Periodica Mathematica Hungarica Vol. 62 (1), 2011, pp. 1–12
DOI: 10.1007/s10998-011-5001-7

ON THE STRONG LAW OF LARGE NUMBERS AND ADDITIVE FUNCTIONS

István Berkes[1], Wolfgang Müller[2] and Michel Weber[3]

[1]Graz University of Technology, Institute of Statistics
Münzgrabenstrasse 11, A-8010 Graz, Austria
E-mail: berkes@tugraz.at

[2]Graz University of Technology, Institute of Statistics
Münzgrabenstrasse 11, A-8010 Graz, Austria
E-mail: w.mueller@tugraz.at

[3]IRMA, Université Louis-Pasteur et C.N.R.S.
7 rue René Descartes, 67084 Strasbourg Cedex, France
E-mail: michel.weber@math.unistra.fr

(*Received September 24, 2010; Accepted October 29, 2010*)

Dedicated to Endre Csáki and Pál Révész
on the occasion of their 75th birthdays

Abstract

Let $f(n)$ be a strongly additive complex-valued arithmetic function. Under mild conditions on f, we prove the following weighted strong law of large numbers: if $X, X_1, X_2, \dots$ is any sequence of integrable i.i.d. random variables, then

$$\lim_{N\to\infty} \frac{\sum_{n=1}^N f(n) X_n}{\sum_{n=1}^N f(n)} = \mathbb{E}X \qquad \text{a.s.}$$

1. Introduction

Let $X, X_1, X_2, \dots$ be i.i.d. integrable random variables and $f(n), n = 1, 2, \dots$ a positive numerical sequence, $F(n) = \sum_{k=1}^n f(k)$. By a classical result of Jamison, Orey and Pruitt [5], under the condition

$$\limsup_{x\to\infty} \frac{1}{x} \#\{n : F(n) \le x f(n)\} < \infty, \tag{1}$$

Mathematics subject classification numbers: 60F15,11A25.
Key words and phrases: strong law of large numbers, weighted i.i.d. sums, strongly additive functions.
[1] Research supported by FWF grant S9603-N23 and OTKA grants K 67961 and K 81928.

0031-5303/2011/$20.00

Akadémiai Kiadó, Budapest
Springer, Dordrecht

we have the weighted strong law

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n}f(k)X_k}{\sum_{k=1}^{n}f(k)}=\mathbb{E}X \qquad \text{a.s.} \tag{2}$$

Conversely, if (2) holds for all i.i.d. sequences $X, X_1, X_2, \dots$ with finite means, then (1) is valid. Note that condition (1) puts a restriction on the distribution of the weight sequence $f(n)$ and not on the magnitude of the weights, as it happens, e.g., in central limit theory. In particular, (1) can fail even for bounded weight sequences $f(n)$, see [5]. Condition (1) is generally difficult to check for irregular sequences $f(n)$, and this leads to the question to study (2) for typical irregular sequences in number theory, for example, additive arithmetic functions. Let $f(n)$, $n=1,2,\dots$, be a real-valued, strongly additive function, i.e., assume that

$$f(mn)=f(m)+f(n) \qquad \text{for } (m,n)=1 \tag{3}$$

and

$$f(p^\alpha)=f(p), \qquad \text{for } p \text{ prime}, \ \alpha=2,3,\dots \tag{4}$$

It follows that

$$f(n)=\sum_{p|n}f(p),$$

so that f is completely determined by its values taken over the primes. A typical example is $\omega(n)$, the number of different prime factors of n. Put

$$A_n=\sum_{p\le n}\frac{f(p)}{p}, \qquad B_n=\sum_{p\le n}\frac{|f(p)|^2}{p}. \tag{5}$$

In [1] we studied the weighted SLLN with coefficients $f(n)$ and proved the following result (see Theorem 1.1 in [1]).

Theorem 1. *Assume that $f\ge 0$ and*

$$B_p\to\infty, \quad f(p)=o(B_p^{1/2}) \quad \text{as } p\to\infty. \tag{6}$$

Then for any i.i.d. sequence $X, X_1, X_2, \dots$ with finite means, the weighted strong law (2) holds.

Condition (6) plays an important role in probabilistic number theory as a nearly optimal sufficient condition for the central limit theorem

$$\lim_{N\to\infty}\frac{1}{N}\#\left\{n\le N : \frac{f(n)-A_N}{B_N^{1/2}}\le x\right\}=\frac{1}{(2\pi)^{1/2}}\int_{-\infty}^{x}e^{-t^2/2}dt \tag{7}$$

(see, e.g., Elliott [3], Kubilius [6]). Halberstam [4] proved that replacing the o by O in (6) the CLT (7) becomes generally false. Note that relation (6) implies the Lindeberg condition

$$\lim_{n\to\infty} \frac{1}{B_n} \sum_{\substack{p<n \\ |f(p)|\geq \varepsilon B_n^{1/2}}} \frac{|f(p)|^2}{p} = 0 \qquad \text{for any } \varepsilon > 0, \tag{8}$$

and, under mild technical assumptions on f, condition (8) is necessary and sufficient for the CLT (7), see again Elliott [3], Kubilius [6]. In [1] we also proved that (6) implies the law of the iterated logarithm corresponding to (2) provided $\mathbb{E}X^2 < \infty$ (see Theorem 1.2 in [1]). We further indicated that if $f(p)$ does not fluctuate too wildly, namely if

$$\sup_{\substack{n\leq p,q\leq n^2 \\ p,q \text{ primes}}} \left|\frac{f(p)}{f(q)}\right| = \mathcal{O}(1), \qquad \text{as } n \to \infty,$$

then Theorem 1 remains valid under condition (8). We raised the question of the validity of Theorem 1 under the sole Lindeberg condition. Recently, Fukuyama and Komatsu [2] answered this question affirmatively.

Theorem 2. *Assume that $f \geq 0$ and the Lindeberg condition* (8) *is satisfied. Then for any i.i.d. sequence $X, X_1, X_2, \dots$ with finite means,* (2) *holds.*

The approach of Fukuyama and Komatsu [2] is elegant and is based on Abel summation. Put

$$G(n) = \sum_{m\leq n} |f(m)|^2. \tag{9}$$

The estimates

$$F(n) \gg nA_n, \qquad G(n) \ll nA_n^2 \tag{10}$$

which for $f \geq 0$ are implied by the Lindeberg condition (8) (see for instance Lemma 2.1 in [1]) are crucial in their proof and the proof remains valid under these sole conditions. Here $\ll$ means the same as the O notation. Observe now that, still assuming $f \geq 0$,

$$F(n) = \sum_{p\leq n} f(p)\lfloor n/p \rfloor = nA_n + O\Big(\sum_{p\leq n} f(p)\Big) = nA_n + O(n(\log n)^{-1/2}B_n^{1/2}), \tag{11}$$

where in the last step we used the prime number theorem and the Cauchy–Schwarz inequality. Thus assuming that

$$B_n \ll A_n^2 \tag{12}$$

it follows that $F(n) \sim nA_n$ and

$$G(n) = \sum_{p\leq n} |f(p)|^2 \lfloor n/p \rfloor + 2 \sum_{2\leq p<q\leq n} f(p)f(q)\lfloor n/pq \rfloor \ll n(B_n + A_n^2) \ll nA_n^2.$$

In other words, (12) implies (10) and thus the proof of Theorem 2 yields the following stronger result:

THEOREM 3. *Let $f \geq 0$ be a strongly additive function satisfying $B_n \ll A_n^2$. Then for any i.i.d. sequence $X, X_1, X_2, \ldots$ with finite means,* (2) *holds.*

Note that the condition $B_n \ll A_n^2$ is weaker than the Lindeberg condition which, as is shown in [1], implies $B_n = o(A_n^2)$. For example, $B_n \ll A_n^2$ is satisfied if $f(p) = (\log p)^\gamma$, $\gamma > 0$, under which the Lindeberg condition and the central limit theorem (7) are false (see Halberstam [4]).

The purpose of the present paper is to extend Theorem 3 to complex-valued additive functions. The following example shows that without $f \geq 0$ the Lindeberg condition (8) is generally not sufficient for the validity of the strong law (2).

EXAMPLE. Let $p_1 < p_2 < \cdots$ be the sequence of the primes and define the function f on the primes by $f(p_{2k-1}) = 1$, $f(p_{2k}) = -1$ $(k = 1, 2, \ldots)$. Then $B_n \sim \log\log n$ and the Lindeberg condition (8) holds, but, as we will show in Section 3, the strong law (2) fails for some i.i.d. sequences $X, X_1, X_2, \ldots$ with finite means.

Call a nonnegative sequence (x_n) *quasi-monotone* if there exists a nonnegative monotone sequence (y_n) with $x_n \ll y_n \ll x_n$. Our main result is the following.

THEOREM 4. *Let f be a complex-valued strongly additive arithmetic function such that $n|A_n| \to \infty$ and $n|A_n|$ is quasi-monotone. Assume that $B_n \to \infty$ and*

$$B_n \ll |A_n|^2, \qquad \sup_{n^h < p \leq n} |f(p)| \ll |A_n| \tag{13}$$

for some $0 < h < 1/4$. Then for any i.i.d. sequence $X, X_1, X_2, \ldots$ with finite means, the weighted strong law (2) *holds.*

Assuming $B_n \ll |A_n|^2$, the second condition of (13) is satisfied if $|f(p)| = O(B_p^{1/2})$, which is weaker than (6) and does not imply the Lindeberg condition, as the example $f(p) = (\log p)^\gamma$, $\gamma > 0$ shows. In analogy with Theorem 3, it is possible that the second condition of (13) can be omitted completely, but this remains open. The quasi-monotonicity of $n|A_n|$ is required because of the use of the Kronecker lemma in the proof; in the case $f \geq 0$ this condition is trivially satisfied. For $f \geq 0$ the first relation of (13) reduces to the condition of Theorem 3, but there is a big difference between the case $f \geq 0$ and the general case. For $f \geq 0$ the Lindeberg condition implies the CLT (7) and also $B_n = o(A_n^2)$ (see, e.g., [1]); this means that the centering factor A_n in (7) dominates the norming factor $B_n^{1/2}$ and consequently $f(n) \sim A_n$ along a set of integers with density 1. In the complex case the Lindeberg condition still implies the CLT (7), but not $B_n = o(|A_n|^2)$ or even $B_n \ll |A_n|^2$ as the example above shows: there A_n is bounded and $B_n \sim \log\log n$. Instead, we get the weaker relation

$$B_n = o(\widetilde{A}_n^2) \quad \text{with} \quad \widetilde{A}_n = \sum_{p \leq n} \frac{|f(p)|}{p}.$$

In the complex case the centering factor in the CLT (7) can be much smaller than the norming factor and in such cases the SLLN (2) may also fail, as the example above shows.

2. Some lemmas

In this section we formulate some lemmas needed for the proof of Theorem 4. Our first lemma extends the classical theorem of Jamison, Orey and Pruitt [5] to complex weights.

LEMMA 1. *Let $f(n)$, $n = 1, 2, \ldots$, be complex numbers and put $F(n) = \sum_{k=1}^{n} f(k)$ and*

$$L(t) = \#\{n : \ |F(n)| \le t|f(n)|\}, \qquad (t \ge 0). \tag{14}$$

Assume that

$$|F(n)| \to \infty \quad \text{and} \quad |F(n)| \text{ is quasi-monotone.} \tag{15}$$

Then the weighted strong law (2) *holds for all i.i.d. sequences $X, X_1, X_2, \ldots$ with finite means if and only if*

$$L(t) \ll t \qquad (t \ge 1). \tag{16}$$

Dropping the quasi-monotonicity of $|F(n)|$, the necessity of (16) *remains valid.*

PROOF. The argument of Jamison, Orey and Pruitt [5] yields the following stronger statement. Assume that there is a nondecreasing monotone sequence $(V(n))_{n\ge1}$ of nonnegative numbers such that for $n \to \infty$

$$V(n) \ll |F(n)| \qquad \text{and} \qquad V(n) \to \infty\,.$$

Then (2) holds for all i.i.d. sequences $X, X_1, X_2, \ldots$ with finite means if $L_V(t) \ll t$ for $t \ge 1$, where

$$L_V(t) = \#\{n : V(n) \le t|f(n)|\}\,.$$

To prove this, set $Y_n = X_n - E(X_1)$ and $Z_n = f(n)Y_n I(|f(n)Y_n| < V(n))$. Then

$$\sum_{n\ge1} P(Z_n \ne f(n)Y_n) \le \sum_{n\ge1} \int_{|f(n)x|\ge V(n)} dP_{Y_1}(x) = E(L_V(|Y_1|)) \ll E(|Y_1|) < \infty\,.$$

Hence (2) follows if

$$\lim_{n\to\infty} \frac{1}{V(n)} \sum_{m=1}^{n} Z_m = 0 \qquad \text{a.s.}$$

Using Kronecker's lemma, this follows from $\sum_{n\geq 1} Z_n/V(n) < \infty$ a.s., and this in turn is true if $\sum_{n\geq 1} \mathrm{Var}\,(Z_n)/V(n)^2 < \infty$. Now

$$\begin{aligned}\sum_{n\geq 1} \frac{\mathrm{Var}\,(Z_n)}{V(n)^2} &= \sum_{n\geq 1} \frac{|f(n)|^2}{V(n)^2} \int_{|f(n)x|<V(n)} x^2\, dP_{Y_1}(x) \\ &= \int x^2 \sum_{|f(n)x|<V(n)} \frac{|f(n)|^2}{V(n)^2}\, dP_{Y_1}(x) \\ &= \int x^2 \int_{(|x|,\infty)} y^{-2}\, dL_V(y)\, dP_{Y_1}(x)\,.\end{aligned}$$

Using the assumption $L_V(t) \ll t$ for $t \geq 1$ and $L_V(t) \leq L_V(1)$ for $0 \leq t < 1$ and partial integration, the last integral is bounded by a constant multiple of $E(|Y_1|) < \infty$. Hence $L_V(x) \ll x$ is sufficient for the validity of the strong law (2). Finally, setting $V(n) = \max_{m\leq n} |F(m)|$, condition (15) implies $V(n) \ll |F(n)|$. Since $L_V(x) \leq L(x)$, condition (16) is sufficient for (2). The necessity part of the argument of Jamison, Orey and Pruitt remains unchanged in the case of complex weights.

Next we need a lemma on divisors of Bernoulli sums. Let $\{\varepsilon_i, i \geq 1\}$ be a Bernoulli sequence, i.e. a sequence of independent random variables such that $\mathbb{P}(\varepsilon_i = 0) = \mathbb{P}(\varepsilon_i = 1) = 1/2$, $(i = 1, 2, \ldots)$. Let $S_n = \sum_{i=1}^n \varepsilon_i$. Consider the elliptic Theta function

$$\Theta(d, m) = \sum_{\ell\in\mathbb{Z}} e^{im\pi\frac{\ell}{d} - \frac{m\pi^2\ell^2}{2d^2}}\,.$$

The following lemma, which is Theorem II from [7], yields precise asymptotics for the probability that a natural number divides S_n.

Lemma 2. *We have the following uniform estimate:*

$$\sup_{2\leq d\leq n} \left|\mathbb{P}\{d|S_n\} - \frac{\Theta(d,n)}{d}\right| = \mathcal{O}((\log n)^{5/2} n^{-3/2}),$$

and

$$\left|\mathbb{P}\{d|S_n\} - \frac{1}{d}\right| = \begin{cases} \mathcal{O}((\log n)^{5/2} n^{-3/2} + \frac{1}{d} e^{-\frac{n\pi^2}{2d^2}}), & \text{if } d \leq \sqrt{n}, \\ \mathcal{O}(1/\sqrt{n}), & \text{if } \sqrt{n} \leq d \leq n. \end{cases}$$

Further, for any $\alpha > 0$ and all $\varepsilon > 0$,

$$\sup_{d<\pi\sqrt{\frac{n}{2\alpha\log n}}} \left|\mathbb{P}\{d|S_n\} - \frac{1}{d}\right| = \mathcal{O}_\varepsilon(n^{-\alpha+\varepsilon})$$

and for any $0 < \rho < 1$ and all $0 < \varepsilon < 1$,

$$\sup_{d<(\pi/\sqrt{2})n^{(1-\rho)/2}} \left|\mathbb{P}\{d|S_n\} - \frac{1}{d}\right| = \mathcal{O}_\varepsilon(e^{-(1-\varepsilon)n^\rho}). \tag{17}$$

Remark. Actually, we will need only relation (17) in the present paper. By using the Poisson summation formula we get

$$\frac{\Theta(d,n)}{d} = \frac{1}{d}\sum_{\ell\in\mathbb{Z}} e^{i\pi n\frac{\ell}{d} - n\pi^2\frac{\ell^2}{2d^2}} = \sqrt{\frac{2}{\pi n}}\sum_{\ell\in\mathbb{Z}} e^{-2(\frac{n}{2d}+\ell)^2\frac{d^2}{n}}$$

and thus

$$\sup_{2\le d\le n}\Big|\mathbb{P}\{d|S_n\} - \sqrt{\frac{2}{\pi n}}\sum_{\ell\in\mathbb{Z}} e^{-2(\frac{n}{2d}+\ell)^2\frac{d^2}{n}}\Big| = \mathcal{O}\Big(\frac{(\log n)^{5/2}}{n^{3/2}}\Big),$$

a further asymptotic formula useful in many situations.

3. Proof of Theorem 4

Let f be a strongly additive, complex-valued arithmetical function satisfying the conditions of Theorem 4. Analogously to (11), we have

$$|F(n)| = n|A_n| + O(n(\log n)^{-1/2}B_n^{1/2}), \tag{18}$$

and thus by $B_n \ll |A_n|^2$ we have $|F(n)| \sim n|A_n|$. Hence the quasi-monotonicity of $n|A_n|$ and $n|A_n| \to \infty$ imply that (15) holds. In view of Lemma 1, for the proof of Theorem 4 it suffices to prove (16) for the function L defined by (14) and to do this, we use the same probabilistic trick as in [1]. Consider a Bernoulli sequence $\{\varepsilon_i,\ i\ge 1\}$ defined on a probability space $(\Omega,\mathcal{A},\mathbb{P})$, possibly different from the probability space supporting the variables $X, X_1, X_2, \dots$ in (2). Put $S_n = \sum_{i=1}^n \varepsilon_i$ and define L by (14). Let $0<\eta<1/2$ and $F_\eta(n) = \inf_{m\ge \eta n}|F(m)|$. Clearly

$$\begin{aligned} L(t) &= \#\{n : |F(n)| \le t|f(n)|\} \\ &\le \#\{n : |F(S_n)| \le t|f(S_n)|\} \le \#\{n : F_\eta(S_n) \le t|f(S_n)|\}, \end{aligned} \tag{19}$$

and this is true for any $t\ge 0$, because the graph of the random walk $\{S_n, n\ge 1\}$ replicates all positive integers with possible multiplicities. We next define

$$\Omega_\eta = \{S_n \ge \eta n \text{ for all } n\ge 1\},$$

and observe that $\mathbb{P}(\Omega_\eta) > 0$. Indeed, $S_n/n \to 1/2$ a.s. and thus there exists an integer $n_0 \ge 2$ such that letting

$$\Omega_\eta^* = \{S_n \ge \eta n \text{ for all } n \ge n_0\}, \qquad A = \{\varepsilon_1 = 1, \dots, \varepsilon_{n_0-1} = 1\},$$

we have $\mathbb{P}(\Omega_\eta^*) > 0$. Let Ω_η^{**} denote the event obtained from Ω_η^* by replacing $\varepsilon_1, \dots, \varepsilon_{n_0-1}$ in the sums S_n by 1. Clearly $\Omega_\eta^* \subseteq \Omega_\eta^{**}$ and thus $\mathbb{P}(\Omega_\eta^{**}) > 0$ and consequently $\mathbb{P}(\Omega_\eta^{**} \cap A) > 0$, since Ω_η^{**} depends on $\varepsilon_{n_0}, \varepsilon_{n_0+1}, \dots$ and thus Ω_η^{**} and A are independent. Now observing that $\Omega_\eta^{**} \cap A \subseteq \Omega_\eta$, relation $\mathbb{P}(\Omega_\eta) > 0$ follows.

Reading (19) on Ω_η gives

$$L(t) \le \#\{n : F_\eta(n) \le t|f(S_n)|\} \quad \text{on } \Omega_\eta \text{ for all } t > 0$$

and consequently

$$\frac{1}{t} L(t) \le \frac{1}{\mathbb{P}(\Omega_\eta)} \mathbb{E}\, \frac{1}{t} \#\{n : F_\eta(n) \le t|f(S_n)|\}. \tag{20}$$

But for all $t > 0$

$$\begin{aligned} &\frac{1}{t}\, \#\{n : F_\eta(n) \le t|f(S_n)|\} \\ &\le 1 + \frac{1}{t} \#\{n \ge t : F_\eta(n) \le t|f(S_n)|\} = 1 + \frac{1}{t} \sum_{n \ge t} \chi\{\, F_\eta^2(n) \le t^2 |f(S_n)|^2\} \\ &\le 1 + t \sum_{n \ge t} \frac{|f(S_n)|^2}{F_\eta^2(n)}. \end{aligned} \tag{21}$$

We now prove the following lemma.

Lemma 3. *For any $0 < h < 1/4$ and sufficiently large n we have*

$$\|f(S_n)\| \le \frac{1}{h} \sup_{n^h < p \le n} |f(p)| + \Big| \sum_{2 \le p \le n^h} \frac{f(p)}{p} \Big| + C \Big(\sum_{2 \le p \le n^h} \frac{|f(p)|^2}{p} \Big)^{1/2}$$

where $\|\cdot\|$ denotes the L_2 norm in $(\Omega, \mathcal{A}, \mathbb{P})$ and C is an absolute constant.

Proof. In the argument that follows, all constants implied by $\ll$ will be absolute. Using the triangular inequality we find

$$\begin{aligned} \|f(S_n)\| \le &\Big\| f(S_n) - \sum_{\substack{2 \le p \le n^h \\ p | S_n}} f(p) \Big\| + \Big\| \sum_{\substack{2 \le p \le n^h \\ p | S_n}} f(p) - \sum_{2 \le p \le n^h} f(p) \mathbb{P}\{p | S_n\} \Big\| + \\ &+ \Big| \sum_{2 \le p \le n^h} f(p) \mathbb{P}\{p|S_n\} - \sum_{2 \le p \le n^h} \frac{f(p)}{p} \Big| + \Big| \sum_{2 \le p \le n^h} \frac{f(p)}{p} \Big|. \end{aligned}$$

Since $f(S_n) = \sum_{p | S_n} f(p)$ and $S_n \le n$, we obtain

$$\Big| f(S_n) - \sum_{\substack{2 \le p \le n^h \\ p | S_n}} f(p) \Big| = \Big| \sum_{\substack{n^h < p \le n \\ p | S_n}} f(p) \Big| \le \sum_{\substack{n^h < p \le n \\ p | S_n}} |f(p)| \le \frac{1}{h} \sup_{n^h < p \le n} |f(p)|.$$

The last bound is justified by the fact that if S_n admits K different prime factors $> n^h$, then $n^{Kh} \le S_n \le n$ implies $Kh \le 1$. Using Lemma 2 with $\rho = 1-2h$ and $\varepsilon = 1/2$, we obtain

$$\begin{aligned}\Big|\sum_{2\le p\le n^h} f(p)\mathbb{P}\{p|S_n\} - \sum_{2\le p\le n^h}\frac{f(p)}{p}\Big| &\le \sum_{2<p\le n^h}|f(p)|\Big|\mathbb{P}\{p|S_n\}-\frac{1}{p}\Big| \\ &\ll e^{-\frac12 n^{1-2h}}\sum_{2\le p\le n^h}|f(p)|.\end{aligned}$$

By Cauchy's inequality

$$\sum_{2\le p\le n^h}|f(p)| \le \Big(\sum_{2\le p\le n^h} p\Big)^{1/2}\Big(\sum_{2\le p\le n^h}\frac{|f(p)|^2}{p}\Big)^{1/2} \le n^h\Big(\sum_{2\le p\le n^h}\frac{|f(p)|^2}{p}\Big)^{1/2},$$

hence

$$\begin{aligned}\Big|\sum_{2\le p\le n^h} f(p)\mathbb{P}\{p|S_n\} - \sum_{2\le p\le n^h}\frac{f(p)}{p}\Big| &\ll e^{-\frac12 n^{1-2h}} n^h\Big(\sum_{2\le p\le n^h}\frac{|f(p)|^2}{p}\Big)^{1/2} \\ &\ll \Big(\sum_{2\le p\le n^h}\frac{|f(p)|^2}{p}\Big)^{1/2}.\end{aligned}$$

We finally prove

$$\Big\|\sum_{\substack{2\le p\le n^h\\ p|S_n}} f(p) - \sum_{2\le p\le n^h} f(p)\mathbb{P}\{p|S_n\}\Big\|^2 \ll \sum_{2<p\le n^h}\frac{|f(p)|^2}{p}. \tag{22}$$

Write the left-hand side of (22) as

$$\begin{aligned}&\mathbb{E}\Big|\sum_{2\le p\le n^h} f(p)(\chi(p|S_n)-\mathbb{P}\{p|S_n\})\Big|^2 \\ &\quad= \sum_{2\le p\le n^h}|f(p)|^2\mathbb{P}\{p|S_n\}(1-\mathbb{P}\{p|S_n\})+ \\ &\qquad +2\Re\Big\{\sum_{2\le p<q\le n^h} f(p)\overline{f(q)}(\mathbb{P}\{pq|S_n\}-\mathbb{P}\{p|S_n\}\mathbb{P}\{q|S_n\})\Big\}.\end{aligned}$$

By relation (17) with $\rho = 1-4h > 0$ and $\varepsilon = 1/2$, we obtain for $p, q \le n^h$

$$\mathbb{P}\{p|S_n\} = \frac{1}{p} + O(e^{-\frac12 n^{1-4h}}), \qquad \mathbb{P}\{pq|S_n\} = \frac{1}{pq} + O(e^{-\frac12 n^{1-4h}}). \tag{23}$$

This also yields

$$|\mathbb{P}\{pq|S_n\} - \mathbb{P}\{p|S_n\}\mathbb{P}\{q|S_n\}| \ll e^{-\frac12 n^{1-4h}}.$$

Thus we conclude

$$\begin{aligned}\Big\| \sum_{\substack{2\le p\le n^h\\ p|S_n}} f(p) - \sum_{2\le p\le n^h} f(p)\mathbb{P}\{p|S_n\}\Big\|^2 \\ &\ll \sum_{2\le p\le n^h} \frac{|f(p)|^2}{p} + e^{-\frac12 n^{1-4h}}\Big(\sum_{2\le p\le n^h}|f(p)|\Big)^2 \\ &\ll \sum_{2\le p\le n^h} \frac{|f(p)|^2}{p} + n^{2h}e^{-\frac12 n^{1-4h}}\sum_{2\le p\le n^h}\frac{|f(p)|^2}{p} \ll \sum_{2\le p\le n^h}\frac{|f(p)|^2}{p},\end{aligned}$$

proving (22). The proof of Lemma 3 is now complete.

□

We can now easily complete the proof of Theorem 4. Observe that the conditions of Theorem 4 imply

$$|F(\eta n)| \gg n\max(|A_{n^h}|, B_{n^h}^{1/2}), \qquad \sup_{n^h<p\le n}|f(p)| \ll |F(\eta n)|/n \tag{24}$$

for any $0<\eta\le 1$, $0<h<1$ and $n\ge n_0(\eta,h)=t_0$. To prove the first relation of (24) note that $\sum_{p\le n}\frac1p = \log\log n + C + o(1)$ for some constant C and thus

$$\sum_{n^h<p\le n}\frac1p = O(1), \qquad \sum_{\eta n<p\le n}\frac1p = o(1).$$

Hence by (13)

$$|A_n - A_{\eta n}| \le \sup_{\eta n<p\le n}|f(p)|\sum_{\eta n<p\le n}\frac1p = o(A_n)$$

giving $A_{\eta n}\sim A_n$ and similarly

$$|A_{n^h} - A_n| \le \sup_{n^h<p\le n}|f(p)|\sum_{n^h<p\le n}\frac1p \ll |A_n|,$$

giving $|A_{n^h}|\ll|A_n|$. Thus $B_n\ll|A_n|^2$, $A_{\eta n}\sim A_n$ and the relation $|F(n)|\sim n|A_n|$ pointed out after (18) imply

$$|F(\eta n)| \gg n|A_{\eta n}| \sim n|A_n| \gg n|A_{n^h}| \gg nB_{n^h}^{1/2}.$$

This proves the first relation of (24) and shows also that the second relation of (24) follows from the second condition of (13).

Let us choose now, as before, $0 < \eta < 1/2$ and $0 < h < 1/4$. Using (21), Lemma 3 and the relation $F_\eta(n) \gg |F(\eta n)|$ implied by the quasi-monotonicity of $|F(n)|$, we get, for $t \ge t_0$,

$$\begin{aligned}
&\mathbb{E}\,\frac{1}{t}\#\{n : F_\eta(n) \le t|f(S_n)|\}\\
&\le 1 + t\sum_{n\ge t}\frac{\mathbb{E}\,|f(S_n)|^2}{F_\eta^2(n)}\\
&\le 1 + Ct\sum_{n\ge t}\frac{1}{|F(\eta n)|^2}\Big\{\sup_{n^h<p\le n}|f(p)|^2 + \Big|\sum_{2\le p\le n^h}\frac{f(p)}{p}\Big|^2 + \sum_{2\le p\le n^h}\frac{|f(p)|^2}{p}\Big\}.
\end{aligned}$$

On using (24) we deduce

$$\sup_{t\ge t_0}\mathbb{E}\,\frac{1}{t}\#\{n : F_\eta(n) \le t|f(S_n)|\} \le C_1$$

with some other constant C_1 which, together with (20), gives $L(t) \ll t$, completing the proof of Theorem 4.

In conclusion we show that for the additive function f defined in the example in Section 1, the weighted SLLN (2) fails for some i.i.d. sequence $X, X_1, X_2, \dots$ with finite means. In this case $A_n \to \alpha = \sum_{k=1}^\infty (-1)^{k-1}/p_k > 0$ and thus $A_n = O(1)$; further $|f(p)| = 1$ for all primes p and $B_n \sim \log\log n$ and thus the Lindeberg condition (8) is satisfied. Thus the CLT (7) is also valid, and it follows that $|f(n)| \gg (\log\log n)^{1/2}$ on a set H of integers with density $\ge 1/2$. Further by (18) we have $|F(n)| \sim \alpha n$ as $n \to \infty$ and thus $|F(n)/f(n)| \ll n/(\log\log n)^{1/2}$ on H. Consequently, for $n \in H$, $n \ge n_0$ the inequality $|F(n)/f(n)| \le t$ is satisfied if $n \le C_1 t(\log\log t)^{1/2}$ for a sufficiently small constant C_1 and thus

$$L(t) = \#\{n : |F(n)/f(n)| \le t\} \gg t(\log\log t)^{1/2}.$$

Therefore by Lemma 1 the weighted SLLN (2) fails for some i.i.d. sequence (X_n) with finite means, as claimed.

References

[1] I. Berkes and M. Weber, A law of the iterated logarithm for arithmetic functions, *Proc. Amer. Math. Soc.*, **135** (2007), 1223–1232.

[2] K. Fukuyama and Y. Komatsu, A law of large numbers for arithmetic functions, *Proc. Amer. Math. Soc.*, **137** (2009), 349–352.

[3] P. D. Elliott, *Probabilistic number theory II*, Springer, New York, 1980.

[4] H. Halberstam, Über additive zahlentheoretische Funktionen, *J. Reine Angew. Math.*, **195** (1955), 210–214.

[5] B. Jamison, S. Orey and W. Pruitt, Convergence of weighted averages of independent random variables, *Z. Wahrscheinlichkeitsth. und Verw. Gebiete*, **4** (1965), 40–44.

[6] J. Kubilius, *Probabilistic methods in the theory of numbers*, Translations of Mathematical Mographs, Vol. 11, Amer. Math. Soc., Providence, R.I., 1964.

[7] M. Weber, Small divisors of Bernoulli sums, *Indag. Math. (N.S.)*, **18** (2007), 281–293.