23868339955752715692132986729285170427530832996153507207

Table 1: The values, multiplied by $p_1 \cdots p_N$, of the smallest signed harmonic sums with the first N primes, with N up to 40. See also OEIS A332399.

N	$\mathfrak{m}_N(0) \cdot p_1 \cdots p_N$
41	3343165792500492306892396976512891068137770193474133826457
42	47233268931962642510303169511493601517566800154537867238057
43	93915329439868205746156163805290441755151986127947916375626793
44	50313439148416324581127610155641150127987318260569172331033593181
45	2035703788246113211455753014584246782664737720644793016891955087197
46	193768861589178044091624877468627581772116464350368833881209864412247
47	4664128549520402650533030541013467806288648880741654578068005845271177
48	252294099680710988063673862003152188841680135741161924018446904086039541
49	1641527055336324967995403445372629420483564255197731535006975381936073433
50	25436424505451332441928319474656471336874167655047366774702187882274894064063
51	1780024077761328763318128562703299120404666081323149178582620236480827415289259
52	115533643751466097619699345183033980786661230484621892531131629910924364040946261
53	34644520573176659229537081198934624126738529150336245449473941125320497104653817109
54	7369668963051661582966392617319633009625522375611294051784365401090471220946387592789
55	1999632582248468763357938742475072167566513418694128163881669512737786988287075374795317
56	151351981933638637742621357138936533979590998748883750430193460129876391573603481014628429
57	15302724902698188450027684974980553939987991074013402437579866232981371846926226684458406969
58	6269085432675155135477773589250562149563926327373176617473379555222137615792922214195964225281
59	429918790837116674905123858093668694474961832761345115366942177591943696826657060080682245858603
60	115809464188499233574522294110279752895686365776568444548440426304978721966632473743873345620708313

Table 2: The values, multiplied by $p_1 \cdots p_N$, of the smallest signed harmonic sums with the first N primes, with N between 41 and 60. See also OEIS A332399.

N	$\Delta_N \cdot p_1 \cdots p_N$
1	1
2	1
3	1
4	2
5	22
6	35
7	263
8	4675
9	24871
10	104006
11	2356081
12	6221080
13	141769355
14	6096082265
15	6928889495
16	367231143235
17	1283811918935
18	78312527055035
19	5246939312687345
20	372532691200801495
21	8815359347599933286
22	223849990729887044174
23	6148176498383067879445
24	179847837287937160817963
25	663024394602752425373130

Table 3: The values, multiplied by $p_1 \cdots p_N$, of the shortest distances Δ_N between different signed harmonic sums with the first N primes, with N up to 25.