

An exact asymptotic for the square variation of partial sum processes

Allison Lewko^{a,1} and Mark Lewko^{b,2}

^a*Department of Computer Science, The University of Texas at Austin, USA. E-mail: alewko@cs.columbia.edu*

^b*Department of Mathematics, University of California Los Angeles and Institute for Advanced Study, USA. E-mail: mlewko@math.ucla.edu*

Received 29 March 2012; revised 18 March 2014; accepted 22 March 2014

Abstract. We establish an exact asymptotic formula for the square variation of certain partial sum processes. Let $\{X_i\}$ be a sequence of independent, identically distributed mean zero random variables with finite variance σ^2 and satisfying a moment condition $\mathbb{E}[|X_i|^{2+\delta}] < \infty$ for some $\delta > 0$. If we let \mathcal{P}_N denote the set of all possible partitions of the interval $[N]$ into subintervals, then we have that $\max_{\pi \in \mathcal{P}_N} \sum_{I \in \pi} |\sum_{i \in I} X_i|^2 \sim 2\sigma^2 N \ln \ln(N)$ holds almost surely. This can be viewed as a variational strengthening of the law of the iterated logarithm and refines results of J. Qian on partial sum and empirical processes. When $\delta = 0$, we obtain a weaker ‘in probability’ version of the result.

Résumé. Nous établissons une formule asymptotique exacte pour la variation quadratique de certains processus de sommes partielles. Soit $\{X_i\}$ une suite de variables indépendantes et identiquement distribuées de moyenne nulle et de variance finie σ^2 satisfaisant une condition de moments $\mathbb{E}[|X_i|^{2+\delta}] < \infty$ pour un $\delta > 0$. Soit \mathcal{P}_N l'ensemble de toutes les partitions possibles de l'intervalle $[N]$ en sous-intervalles, alors nous montrons que presque sûrement $\max_{\pi \in \mathcal{P}_N} \sum_{I \in \pi} |\sum_{i \in I} X_i|^2 \sim 2\sigma^2 N \ln \ln(N)$. Ceci peut être interprété comme une amélioration de la loi du logarithme itéré et précise les résultats de J. Qian sur les sommes partielles et les processus empiriques. Quand $\delta = 0$, nous obtenons une version plus faible, en probabilité, de ce résultat.

MSC: 60G50

Keywords: Square variation; Law of the iterated logarithm; Random walks

1. Introduction

Let $\{X_i\}$ be a sequence of independent, identically distributed random variables with mean $\mu < \infty$. The strong law of large numbers asserts that

$$\sum_{i=1}^N X_i \sim N\mu$$

¹Supported by a National Defense Science and Engineering Graduate Fellowship.

²Supported by a NSF Postdoctoral Fellowship, DMS-1204206 and the IAS Fund for Math.

almost surely. Without loss of generality, one can assume that X_i are mean zero by considering $X_i - \mu$. If we further assume a finite variance, that is $\mathbb{E}[|X_i|^2] = \sigma^2 < \infty$, the Hartman–Wintner law of the iterated logarithm [7] gives an exact error estimate for the strong law of large numbers. More precisely,

$$\left| \sum_{i=1}^N X_i \right|^2 \leq (2 + o(1))\sigma^2 N \ln \ln(N), \tag{1}$$

holds a.s., where the constant 2 cannot be replaced by a smaller constant. That is, the quantity $\sum_{i=1}^N X_i$ gets as large/small as $\pm\sqrt{(2 - \varepsilon)\sigma^2 N \ln \ln(N)}$ infinitely often (a.s.).

The purpose of our current work is to prove a more delicate variational asymptotic that refines the law of the iterated logarithm and captures more subtle information about the oscillations of sums of i.i.d. random variables about their expected values. More precisely, we prove the following theorem. We let \mathcal{P}_N denote the set of all possible partitions of the interval/progression of integers $[N] := [1, 2, \dots, N]$ into subintervals, and we consider each $\pi \in \mathcal{P}_N$ as a collection of disjoint intervals, the union of which is $[N]$. We write $I \in \pi$ to denote that the interval $I \subseteq [N]$ belongs to π .

Theorem 1. *Let $\{X_i\}$ be a sequence of independent, identically distributed mean zero random variables with variance σ^2 and satisfying $\mathbb{E}[|X_i|^{2+\delta}] < \infty$ for some $\delta > 0$. Then we have almost surely:*

$$\max_{\pi \in \mathcal{P}_N} \sum_{I \in \pi} \left| \sum_{i \in I} X_i \right|^2 \sim 2\sigma^2 N \ln \ln(N). \tag{2}$$

Choosing the partition π to contain a single interval $J = [1, N]$ immediately recovers (1), the upper bound in the law of the iterated logarithm.

It is unclear if the assumption $\mathbb{E}[|X_i|^{2+\delta}] < \infty$ can be removed. The analysis here should be able to be pushed further to handle a condition of the form $\mathbb{E}[|X_i|^2 \varphi(X_i)] < \infty$ where $\varphi(x)$ is a positive increasing function that grows slower than $|x|^\delta$ for any $\delta > 0$ (this requires analogous refinements of Lemma 9 and Lemma 11), however this will not extend to the case when $\varphi(x)$ is bounded. Note, however, that using Theorem 1 and a truncation argument (similar to that used in Section 3) the inequality

$$\max_{\pi \in \mathcal{P}_N} \sum_{I \in \pi} \left| \sum_{i \in I} X_i \right|^2 \leq C\sigma^2 N \ln \ln(N),$$

with any absolute constant C , is sufficient to establish the exact asymptotic in the general case, $\delta = 0$. Without an auxiliary moment condition, we are able to establish the following weaker ‘in probability’ result.

Theorem 2. *Let $\{X_i\}$ be a sequence of independent, identically distributed mean zero random variables with finite variance σ^2 . We then have that*

$$\frac{\max_{\pi \in \mathcal{P}_N} \sum_{I \in \pi} \left| \sum_{i \in I} X_i \right|^2}{2\sigma^2 N \ln \ln(N)} \xrightarrow{P} 1.$$

Here, \xrightarrow{P} denotes convergence in probability.

For the rest of the paper, we will denote $\{X_i\}_{i=1}^N$ more concisely as $\{X_i\}^N$ and write $\|\{X_i\}^N\|_{V^2}$ for the quantity $\sqrt{\max_{\pi \in \mathcal{P}_N} \sum_{I \in \pi} \left| \sum_{i \in I} X_i \right|^2}$. As indicated by the notation, this expression satisfies the triangle inequality: $\|\{X_i + Y_i\}^N\|_{V^2} \leq \|\{X_i\}^N\|_{V^2} + \|\{Y_i\}^N\|_{V^2}$. (This can be easily verified.)

1.1. Previous work

In [17], Taylor proved an almost sure asymptotic for the path variation of Brownian motion. This is closely related to our main result in the case that X_i are normally distributed. The question of the asymptotic order of $\|\{X_i\}^N\|_{V^2}$

in the case of more general X_i is implicit in work of Dudley [4] and Bretagnolle [1], where related questions about the p -variation of processes are studied. Dudley’s interest in these p -variation norms stems from the fact that they majorize the (more classical) sup norm, but in many cases have nicer differentiability properties.

The most recent result we are aware of concerning our specific question appears in the work of J. Qian [14]. There it is shown that

Theorem 3. *Let $\{X_i\}_{i=1}^\infty$ be i.i.d. mean zero random variables with variance σ^2 . Then for some constant c we have that $\mathbb{P}[\max_{\pi \in \mathcal{P}_N} \sum_{I \in \pi} |\sum_{i \in I} X_i|^2 \leq c\sigma^2 N \ln \ln(N)] \rightarrow 0$ as $N \rightarrow \infty$.*

Theorem 4. *Let $\{X_i\}_{i=1}^\infty$ be i.i.d. mean zero random variables with variance σ^2 and $\mathbb{E}[|X_i|^{2+\delta}] < \infty$ for some $\delta > 0$. For some constant c' we have that $\mathbb{P}[\max_{\pi \in \mathcal{P}_N} \sum_{I \in \pi} |\sum_{i \in I} X_i|^2 \geq c'\sigma^2 N \ln \ln(N)] \rightarrow 0$ as $N \rightarrow \infty$.*

These results have been used in [9], [10], [11], [12] and [13] to show that that certain variation operators which generalize (and majorize) classical maximal operators arising in harmonic analysis and analytic number theory are unbounded on certain L^p spaces.

Notice that Qian’s upper bound does not require a moment condition with $\delta > 0$, but her lower bound does. Our results improve on these by establishing the exact constants as well as improving convergence in probability to almost sure convergence when $\delta > 0$. When only the second moment is finite, we obtain convergence in probability results with the exact asymptotic constant.

The high-level organization of our proof follows Taylor’s strategy of splitting the random variables associated to dyadic intervals into good, medium and bad pieces. Roughly speaking, we then approximate the random variables associated to each of these dyadic intervals by an appropriately normalized Gaussian and an error term. While the ideas from [17] are generally applicable to the Gaussian part, a fair amount of additional analysis is required to control the contribution from the error terms. For instance, very short intervals require a separate argument that has no analog in [17].

1.2. Notation

In our proofs below, we will often fix a positive integer N and use $[N]$ to denote the set of integers $\{1, \dots, N\}$. When we say I is a subinterval of $[N]$, we mean that $I = (i_s, i_e]$ for some real numbers $0 \leq i_s \leq i_e \leq N$. We denote the length of the interval I by $|I|$ (i.e. $|I| = i_e - i_s$). When i_s and i_e are integers, this quantity is equal to the number of positive integers contained in I . When we say $I' \subseteq I$ (I' is a subinterval of I), we mean that $I' = (i'_s, \dots, i'_e]$ for some real numbers i'_s, i'_e satisfying $i_s \leq i'_s \leq i'_e \leq i_e$.

We will consider independent, identically distributed random variables X_1, X_2, \dots . We will routinely use S_ℓ to denote the partial sum $S_\ell = X_1 + \dots + X_\ell$. For an interval I , we use S_I to denote $S_I := \sum_{i \in I} X_i$ (i.e. the partial sum of the variables X_i whose indices i are contained in the interval I). We use \ln to denote the natural logarithm and \log to denote the base 2 logarithm. We use $\exp(x)$ as an alternate notation for e^x .

In our proofs, we will often refer to “constants” whose values depend on δ (and only on δ). We will often not reflect this dependence in our notation. Throughout our proofs, δ should be thought of as a fixed, positive constant.

2. The upper bound

In this section, we prove the following theorem:

Theorem 5. *We let X_1, X_2, \dots denote independent, identically distributed random variables with $\mathbb{E}[X_i] = 0$ and $\text{Var}[X_i] = 1$. We further assume that $\mathbb{E}[|X_i|^{2(1+\delta)}] < \infty$, for some $\delta > 0$. Then, for every $\varepsilon > 0$,*

$$\limsup_{N \rightarrow \infty} \frac{\|\{X_i\}^N\|_{V^2}^2}{N \ln \ln N} \leq 2(1 + \varepsilon) \quad a.s.$$

The proof of this theorem will proceed in several stages. First, we will fix $\varepsilon > 0$ and classify intervals $I \subseteq [N]$ into three disjoint categories, “good,” “medium,” and “bad.” (This same strategy is used in [17].) We say an interval I is “good” if

$$S_I^2 \leq (2 + \varepsilon)|I| \ln \ln N. \tag{3}$$

We say it is “medium” if

$$(2 + \varepsilon)|I| \ln \ln N < S_I^2 \leq B|I| \ln \ln N, \tag{4}$$

and it is “bad” if

$$S_I^2 > B|I| \ln \ln N. \tag{5}$$

The precise value of the parameter B will be chosen later. For now, we simply think of it as a constant depending only on δ .

We will deal with each class of intervals separately. We begin by considering the contribution of the bad intervals to the value of $\|\{X_i\}^N\|_{V^2}^2$.

2.1. The bad intervals

To suitably bound the contribution of the bad intervals, we will begin by essentially reducing the space of allowable partitions. We assume for simplicity that N is a power of two, denoted by $N = 2^n$. We will later argue that our results extend to all positive integers N . When $N = 2^n$, we consider intervals of the form

$$((c - 1)2^i, c2^i], \quad i \in \{0, 1, \dots, n\}, c \in \{1, \dots, 2^{n-i}\}.$$

This gives us $n + 1 = \log(N) + 1$ levels of intervals, where the i th level contains 2^{n-i} disjoint intervals, each of size 2^i . We now augment this family of intervals by adding “half shifts” of each level. More precisely, we also consider intervals of the form

$$((c - 1)2^i + 2^{i-1}, c2^i + 2^{i-1}], \quad i \in \{1, \dots, n - 1\}, c \in \{1, \dots, 2^{n-i} - 1\}.$$

This approximately doubles our total number of intervals. We will call the first family of intervals \mathcal{F} , and the second family of intervals \mathcal{F}_s . \mathcal{F} includes $n + 1$ levels of intervals, while \mathcal{F}_s includes $n - 1$ levels of intervals. Within each family at each level, the intervals are disjoint.

We now show:

Lemma 6. *Let $I' \subseteq [N]$ denote an arbitrary interval. There exists some interval $I \in \mathcal{F} \cup \mathcal{F}_s$ such that $I' \subseteq I$ and $|I| < 4|I'|$.*

Proof. We define i to be the non-negative integer satisfying $2^{i-1} < |I'| \leq 2^i$. If $i \geq n - 1$, then $I := [1, \dots, N] \in \mathcal{F}$ suffices. For $i < n - 1$, we consider the intervals in $\mathcal{F} \cup \mathcal{F}_s$ of length 2^{i+1} . There are two cases: either I' is contained in some $I \in \mathcal{F}$ of size 2^{i+1} (and therefore, we are done), or I' must contain a right endpoint of some $I' \in \mathcal{F}$ of size 2^{i+1} . In other words, I' contains $c2^{i+1}$ for some $c \in \{1, \dots, 2^{n-i-1} - 1\}$. Since $|I'| \leq 2^i$, this implies that $I' \subseteq (c2^{i+1} - 2^i, c2^{i+1} + 2^i)$. We can alternatively express this as:

$$I' \subseteq ((c - 1)2^{i+1} + 2^i, c2^{i+1} + 2^i] \in \mathcal{F}_s.$$

So in both cases, we have an $I \in \mathcal{F} \cup \mathcal{F}_s$ such that $I' \subseteq I$ and $|I| < 4|I'|$. □

For the purpose of bounding the contribution of bad intervals, this allows us to consider only intervals in $\mathcal{F} \cup \mathcal{F}_s$ (to some extent). More precisely, for each interval $I \in \mathcal{F} \cup \mathcal{F}_s$, we will consider the random variable

$$\max_{I' \subseteq I} S_{I'}^2.$$

If some $I' \subseteq I$ of size $|I'| > \frac{1}{4}|I|$ is bad, meaning that $|S_{I'}|^2 > B|I'| \ln \ln N$, then

$$\max_{I' \subseteq I} S_{I'}^2 > \frac{B}{4}|I| \ln \ln N.$$

To enable us to later consider values of N which are not powers of two, we actually consider “badness” with respect to $N/2$ instead of N . More precisely, we let $\mathbb{I}_{\text{Bad}, I}$ denote the indicator variable of the event

$$\max_{I' \subseteq I} S_{I'}^2 > \frac{B}{8}|I| \ln \ln N$$

for a particular interval I . Here we have used the very loose bound that $\ln \ln(N/2) \geq \frac{1}{2} \ln \ln N$, for $N \geq 4$. Note that I can contain a subinterval of size $> \frac{1}{4}|I|$ which is bad with respect to $N/2$ (or anything between $N/2$ and N) only when this event occurs.

It is then clear that the contribution of the bad intervals to the value of $\|\{X_i\}^N\|_{V_2}^2$ is upper bounded by:

$$3 \sum_{I \in \mathcal{F} \cup \mathcal{F}_s} \max_{I' \subseteq I} S_{I'}^2 \cdot \mathbb{I}_{\text{Bad}, I}. \tag{6}$$

To see this, note that each bad interval I' in the partition achieving the maximal value is contained in some $I \in \mathcal{F} \cup \mathcal{F}_s$ such that $\mathbb{I}_{\text{Bad}, I} = 1$. Since the intervals in the maximal partition must be disjoint, each such I will only be associated with at most 3 I' 's. Thus, to control the contribution of the bad intervals, it suffices to prove a suitable upper bound on (6) that holds almost surely. As a shorthand notation, we define the variable $Y_I := \max_{I' \subseteq I} S_{I'}^2$. Then the quantity we need to bound can be written a bit more succinctly as:

$$3 \sum_{I \in \mathcal{F} \cup \mathcal{F}_s} Y_I \cdot \mathbb{I}_{\text{Bad}, I}. \tag{7}$$

To bound (7), we will rely on several standard lemmas.

Lemma 7 (Etemadi’s inequality – Theorem 1 in [5]). *Let X_1, X_2, \dots, X_k denote independent random variables and let $a > 0$. Let $S_\ell := X_1 + \dots + X_\ell$ denote the partial sum. Then:*

$$\mathbb{P}\left[\max_{1 \leq \ell \leq k} |S_\ell| \geq 3a\right] \leq 3 \max_{1 \leq \ell \leq k} \mathbb{P}[|S_\ell| \geq a].$$

Lemma 8 (Doob). *Let $\{M_i\}_{i=1}^L$ be a submartingale taking non-negative real values, and $p > 1$. Then:*

$$\mathbb{E}\left[\left(\max_{1 \leq \ell \leq L} M_\ell\right)^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[M_L^p].$$

Lemma 9 (Rosenthal’s inequality – Theorem 3 in [16]). *Let $2 < p < \infty$. Then there exists a constant K_p depending only on p , so that if X_1, \dots, X_ℓ are independent random variables with $\mathbb{E}[X_i] = 0$ for all i and $\mathbb{E}[|X_i|^p] < \infty$ for all i , then:*

$$\left(\mathbb{E}[|S_\ell|^p]\right)^{1/p} \leq K_p \max\left\{\left(\sum_{i=1}^{\ell} \mathbb{E}[|X_i|^p]\right)^{1/p}, \left(\sum_{i=1}^{\ell} \mathbb{E}[|X_i|^2]\right)^{1/2}\right\}.$$

We now prove:

Lemma 10. *For any interval $I \in \mathcal{F} \cup \mathcal{F}_s$,*

$$\mathbb{E}[|Y_I|^{1+\delta}] \leq C|I|^{1+\delta},$$

where C is a constant depending only on δ .

Proof. For each $I = (i_s, \dots, i_e]$, we define the notation $S_{I,k} := X_{i_s+1} + \dots + X_{i_s+k}$ to denote the partial sum of the first k variables in the interval I (for values of k from 1 to $|I|$). We also define the random variable \tilde{Y}_I as:

$$\tilde{Y}_I := \max_{1 \leq k \leq |I|} S_{I,k}^2.$$

We observe that $\mathbb{E}[|Y_I|^{1+\delta}] \leq 4^{1+\delta} \mathbb{E}[|\tilde{Y}_I|^{1+\delta}]$. To see this, consider an arbitrary interval $I' = (i'_s, \dots, i'_e] \subseteq I = (i_s, \dots, i_e]$, where $i'_s \neq i_s$. We let k_1 denote the number of integers contained in $(i_s, i'_e]$ and we let k_2 denote the number of integers contained in $(i_s, i'_s]$. Then $S_{I,k_1} = S_{I'} + S_{I,k_2}$, so

$$|S_{I'}| \leq 2 \max\{|S_{I,k_1}|, |S_{I,k_2}|\}.$$

This implies $\mathbb{E}[|Y_I|^{1+\delta}] \leq 4^{1+\delta} \mathbb{E}[|\tilde{Y}_I|^{1+\delta}]$. (In fact, it implies the stronger fact that $Y_I \leq 4\tilde{Y}_I$ always holds, and we will use this again later.)

Now, $\{S_{I,k}\}$ is a martingale, so $\{|S_{I,k}|\}$ is a submartingale (by Jensen's inequality). Thus, by Lemma 8:

$$\mathbb{E}[|\tilde{Y}_I|^{1+\delta}] = \mathbb{E}\left[\left(\max_{1 \leq k \leq |I|} |S_{I,k}|\right)^{2(1+\delta)}\right] \leq C' \mathbb{E}[|S_I|^{2(1+\delta)}] \quad (8)$$

for some constant C' depending only on δ .

Applying Lemma 9, we see that:

$$\mathbb{E}[|S_I|^{2(1+\delta)}] \leq K' |I|^{1+\delta}, \quad (9)$$

for some constant K' depending only on δ . Combining (9) with (8) (and recalling that $Y_I \leq 4\tilde{Y}_I$), we have shown:

$$\mathbb{E}[|Y_I|^{1+\delta}] \leq C |I|^{1+\delta}$$

for some constant C depending only on δ . □

Our next goal is to derive a suitable upper bound on the quantity $\mathbb{E}[Y_I \mathbb{I}_{\text{Bad}, I}]$ for every interval I . To do this, we will need one more standard lemma.

Lemma 11 (Berry–Esseen theorem³). *Let X_1, X_2, \dots be independent random variables with $\mathbb{E}[X_i] = 0$, $\mathbb{E}[X_i^2] = 1$, and $\mathbb{E}[|X_i|^{2+\gamma}] \leq M$ for all i , and some $\gamma \in (0, 1]$. Then there exists a universal constant C_γ such that for all positive integers k :*

$$\sup_{-\infty < x < \infty} \left| \mathbb{P}\left[\frac{S_k}{\sqrt{k}} < x\right] - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-y^2/2} dy \right| \leq C_\gamma \left(\frac{M}{k^{\gamma/2}}\right).$$

To upper bound $\mathbb{E}[Y_I \mathbb{I}_{\text{Bad}, I}]$, we begin by applying Hölder's inequality with $p := 1 + \delta$ and q defined so that $\frac{1}{p} + \frac{1}{q} = 1$. This gives us:

$$\mathbb{E}[Y_I \mathbb{I}_{\text{Bad}, I}] = \mathbb{E}[|Y_I| \mathbb{I}_{\text{Bad}, I}] \leq (\mathbb{E}[|Y_I|^{1+\delta}])^{1/(1+\delta)} (\mathbb{E}[\mathbb{I}_{\text{Bad}, I}^q])^{1/q}.$$

Applying Lemma 10, we see that

$$(\mathbb{E}[|Y_I|^{1+\delta}])^{1/(1+\delta)} \leq C^{1/(1+\delta)} |I|.$$

Since $\mathbb{I}_{\text{Bad}, I}$ only takes values in $\{0, 1\}$, we also have

$$\mathbb{E}[\mathbb{I}_{\text{Bad}, I}^q] = \mathbb{E}[\mathbb{I}_{\text{Bad}, I}] = \mathbb{P}[\mathbb{I}_{\text{Bad}, I} = 1].$$

³This can be found in [2], p. 322, for example.

We now consider

$$\mathbb{P}[\mathbb{I}_{\text{Bad}, I} = 1] = \mathbb{P}\left[Y_I > \frac{B}{8}|I| \ln \ln N\right].$$

We recall the definition of \tilde{Y}_I and the fact that $Y_I \leq 4\tilde{Y}_I$ from the proof of Lemma 10. We then have:

$$\mathbb{P}\left[Y_I > \frac{B}{8}|I| \ln \ln N\right] \leq \mathbb{P}\left[\tilde{Y}_I > \frac{B}{32}|I| \ln \ln N\right] = \mathbb{P}\left[\max_{1 \leq k \leq |I|} |S_{I,k}| > \sqrt{\frac{B}{32}}|I|^{1/2}(\ln \ln N)^{1/2}\right].$$

By Lemma 7, this quantity is

$$\leq 3 \max_{1 \leq k \leq |I|} \mathbb{P}\left[|S_{I,k}| \geq \frac{\sqrt{B}}{12\sqrt{2}}|I|^{1/2}(\ln \ln N)^{1/2}\right].$$

We will bound this probability using Chebyshev’s inequality for values of k which are $< |I|^{1/2}$, and using Lemma 11 for larger values of k .

For $k < |I|^{1/2}$, we apply Chebyshev’s inequality to obtain:

$$\mathbb{P}\left[|S_{I,k}| \geq \frac{\sqrt{B}}{12\sqrt{2}}|I|^{1/2}(\ln \ln N)^{1/2}\right] \leq \frac{288\mathbb{E}[|S_{I,k}|^2]}{B|I| \ln \ln N}.$$

We note that $\mathbb{E}[|S_{I,k}|^2] = \mathbb{E}[S_{I,k}^2] = k$ (recall that $S_{I,k}$ is a sum of k independent random variables, each with mean 0 and variance 1). Since $k < |I|^{1/2}$, this gives us:

$$\mathbb{P}\left[|S_{I,k}| \geq \frac{\sqrt{B}}{12\sqrt{2}}|I|^{1/2}(\ln \ln N)^{1/2}\right] \leq \frac{288}{B|I|^{1/2} \ln \ln N}.$$

For $|I|^{1/2} \leq k \leq |I|$, we apply Lemma 11 to obtain:

$$\begin{aligned} \mathbb{P}\left[|S_{I,k}| \geq \frac{\sqrt{B}}{12\sqrt{2}}|I|^{1/2}(\ln \ln N)^{1/2}\right] &\leq \mathbb{P}\left[\frac{|S_{I,k}|}{\sqrt{k}} > \frac{\sqrt{B \ln \ln N}}{12\sqrt{2}}\right] \\ &\leq \frac{1}{\sqrt{2\pi}} \int_{\sqrt{B \ln \ln N}/(12\sqrt{2})}^{\infty} e^{-y^2/2} dy + \frac{D}{k^\delta} \end{aligned}$$

for some constant D depending on δ (we are applying the lemma with $\gamma = 2\delta$, and $\mathbb{E}[|X_i|^{2+2\delta}]$ is a constant).

To bound the integral, we proceed as follows (assuming that N is large enough so that $\frac{\sqrt{B \ln \ln N}}{12} \geq 1$):

$$\int_{\sqrt{B \ln \ln N}/(12\sqrt{2})}^{\infty} e^{-y^2/2} dy \leq \int_{\sqrt{B \ln \ln N}/(12\sqrt{2})}^{\infty} ye^{-y^2/2} dy = -e^{-y^2/2} \Big|_{\sqrt{B \ln \ln N}/(12\sqrt{2})}^{\infty} = \left(\frac{1}{\ln N}\right)^{B/576}.$$

Thus, for $|I|^{1/2} \leq k \leq |I|$, we have shown:

$$\mathbb{P}\left[|S_{I,k}| \geq \frac{\sqrt{B}}{12\sqrt{2}}|I|^{1/2}(\ln \ln N)^{1/2}\right] \leq \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\ln N}\right)^{B/576} + \frac{D}{|I|^{\delta/2}}.$$

We define $\sigma = \min\{1/2, \delta/2\}$. Then, for some constant D' depending on δ and for all k , we have:

$$\mathbb{P}\left[|S_{I,k}| \geq \frac{\sqrt{B}}{12\sqrt{2}}|I|^{1/2}(\ln \ln N)^{1/2}\right] \leq \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\ln N}\right)^{B/576} + \frac{D'}{|I|^\sigma}.$$

Thus,

$$\mathbb{P}[\mathbb{I}_{\text{Bad}, I} = 1] \leq \frac{3}{\sqrt{2\pi}} \left(\frac{1}{\ln N}\right)^{B/288} + \frac{3D'}{|I|^\sigma}.$$

Putting everything together, we have that

$$\mathbb{E}[Y_I \mathbb{I}_{\text{Bad}, I}] \leq C^{1/(1+\delta)} |I| \left(\frac{3}{\sqrt{2\pi}} \left(\frac{1}{\ln N}\right)^{B/576} + \frac{3D'}{|I|^\sigma}\right)^{1-1/(1+\delta)}. \tag{10}$$

Next, we show that the contribution of intervals I satisfying $|I| \geq (\log(N))^d$ to the quantity (7) is not too large, where we define the parameter $d := \frac{2}{\sigma(1-1/(1+\delta))}$. For this, we will use (10) and Kronecker’s lemma.

Lemma 12 (Kronecker’s lemma). *Let a_1, a_2, \dots be a sequence of real numbers such that $a_1 \leq a_2 \leq a_3 \leq \dots$ and $a_j \rightarrow \infty$ as $j \rightarrow \infty$. Then if x_1, x_2, \dots is a sequence of real numbers such that $\sum_{j=1}^k x_j/a_j$ converges,*

$$a_k^{-1} \sum_{j=1}^k x_j \rightarrow 0.$$

We now prove:

Lemma 13. *Let m denote a positive integer. Then:*

$$\frac{1}{2^m \ln \ln(2^m)} \sum_{\substack{I \in \mathcal{F} \cup \mathcal{F}_s \\ |I| \geq (\log(2^m))^d}} Y_I \mathbb{I}_{\text{Bad}, I} \rightarrow 0 \quad \text{a.s. as } m \rightarrow \infty,$$

where the indicator variable $\mathbb{I}_{\text{Bad}, I}$ and $\mathcal{F} \cup \mathcal{F}_s$ are defined with respect to $N = 2^m$.

Proof. For each interval $I \in \mathcal{F} \cup \mathcal{F}_s$, there is a minimal value of n such that $I \subseteq [2^n]$. We order our sum over the I ’s according to their associated values of n (and otherwise arbitrarily): i.e. we first sum terms for I ’s with $n = 2$, then with $n = 3$, and so on, and we only include those I ’s satisfying $|I| \geq (\log(2^n))^d$. (We will ignore the very small number of terms with $n = 1$ for convenience.) We let I_1, I_2, I_3, \dots denote the resulting ordered sequence of all intervals I which are contained in $\mathcal{F} \cup \mathcal{F}_s$ for some value of N (which is a power of 2) and also satisfy $|I| \geq (\log(2^n))^d$. For each such I_i , we define

$$a_i := 2^n \ln \ln(2^n),$$

where n is defined from I_i as above. Since we consider N going to infinity, we get an infinite sequence of I ’s. For a fixed N , we let I_{N_e} denote the final interval $I \subseteq [N]$ appearing in the infinite sequence.

We also define a new indicator variable, $\mathbb{I}_{\text{Bad}, I, n}$, which indicates the event that an interval I is “bad” with respect to the value $N := 2^n$ (where n is defined from I as above). Note that when $\mathbb{I}_{\text{Bad}, I}$ is the indicator for I being “bad” with respect to a larger N , then $\mathbb{I}_{\text{Bad}, I} = 1$ implies that $\mathbb{I}_{\text{Bad}, I, n} = 1$ as well. We then have (for any N which is a power of 2):

$$\frac{1}{N \ln \ln N} \sum_{\substack{I \in \mathcal{F} \cup \mathcal{F}_s \\ |I| \geq (\log(N))^d}} Y_I \mathbb{I}_{\text{Bad}, I} \leq \frac{1}{a_{N_e}} \sum_{i=1}^{N_e} Y_{I_i} \mathbb{I}_{\text{Bad}, I_i, n}.$$

It is a bit easier to work with the variables $Y_I \mathbb{I}_{\text{Bad}, I, n}$ than the original variables $Y_I \mathbb{I}_{\text{Bad}, I}$, since the latter depend on the value of N , and the former do not. Hence we can now think of N going to infinity just in terms of adding more

random variables to our sum, instead of needing to change the definition of all the random variables with each change of N .

Thus, it suffices for us to prove that:

$$\lim_{M \rightarrow \infty} \frac{1}{a_M} \sum_{i=1}^M Y_{I_i} \mathbb{I}_{\text{Bad}, I_i, n} = 0 \quad \text{a.s.}$$

By Lemma 12 (note that $a_M \rightarrow \infty$ as $M \rightarrow \infty$), this follows if the sum

$$\sum_{i=1}^{\infty} \frac{1}{a_i} Y_{I_i} \mathbb{I}_{\text{Bad}, I_i, n}$$

converges almost surely. This in turn follows if:

$$\sum_{i=1}^{\infty} \frac{1}{a_i} \mathbb{E}[|Y_{I_i} \mathbb{I}_{\text{Bad}, I_i, n}|] < \infty. \tag{11}$$

For a fixed n , we will have contributions from intervals of size 2^j for values of j ranging from $d \log n$ to n . Note that since we only consider values of $n \geq 2$, we will have $j \geq 1$. By (10), the sum of the expectations $\mathbb{E}[|Y_I \mathbb{I}_{\text{Bad}, I, n}|]$ for intervals $|I| = 2^j$ with the value of n is at most:

$$C' \cdot 2^n \cdot \left(\left(\frac{1}{n} \right)^{B/288} + \frac{D''}{2^{\sigma j}} \right)^{1-1/(1+\delta)},$$

where D'' and C' are constants depending only on δ . We then see that (11) is dominated by:

$$C' \sum_{n=2}^{\infty} \frac{1}{\ln n} \sum_{j=d \log n}^n \left(\left(\frac{1}{n} \right)^{B/288} + \frac{D''}{2^{\sigma j}} \right)^{1-1/(1+\delta)}. \tag{12}$$

We now note that for any positive real values x, y, γ , we have

$$(x + y)^\gamma \leq (2 \max\{x, y\})^\gamma = 2^\gamma \max\{x^\gamma, y^\gamma\} \leq 2^\gamma (x^\gamma + y^\gamma).$$

Applying this to (12), we see it is

$$\leq C'' \sum_{n=2}^{\infty} \frac{1}{\ln n} \sum_{j=d \log n}^n \left(\left(\frac{1}{n} \right)^{B/576(1-1/(1+\delta))} + \frac{D'''}{2^{\sigma' j}} \right), \tag{13}$$

where C'' , D''' and σ' are constants depending on δ . More specifically, $\sigma' = \sigma(1 - \frac{1}{1+\delta})$.

We split this into two pieces, and first consider the sum:

$$D''' \sum_{n=2}^{\infty} \frac{1}{\ln n} \sum_{j=d \log n}^n \frac{1}{2^{\sigma' j}}.$$

We note that

$$\sum_{j=d \log n}^{\infty} \frac{1}{2^{\sigma' j}} \leq K_{\sigma'} \frac{1}{n^{d\sigma'}}$$

for some constant $K_{\sigma'}$ depending on σ' . Therefore,

$$\sum_{n=2}^{\infty} \frac{1}{\ln n} \sum_{j=d \log n}^n \frac{1}{2^{\sigma' j}} \leq K_{\sigma'} \sum_{n=2}^{\infty} \frac{1}{n^{d\sigma'} \ln n}.$$

Since $d = \frac{2}{\sigma'}$, this sum converges.

Next we consider the sum

$$\sum_{n=2}^{\infty} \frac{1}{\ln n} \cdot n^{-(B/576)(1-1/(1+\delta))} \sum_{j=d \log n}^n 1.$$

Since $\sum_{j=d \log n}^n 1 \leq n$, it suffices to consider

$$\sum_{n=2}^{\infty} \frac{1}{\ln n} \cdot n^{1-(B/576)(1-1/(1+\delta))}.$$

At this point, we choose the value of B so that

$$\frac{B}{576} \left(1 - \frac{1}{1 + \delta}\right) - 1 > 1.$$

This ensures that the sum converges, and the proof of the lemma is complete. □

To conclude our treatment of the bad intervals, we must also show that the contribution of intervals I with $|I| < (\log(N))^d$ is not too large. To do this, we return to considering a fixed value of N (which is a power of 2) and the indicator variables $\mathbb{I}_{\text{Bad}, I}$ are all with respect to this N . For each $I \in \mathcal{F} \cup \mathcal{F}_s$, we define the random variable $Z_I := Y_I \mathbb{I}_{\text{Bad}, I}$. We first consider intervals $I \in \mathcal{F}$ of a fixed size $2^i < (\log(N))^d$. There are $L := N \cdot 2^{-i}$ such intervals, and we denote the associated random variables Z_I as Z_1, \dots, Z_L . We prove:

Lemma 14.

$$\mathbb{P} \left[\left| \sum_{j=1}^L Z_j - \sum_{j=1}^L \mathbb{E}[Z_j] \right| \geq L \right] \leq K(2^i)^{1+2\delta} N^{-\delta},$$

where K is a constant depending only on δ and not on i or N .

Proof. We define $\bar{Z}_j := Z_j \mathbb{I}_{|Z_j| \leq L}$. In other words, \bar{Z}_j is the truncation of the (non-negative) random variable Z_j at the value L . We can then write:

$$\mathbb{P} \left[\left| \sum_{j=1}^L Z_j - \sum_{j=1}^L \mathbb{E}[Z_j] \right| \geq L \right] = \mathbb{P} \left[\left| \sum_{j=1}^L (\bar{Z}_j - \mathbb{E}[\bar{Z}_j]) + \sum_{j=1}^L (Z_j - \bar{Z}_j - \mathbb{E}[Z_j] + \mathbb{E}[\bar{Z}_j]) \right| \geq L \right].$$

We consider

$$\mathbb{E}[Z_j] - \mathbb{E}[\bar{Z}_j] = \int_L^{\infty} \mathbb{P}[Z_j > t] dt + L\mathbb{P}[Z_j \geq L].$$

By Chebyshev's inequality,

$$\mathbb{P}[Z_j > t] \leq \frac{\mathbb{E}[|Z_j|^{1+\delta}]}{t^{1+\delta}}.$$

Inserting this into the previous equality, we obtain:

$$|\mathbb{E}[Z_j] - \mathbb{E}[\bar{Z}_j]| \leq \mathbb{E}[|Z_j|^{1+\delta}] \int_L^\infty t^{-1-\delta} dt = \left(\frac{1}{\delta} + 1\right) \frac{\mathbb{E}[|Z_j|^{1+\delta}]}{L^\delta}.$$

We now see that:

$$\begin{aligned} & \mathbb{P}\left[\left|\sum_{j=1}^L (\bar{Z}_j - \mathbb{E}[\bar{Z}_j]) + \sum_{j=1}^L (Z_j - \bar{Z}_j - \mathbb{E}[Z_j] + \mathbb{E}[\bar{Z}_j])\right| \geq L\right] \\ & \leq L\mathbb{P}[|Z_j| > L] + \mathbb{P}\left[\left|\sum_{j=1}^L (\bar{Z}_j - \mathbb{E}[\bar{Z}_j])\right| \geq L - L\left(\frac{1}{\delta} + 1\right) \frac{\mathbb{E}[|Z_j|^{1+\delta}]}{L^\delta}\right]. \end{aligned}$$

Here, we have applied the union bound, the fact that the Z_j 's are identically distributed, and that $Z_j - \bar{Z}_j = 0$ when $|Z_j| \leq L$.

We will bound these two quantities separately. First, by applying Chebyshev's inequality, we have:

$$L\mathbb{P}[|Z_j| > L] \leq L^{-\delta} \mathbb{E}[|Z_j|^{1+\delta}]. \tag{14}$$

To bound the second quantity, we let $L' := L - L\left(\frac{1}{\delta} + 1\right) \frac{\mathbb{E}[|Z_j|^{1+\delta}]}{L^\delta}$ and we let $\tilde{Z}_j := \bar{Z}_j - \mathbb{E}[\bar{Z}_j]$. By another application of Chebyshev's inequality,

$$\mathbb{P}\left[\left|\sum_{j=1}^L \tilde{Z}_j\right| \geq L'\right] \leq \frac{\mathbb{E}[(\sum_{j=1}^L \tilde{Z}_j)^2]}{(L')^2}.$$

Since the random variables \tilde{Z}_j are independent and mean zero,

$$\mathbb{E}\left[\left(\sum_{j=1}^L \tilde{Z}_j\right)^2\right] = L\mathbb{E}[\tilde{Z}_j^2].$$

We then observe:

$$\mathbb{E}[\tilde{Z}_j^2] = 2 \int_0^\infty x\mathbb{P}[|\tilde{Z}_j| > x] dx \leq 2 + 2 \int_1^\infty x\mathbb{P}[|\tilde{Z}_j| > x] dx. \tag{15}$$

We recall that $\tilde{Z}_j := \bar{Z}_j - \mathbb{E}[\bar{Z}_j]$. Since \bar{Z}_j is a non-negative random variable which is $\leq L$, we have that $|\tilde{Z}_j| \leq L$ as well. Therefore, (15) becomes:

$$\begin{aligned} 2 + 2 \int_1^L x\mathbb{P}[|\tilde{Z}_j| > x] dx & \leq 2 + 2 \int_1^L x \cdot \frac{\mathbb{E}[|\tilde{Z}_j|^{1+\delta}]}{x^{1+\delta}} dx = 2 + 2\mathbb{E}[|\tilde{Z}_j|^{1+\delta}] \int_1^L x^{-\delta} dx \\ & = 2 + \frac{2}{1-\delta} \mathbb{E}[|\tilde{Z}_j|^{1+\delta}] (L^{1-\delta} - 1). \end{aligned}$$

To put this all together and simplify our expressions, we recall that $L := N \cdot 2^{-i}$, where $2^i < (\log N)^d$. Thus, $L > N(\log N)^{-d}$. This is very large compared to the value of $\mathbb{E}[|Z_j|^{1+\delta}]$, which is $\leq \mathbb{E}[|Y_I|^{1+\delta}]$ for the associated interval I . Recall from Lemma 10 that $\mathbb{E}[|Y_I|^{1+\delta}]$ is at most $C|I|^{1+\delta}$, and $|I| = 2^i < (\log N)^d$. This means that for sufficiently large N , we can loosely bound L' as:

$$L' = L - L\left(\frac{1}{\delta} + 1\right) \frac{\mathbb{E}[|Z_j|^{1+\delta}]}{L^\delta} > \frac{L}{2}.$$

We then have:

$$\mathbb{P}\left[\left|\sum_{j=1}^L \tilde{Z}_j\right| \geq L'\right] \leq \frac{\mathbb{E}[(\sum_{j=1}^L \tilde{Z}_j)^2]}{(L')^2} \leq \frac{4\mathbb{E}[(\sum_{j=1}^L \tilde{Z}_j)^2]}{L^2} \leq K'\mathbb{E}[|\tilde{Z}_j|^{1+\delta}]L^{-\delta},$$

for some constant K' depending on δ . Combining this with (14), we have

$$\mathbb{P}\left[\left|\sum_{j=1}^L Z_j - \sum_{j=1}^L \mathbb{E}[Z_j]\right| \geq L\right] \leq K''(\max\{\mathbb{E}[|\tilde{Z}_j|^{1+\delta}], \mathbb{E}[|Z_j|^{1+\delta}]\})L^{-\delta},$$

where $K'' := K' + 1$.

Now we consider the quantity $\mathbb{E}[|\tilde{Z}_j|^{1+\delta}]$. We note that:

$$\mathbb{E}[|\tilde{Z}_j|^{1+\delta}] = \mathbb{E}[|\bar{Z}_j - \mathbb{E}[\bar{Z}_j]|^{1+\delta}].$$

Since \bar{Z}_j is a non-negative random variable, $|\bar{Z}_j - \mathbb{E}[\bar{Z}_j]| \leq \max\{\bar{Z}_j, \mathbb{E}[\bar{Z}_j]\}$. Then, $(\max\{\bar{Z}_j, \mathbb{E}[\bar{Z}_j]\})^{1+\delta} \leq \bar{Z}_j^{1+\delta} + (\mathbb{E}[\bar{Z}_j])^{1+\delta}$. Since $g(x) = x^{1+\delta}$ is a convex function on $[0, \infty)$, Jensen's inequality implies that $(\mathbb{E}[\bar{Z}_j])^{1+\delta} \leq \mathbb{E}[\bar{Z}_j^{1+\delta}]$. Therefore,

$$\mathbb{E}[|\tilde{Z}_j|^{1+\delta}] \leq 2\mathbb{E}[\bar{Z}_j^{1+\delta}] \leq 2\mathbb{E}[Z_j^{1+\delta}].$$

Since $Z_j \leq Y_I$ for the associated interval I , we have:

$$\mathbb{P}\left[\left|\sum_{j=1}^L Z_j - \sum_{j=1}^L \mathbb{E}[Z_j]\right| \geq L\right] \leq 2K''\mathbb{E}[|Y_I|^{1+\delta}]L^{-\delta} \leq K|I|^{1+\delta}L^{-\delta},$$

for some constant K depending on δ , by Lemma 10. Since $L := N/|I|$, we can rewrite this as $K|I|^{1+2\delta}N^{-\delta}$. Recalling that $|I| = 2^i$, we have:

$$\mathbb{P}\left[\left|\sum_{j=1}^L Z_j - \sum_{j=1}^L \mathbb{E}[Z_j]\right| \geq L\right] \leq K(2^i)^{1+2\delta}N^{-\delta}.$$

□

Now, we fix N and consider summing these error probabilities in Lemma 14 for all i such that $2^i < (\log(N))^d$. We also fix a value δ' such that $0 < \delta' < \delta$. Since the number of such i and all of the terms except $N^{-\delta}$ are polylogarithmic in N , we get that, for all N sufficiently large with respect to δ (and δ, δ'):

$$\sum_{i=0}^{d \log \log N} \mathbb{P}\left[\left|\sum_{\substack{I \in \mathcal{F} \\ |I|=2^i}} Y_I \mathbb{1}_{\text{Bad}, I} - \sum_{\substack{I \in \mathcal{F} \\ |I|=2^i}} \mathbb{E}[Y_I \mathbb{1}_{\text{Bad}, I}]\right| \geq N \cdot 2^{-i}\right] \leq N^{-\delta'}. \tag{16}$$

We will refer to the levels of \mathcal{F} and \mathcal{F}_s with $2^i < (\log N)^d$ as the “low” levels. The left hand side of (16) is an upper bound on the probability of the contribution of *any* low level of \mathcal{F} to the quantity (7) exceeding its expectation by more than $N \cdot 2^{-i}$. The very same argument can be applied to the low levels of \mathcal{F}_s . Since we are considering only values of N which are powers of 2, and

$$\sum_{n=1}^{\infty} 2^{-n\delta'} < \infty,$$

we can apply the Borel–Cantelli lemma to conclude that almost surely, only finitely many values of N will have a low level which contributes more than $N \cdot 2^{-i}$ plus its expected contribution.

When the contributions of all the low levels obey this bound, we have:

$$\sum_{\substack{I \in \mathcal{F} \cup \mathcal{F}_s \\ |I| < (\log N)^d}} Y_I \mathbb{I}_{\text{Bad}, I} \leq \sum_{i=0}^{d \log \log N} N \cdot 2^{-i} + \sum_{\substack{I \in \mathcal{F} \cup \mathcal{F}_s \\ |I| < (\log N)^d}} \mathbb{E}[Y_I \mathbb{I}_{\text{Bad}, I}]. \tag{17}$$

We observe:

$$\sum_{i=0}^{d \log \log N} N \cdot 2^{-i} < N \sum_{i=0}^{\infty} 2^{-i} = 2N.$$

We will bound the second quantity using (10). Recalling that B was fixed so that $\frac{B}{576}(1 - \frac{1}{1+\delta}) > 2$, (10) implies:

$$\mathbb{E}[Y_I \mathbb{I}_{\text{Bad}, I}] \leq C'' |I| \left(\frac{1}{\ln^2 N} + \frac{D'''}{|I|^{\sigma'}} \right),$$

where C'' , D''' , d , σ' are constants depending on δ , δ . We note that this is merely a restatement of (13). Thus, we have:

$$\sum_{\substack{I \in \mathcal{F} \cup \mathcal{F}_s \\ |I| < (\log N)^d}} \mathbb{E}[Y_I \mathbb{I}_{\text{Bad}, I}] \leq 2C'' N \sum_{i=0}^{d \log \log N} \frac{1}{\ln^2 N} + \frac{D'''}{2^{i\sigma'}}. \tag{18}$$

Next, we observe that

$$2C'' N \sum_{i=0}^{d \log \log N} \frac{1}{\ln^2 N} \leq C''' \frac{N(\ln \ln N)}{\ln^2 N},$$

for some constant C''' depending on δ . Finally, we note that

$$\sum_{i=0}^{\infty} \frac{1}{2^{i\sigma'}} < \infty.$$

Putting these results together with Lemma 13, we have proven that quantity (7) divided by $N \ln \ln N$ goes to zero as N goes to infinity, for N 's which are powers of 2. To achieve a result for all values of N , we consider an N which is an arbitrary positive integer, and let N' the smallest power of 2 such that $N \leq N'$. Then $N' < 2N$, and $N' \ln \ln N' < 3N \ln \ln N$ for instance (for all but very small N). Next, we claim that the contribution of the bad intervals to $\|\{X_i\}^N\|_{V_2}^2$ is bounded by quantity (7) for N' . To see this, note that any bad interval in the maximal partition of $[N]$ will fall in an interval $I \in \mathcal{F} \cup \mathcal{F}_s$ for N' (of size less than 4 times the size of the bad interval), and this I must have $\mathbb{I}_{\text{Bad}, I} = 1$ (recall that we defined these indicator variables to detect “badness” with respect to $N'/2 < N$ for any subintervals of sufficient size). Thus, we have proven:

Theorem 15. *For a positive integer N , we let \mathcal{B}_N denote the contribution of intervals I such that $S_I^2 > B|I| \ln \ln N$ to the quantity $\|\{X_i\}^N\|_{V_2}^2$. Then:*

$$\frac{\mathcal{B}_N}{N \ln \ln N} \rightarrow 0 \quad \text{a.s. as } N \rightarrow \infty.$$

This concludes our treatment of the bad intervals.

2.2. *The medium intervals*

We first reduce to the case of bounded X_i 's (i.e. $|X_i| \leq M$ for some constant M). We consider the truncation of an unbounded X_i , denoted by $\bar{X}_i := X_i \mathbb{1}_{|X_i| \leq M}$. We define $Z_i := X_i - \bar{X}_i$. We then have:

$$\|\{X_i\}^N\|_{V^2}^2 = \|\{(\bar{X}_i - \mathbb{E}[\bar{X}_i]) + (Z_i - \mathbb{E}[Z_i])\}^N\|_{V^2}^2.$$

By the triangle inequality for the ℓ^2 norm, this is: $\leq \|\{\bar{X}_i - \mathbb{E}[\bar{X}_i]\}^N\|_{V^2} + \|\{Z_i - \mathbb{E}[Z_i]\}^N\|_{V^2}$.

Now, $\frac{Z_i - \mathbb{E}[Z_i]}{\sqrt{\text{Var}[Z_i]}}$ are identically distributed random variables with expectation equal to 0, variance equal to 1, and finite $2(1 + \delta)$ moment. We can therefore apply Theorem 15 to conclude:

$$\limsup_{N \rightarrow \infty} \frac{\|\{Z_i - \mathbb{E}[Z_i]\}^N\|_{V^2}^2}{N \ln \ln N} \leq B \text{Var}[Z_i] \quad \text{a.s.}$$

We recall that B is a constant depending only on δ . Since $\mathbb{E}[X_i^2] = 1$, we can choose M sufficiently large so that $B \text{Var}[Z_i]$ is much smaller than ε , say $< \varepsilon/2$ (and $\text{Var}[\bar{X}_i]$ is very close to 1). It then suffices to bound the contribution of the medium intervals for the variables $\frac{\bar{X}_i - \mathbb{E}[\bar{X}_i]}{\sqrt{\text{Var}[\bar{X}_i]}}$ as $\leq 2(1 + \varepsilon/2)N \ln \ln N$. These variables are bounded, identically distributed, mean zero, variance 1, and have finite $2(1 + \delta)$ moment.

To control the contribution of the medium intervals for bounded variables, we will first choose parameters $\varepsilon_0, \varepsilon_1 > 0$ such that $\varepsilon_0 < \varepsilon, \varepsilon_1^2 > \varepsilon_0$, and:

$$(1 + \varepsilon)(1 + \varepsilon_0)^{-1}(1 - \varepsilon_1)^2 > 1. \tag{19}$$

The reason for this constraint will become clear later. We also define ε' by $(1 + \varepsilon')^2 = 1 + \varepsilon_0$. We let M denote the bound on the absolute values of the X_i 's. For a fixed positive integer N , we let $N' = (1 + \varepsilon')^n$ denote the real number which is the smallest integral power of $1 + \varepsilon'$ that is $\geq N$. We now define a family of real intervals which we will denote by \mathcal{H} . This will be similar to $\mathcal{F} \cup \mathcal{F}_s$, the family of intervals we considered in our analysis of the bad intervals, but now our interval lengths will be powers of $1 + \varepsilon'$ instead of powers of two.

\mathcal{H} will be a union of several families of intervals, denoted $\mathcal{H}_0, \dots, \mathcal{H}_h$, where h is a function of ε' . \mathcal{H}_0 consists of all intervals of the form:

$$((c - 1)(1 + \varepsilon')^i, c(1 + \varepsilon')^i], \quad i \in \{0, 1, \dots, n\}, c \in \mathbb{Z} \cap [1, (1 + \varepsilon')^{n-i}].$$

For $j \neq 0$, \mathcal{H}_j consists of all intervals of the form:

$$\begin{aligned} &((c - 1)(1 + \varepsilon')^i + j\varepsilon'(1 + \varepsilon')^{i-1}, c(1 + \varepsilon')^i + j\varepsilon'(1 + \varepsilon')^{i-1}], \\ &i \in \{1, \dots, n - 1\}, c \in \mathbb{Z} \cap [1, (1 + \varepsilon')^{n-i} - 1]. \end{aligned}$$

We will refer to intervals of size $(1 + \varepsilon')^i$ as belonging to level i in each \mathcal{H}_j . We set h (the maximum value of j) to be $\frac{1+\varepsilon'}{\varepsilon'} - 1$. Essentially, we are taking the intervals in \mathcal{H}_0 and shifting them by multiples of $\varepsilon'(1 + \varepsilon')^{i-1}$ and stopping when we would reach the next interval. We now prove:

Lemma 16. *Let $I' \subseteq [N]$ denote an arbitrary interval. There exists some interval $I \in \mathcal{H}$ such that $I' \subseteq I$ and $|I| < (1 + \varepsilon')^2|I'|$.*

Proof. We proceed similarly to the proof of Lemma 6. We define the integer i by the inequality $(1 + \varepsilon')^{i-1} < |I'| \leq (1 + \varepsilon')^i$. If $i \geq n - 1$, then $I = (0, (1 + \varepsilon')^n]$ suffices. Otherwise, we consider intervals in \mathcal{H} of size $(1 + \varepsilon')^{i+1}$. There exist integers c, k such that the leftmost endpoint of I is $> c(1 + \varepsilon')^{i+1} + k\varepsilon'(1 + \varepsilon')^i$ and $\leq c(1 + \varepsilon')^{i+1} + (k + 1)\varepsilon'(1 + \varepsilon')^i$ and $c \in [0, (1 + \varepsilon')^{n-i-1} - 1], k < \frac{1+\varepsilon'}{\varepsilon}$. Since $|I'| \leq (1 + \varepsilon')^i$, this implies that:

$$I' \subseteq (c(1 + \varepsilon')^{i+1} + k\varepsilon'(1 + \varepsilon')^i, c(1 + \varepsilon')^{i+1} + (k + 1)\varepsilon'(1 + \varepsilon')^i + (1 + \varepsilon')^i].$$

We can rewrite the containing interval as:

$$I := (c(1 + \varepsilon')^{i+1} + k\varepsilon'(1 + \varepsilon')^i, (c + 1)(1 + \varepsilon')^{i+1} + k\varepsilon'(1 + \varepsilon')^i] \in \mathcal{H}_k.$$

This I satisfies $|I| = (1 + \varepsilon')^{i+1} < (1 + \varepsilon')^2|I'|$, since $|I'| > (1 + \varepsilon')^{i-1}$. □

We let $N'' := (1 + \varepsilon')^{n-1}$. We will refer to an interval $I \in \mathcal{H}$ as “medium” if

$$\max_{\substack{I' \subseteq I \\ |I'| > (1 + \varepsilon_0)^{-1}|I|}} S_{I'}^2 > 2(1 + \varepsilon)(1 + \varepsilon_0)^{-1}|I| \ln \ln N''.$$

We note that $I \in \mathcal{H}$ can contain a subinterval I' of size $|I'| > (1 + \varepsilon')^{-2}|I|$ that is medium (in the sense of (4)) with respect to some $N'' < N \leq N'$ only when I satisfies this new condition for being “medium.” We let $\mathbb{I}_{I, \text{Med}}$ be the indicator variable for the event that $I \in \mathcal{H}$ is “medium.” Now, for any N between N'' and N' , the total length of the medium intervals (in the sense of (4)) in the maximal partition for N is $\leq \sum_{I \in \mathcal{H}} |I| \cdot \mathbb{I}_{I, \text{Med}}$. We will upper bound this quantity. We begin by upper bounding $\mathbb{P}[\mathbb{I}_{I, \text{Med}} = 1]$, for which we employ the following lemma.

Lemma 17. *Let $\{X_i\}$ be a sequence of independent mean 0 random variables such that $|X_i| \leq M \forall i$. Then*

$$\mathbb{P}\left[\max_{1 \leq l \leq L} \left| \sum_{i=1}^l X_i \right| > t\right] \leq C e^{-t^2/2/(\sum_{i=1}^L \mathbb{E}X_i^2 + Mt/3)}$$

for some absolute constant C .

While this statement is certainly not new (all of the essential ideas are contained in Hoeffding’s paper [8]), we have been unable to locate a reference for this precise formulation, so we have included a proof in the [Appendix](#). The key ingredient in obtaining this maximal form is Doob’s maximal inequality for martingales.

Corollary 18. *Let $\{X_i\}$ be a sequence of independent mean 0 random variables with $\text{Var}[X_i] = 1$ and $|X_i| \leq M$ for all i . Then*

$$\mathbb{P}\left[\max_{\substack{I' \subseteq [L] \\ |I'| > (1 + \varepsilon_0)^{-1}L}} |S_{I'}| > t\right] \leq C' e^{-t^2(1 - \varepsilon_1)^2/2/(L + Mt\varepsilon_1/3\varepsilon_0)}$$

for some absolute constant C' .

Proof. We apply the union bound to conclude:

$$\mathbb{P}\left[\max_{\substack{I' \subseteq [L] \\ |I'| > (1 + \varepsilon_0)^{-1}L}} |S_{I'}| > t\right] \leq \mathbb{P}\left[\max_{l \leq L} |S_l| > (1 - \varepsilon_1)t\right] + \mathbb{P}[|S_{\varepsilon_0 L}| > \varepsilon_1 t].$$

To see this, note that any subinterval $I' \subseteq [L]$ of size $> (1 + \varepsilon_0)^{-1}L$ must have its left endpoint be $< (1 - \frac{1}{1 + \varepsilon_0})L < \varepsilon_0 L$. By Lemma 17, this is:

$$\leq C \exp\left(-\frac{t^2(1 - \varepsilon_1)^2}{2(L + Mt(1 - \varepsilon_1)/3)}\right) + C \exp\left(-\frac{t^2\varepsilon_1^2}{2(\varepsilon_0 L + Mt\varepsilon_1/3)}\right).$$

Recalling that $\varepsilon_1^2 > \varepsilon_0$, this is:

$$\leq 2C \exp\left(-\frac{t^2(1 - \varepsilon_1)^2}{2(L + Mt\varepsilon_1/3\varepsilon_0)}\right).$$

□

Since our variables X_i are bounded in absolute value by M , the maximum of $S_{I'}^2$ for subintervals $I' \subseteq I$ is always bounded as $\leq (M|I|)^2$. Thus, for intervals I which are too small with respect to N'' , I cannot possibly be medium. More specifically, $\mathbb{I}_{I,\text{Med}}$ can only be equal to 1 when $|I| > \frac{2(1+\varepsilon)}{M^2(1+\varepsilon_0)} \ln \ln N''$. Thus, we can assume N (and hence N'') is sufficiently large so that applying Corollary 18 yields:

$$\mathbb{P}[\mathbb{I}_{I,\text{Med}} = 1] \leq C' \exp(-(1 + \varepsilon_2) \ln \ln N'') = C' (\ln N'')^{-(1+\varepsilon_2)}, \tag{20}$$

for some positive ε_2 . To see this, note that we are applying the corollary with the value

$$t = \sqrt{2(1 + \varepsilon)(1 + \varepsilon_0)^{-1}|I| \ln \ln N''},$$

and by (19), $(1 + \varepsilon)(1 + \varepsilon_0)^{-1}(1 - \varepsilon_1)^2 > 1$. We consider N'' large enough so that L dominates the term $Mt\varepsilon_1/3\varepsilon_0$ in the denominator. Here, L is the number of integers in the interval I , which is asymptotically equal to $|I|$ (the length of the real interval I).

We consider intervals in level i of \mathcal{H}_j (where i is large enough so that these intervals can be medium). We let k_{ij} denote the number of such intervals, and $k_{ij} \sim (1 + \varepsilon')^{n-i}$. We note that the indicator random variables of these intervals $\mathbb{I}_{I,\text{Med}}$ are independent, because the intervals are disjoint. We will let ℓ_{ij} denote the total length of all the medium intervals in level i of H_j . By a Chernoff bound:

$$\mathbb{P}\left[\sum_{\substack{I \in H_j \\ \text{level } i}} \mathbb{I}_{I,\text{Med}} \geq 2C'k_{ij}(\ln N'')^{-(1+\varepsilon_2)}\right] \leq \exp(-C'k_{ij}(\ln N'')^{-(1+\varepsilon_2)}/3). \tag{21}$$

When the event

$$\sum_{\substack{I \in H_j \\ \text{level } i}} \mathbb{I}_{I,\text{Med}} \leq 2C'k_{ij}(\ln N'')^{-(1+\varepsilon_2)}$$

occurs, we have $\ell_{ij} \leq 2C'N'(\ln N'')^{-(1+\varepsilon_2)}$. We will call this event E_{ij} .

Since $k_{ij} \sim (1 + \varepsilon')^{n-i}$ and $\ln N'' \sim n$, we can choose a constant d large enough so that:

$$\sum_n \sum_{i < n - d \ln n} \exp(-C'k_{ij}(\ln N'')^{-(1+\varepsilon_2)}/3) \leq \sum_n \sum_{i < n - d \ln n} \exp(-C''(1 + \varepsilon')^{n-i}n^{-1-\varepsilon_2}) < \infty.$$

By the Borel–Cantelli lemma, we then have that almost surely, *all* of the events E_{ij} (for $i < n - d \ln n$) occur when n is sufficiently large (i.e. N' is sufficiently large).

To address the medium intervals in \mathcal{H}_j for levels with $i \geq n - d \ln n$, we note there are $d \ln n \sim d \ln \ln N'$ such levels, and by (20), the expected length of the medium intervals on each level is at most $C'N'(\ln N'')^{-(1+\varepsilon_2)}$. Therefore, for each such i , by Markov’s inequality (for $\varepsilon_3 < \varepsilon_2$):

$$\mathbb{P}[\ell_{ij} \geq C'N'(\ln N'')^{-\varepsilon_2+\varepsilon_3}] \leq \frac{1}{(\ln N'')^{1+\varepsilon_3}}.$$

Now, the quantity $\frac{d \ln n}{(\ln N'')^{1+\varepsilon_3}} \sim \frac{d \ln n}{(n \ln(1+\varepsilon'))^{1+\varepsilon_3}}$ converges when we sum over n . Hence, another application of the Borel–Cantelli lemma tells us that, almost surely, for n sufficiently large we will have $\ell_{ij} \leq C'N'(\ln N'')^{-\varepsilon_2+\varepsilon_3}$ for all $i \geq n - d \ln n$.

Putting everything together, we have that almost surely, for sufficiently large n :

$$\sum_{i,j} \ell_{ij} \leq \frac{2C'h}{\ln(1 + \varepsilon')} N' \ln N' (\ln N'')^{-(1+\varepsilon_2)} + C'h(d \ln n)N'(\ln N'')^{-\varepsilon_2+\varepsilon_3}.$$

Since $\ln N'' \sim \ln N'$ and $\ln n \sim \ln \ln N'$, we see that this entire quantity is $o(N)$. Thus, we have proven:

Theorem 19. *Almost surely, for sufficiently large N , the length of the intervals in the maximal partition for N which are medium in the sense of (4) is $o(N)$.*

This completes our proof of Theorem 5, since the total contribution of the medium intervals is at most $B \ln \ln N$ times the length of the medium intervals, and the contribution of the good intervals is at most $(2 + \varepsilon)N \ln \ln N$.

3. The lower bound

We now prove the following theorem:

Theorem 20. *We let X_1, X_2, \dots denote independent, identically distributed random variables with $\mathbb{E}[X_i] = 0$, $\text{Var}[X_i] = 1$, and $\mathbb{E}[|X_i|^{2(1+\delta)}] < \infty$ for some $\delta > 0$. Then, for every $\varepsilon > 0$,*

$$\liminf_{N \rightarrow \infty} \frac{\|\{X_i\}^N\|_{V^2}^2}{N \ln \ln N} \geq 2(1 - \varepsilon) \quad \text{a.s.}$$

We first argue that it suffices to prove this theorem when the random variables X_i are bounded (i.e. $|X_i| \leq M$ for some constant M). To see why, we again consider the truncation of an unbounded X_i , denoted by $\bar{X}_i := X_i \mathbb{1}_{|X_i| \leq M}$, and we define $Z_i := X_i - \bar{X}_i$. We fix values $\varepsilon_1, \varepsilon_2$ sufficiently small and a value of M sufficiently high such that:

$$\sqrt{2(1 - \varepsilon_2) \text{Var}[\bar{X}_i]} - \sqrt{2(1 + \varepsilon_1) \text{Var}[Z_i]} \geq \sqrt{2(1 - \varepsilon)}. \tag{22}$$

There exists such a choice of M, ε_1 , and ε_2 because $\mathbb{E}[X_i^2] < \infty$, so choosing M sufficiently large will make $\text{Var}[\bar{X}_i]$ sufficiently close to 1 and $\text{Var}[Z_i]$ sufficiently close to 0.

Now, $\frac{Z_i - \mathbb{E}[Z_i]}{\sqrt{\text{Var}[Z_i]}}$ is a random variable with expectation equal to 0, variance equal to 1, and finite $2(1 + \delta)$ moment. We can hence apply Theorem 5 with ε_1 to conclude that:

$$\limsup_{N \rightarrow \infty} \frac{\|\{Z_i - \mathbb{E}[Z_i]\}^N\|_{V^2}^2}{N \ln \ln N} \leq 2(1 + \varepsilon_1) \text{Var}[Z_i] \quad \text{a.s.} \tag{23}$$

We note that $\|\{\bar{X}_i - \mathbb{E}[\bar{X}_i]\}^N\|_{V^2}^2 = \|\{X_i - (Z_i - \mathbb{E}[Z_i])\}^N\|_{V^2}^2$. (Here, we have used the fact that $\mathbb{E}[X_i] = 0$, so $\mathbb{E}[\bar{X}_i] = -\mathbb{E}[Z_i]$.) Employing the triangle inequality, we have:

$$\|\{X_i\}^N\|_{V^2} \geq \|\{\bar{X}_i - \mathbb{E}[\bar{X}_i]\}^N\|_{V^2} - \|\{Z_i - \mathbb{E}[Z_i]\}^N\|_{V^2}. \tag{24}$$

If we prove Theorem 20 for bounded variables, we can apply it to the $\bar{X}_i - \mathbb{E}[\bar{X}_i]$'s with ε_2 . Note that these are mean zero random variables with finite variance and finite $2(1 + \delta)$ moment, and we can divide them by $\sqrt{\text{Var}[\bar{X}_i]}$ to make them have variance equal to one. This gives us:

$$\liminf_{N \rightarrow \infty} \frac{\|\{\bar{X}_i\}^N\|_{V^2}}{\sqrt{N \ln \ln N}} \geq \sqrt{2(1 - \varepsilon_2) \text{Var}[\bar{X}_i]} \quad \text{a.s.}$$

Combining this with (24), (23), and (22), we have that

$$\liminf_{N \rightarrow \infty} \frac{\|\{X_i\}^N\|_{V^2}^2}{N \ln \ln N} \geq 2(1 - \varepsilon) \quad \text{a.s.}$$

Therefore, we may assume from this point forward that the X_i 's are bounded in absolute value by a constant M .

To prove the theorem for bounded variables X_i , we will use an inequality proven in [6] and a general strategy motivated by [17]. We begin by following the approach in [6] for proving the lower bound portion of the law of the iterated logarithm for bounded random variables. We fix a value $0 < \alpha < 1/2$ and a parameter ε_3 whose value will be set later.

We consider the sequence of integers m_1, m_2, \dots defined by $m_k := s^k$, for some suitably large integer $s > 1$. As usual, we let S_{m_k} denote the sum $X_1 + \dots + X_{m_k}$. We note the following inequality from [6], p. 282:

Lemma 21. *For every $\varepsilon' > 0$, when s is sufficiently large with respect to ε' , there exists $0 < \gamma < 1$ such that for all sufficiently large k ,*

$$\mathbb{P}[S_{m_k} - S_{m_{k-1}} \geq (1 - \varepsilon')\sqrt{2(m_k - m_{k-1}) \ln \ln m_k}] \geq \frac{1}{k^\gamma}.$$

We fix a value of s and a value ε_4 such that:

$$(1 - \varepsilon_4)\sqrt{1 - 1/s} > \sqrt{1 - \varepsilon_3}, \tag{25}$$

and s is sufficiently large with respect to ε_4 to apply Lemma 21. We consider values of N which are powers of s , i.e. $N = s^n$ for some integer n .

We now prove:

Lemma 22. *For $N = s^n$ sufficiently large and for each fixed $j \in [N]$,*

$$\mathbb{P}\left[\sup_{1 \leq i_1 \leq i_2 \leq N^{1-\alpha}} \frac{(S_{i_1+j} - S_j)^2 + (S_{i_2+j} - S_{i_1+j})^2}{i_2} < 2(1 - \varepsilon_3) \ln \ln N\right] \leq \exp(-cn^{1-\gamma}), \tag{26}$$

for some constants $c, \gamma > 0$ independent of n with $\gamma < 1$.

Proof. We consider values of i_2 which are $= s^k$ for $n/2 \leq k \leq n(1 - \alpha)$ (i.e. $i_2 = m_k$). For each such k , Lemma 21 implies that:

$$\mathbb{P}[S_{j+m_k} - S_{j+m_{k-1}} > (1 - \varepsilon_4)\sqrt{2(m_k - m_{k-1}) \ln \ln m_k}] \geq \frac{1}{k^\gamma}. \tag{27}$$

Now, we suppose that for some $n/2 \leq k \leq n(1 - \alpha)$, we have

$$S_{j+m_k} - S_{j+m_{k-1}} > (1 - \varepsilon_4)\sqrt{2(m_k - m_{k-1}) \ln \ln m_k}.$$

We will call this event E_k , and we denote its complement by \bar{E}_k . When E_k occurs, we consider $i_2 := m_k$ and $i_1 := m_{k-1}$. Note that $i_2 \leq N^{1-\alpha}$. Then:

$$(S_{i_1+j} - S_j)^2 + (S_{i_2+j} - S_{i_1+j})^2 \geq 2(1 - \varepsilon_4)^2(m_k - m_{k-1}) \ln \ln m_k.$$

Using (25), $m_k - m_{k-1} = s^k(1 - 1/s)$, and $k \geq n/2$, this is $> 2(1 - \varepsilon_3)s^k \ln \ln s^k$. In fact, since $\ln \ln N = \ln \ln s^n$ and $\ln \ln s^{n/2} = \ln \ln s^n + \ln 1/2$, for sufficiently large N we have:

$$(S_{i_1+j} - S_j)^2 + (S_{i_2+j} - S_{i_1+j})^2 \geq 2(1 - \varepsilon_3)i_2 \ln \ln N.$$

Therefore, the probability on the left hand side of (26) is at most the probability that the event E_k fails to occur for all $n/2 \leq k \leq n(1 - \alpha)$. Observe that these events for different k 's are independent, because they involve disjoint sets of the variables X_i . Thus, by Lemma 21,

$$\mathbb{P}[\bar{E}_k \ \forall k \in [n/2, (1 - \alpha)n]] \leq \prod_{k=\lceil n/2 \rceil}^{\lfloor n(1-\alpha) \rfloor} \left(1 - \frac{1}{k^\gamma}\right) \leq \prod_{k=\lceil n/2 \rceil}^{\lfloor n(1-\alpha) \rfloor} e^{-1/k^\gamma},$$

since for positive real numbers x , $1 - x \leq e^{-x}$. This can be rewritten as:

$$\exp\left(-\sum_{k=\lceil n/2 \rceil}^{\lfloor n(1-\alpha) \rfloor} \frac{1}{k^\gamma}\right).$$

We observe that

$$\sum_{k=\lceil n/2 \rceil}^{\lfloor n(1-\alpha) \rfloor} \frac{1}{k^\gamma} \geq \frac{n(1-\alpha-1/2)-2}{n^\gamma(1-\alpha)^\gamma}.$$

Thus, we can choose a constant c independent of n such that

$$\sum_{k=\lceil n/2 \rceil}^{\lfloor n(1-\alpha) \rfloor} \frac{1}{k^\gamma} \geq cn^{1-\gamma}.$$

Therefore, we have shown that:

$$\mathbb{P} \left[\sup_{1 \leq i_1 \leq i_2 \leq N^{1-\alpha}} \frac{(S_{i_1+j} - S_j)^2 + (S_{i_2+j} - S_{i_1+j})^2}{i_2} < 2(1 - \varepsilon_3) \ln \ln N \right] \leq \exp(-cn^{1-\gamma}). \quad \square$$

Of course, we need to address *all* values of N and not just those that are powers of s . To allow us to handle general values of N , we introduce the following the family of intervals, which we denote by \mathcal{L} . We let C denote a sufficiently large constant which is a power of s (just how large it should be will be determined later). We define \mathcal{L} as a union of C different interval families, denoted by $\mathcal{L}_0, \dots, \mathcal{L}_{C-1}$. \mathcal{L}_0 consists of all intervals of the form: $(1 + s + \dots + s^{k-1}, 1 + s + \dots + s^k]$, where k is a non-negative integer. We note that these intervals are disjoint, and that for each k , there is exactly one interval of size s^k . More generally, \mathcal{L}_i consists of all intervals of the form:

$$\left(1 + s + \dots + s^{k-1} + \frac{is^{k+1}}{C}, 1 + s + \dots + s^k + \frac{is^{k+1}}{C} \right]$$

(for k 's large enough so that C divides s^{k+1} when $i \neq 0$). Each \mathcal{L}_i is a union of disjoint intervals, one of size s^k for each (large enough) k . To visualize these intervals, first consider \mathcal{L}_0 as an infinite stretch of intervals, starting with one of size 1, then one of size s , then one of size s^2 , and so on, each beginning just where the previous one left off. Now imagine dividing each of the intervals in \mathcal{L}_0 into C pieces of equal size. Then \mathcal{L}_1 can be thought of as a copy of \mathcal{L}_0 where the intervals have been shifted so that they now end at what used to be the end of the first piece of the next interval. In \mathcal{L}_2 , they are shifted so that they end at what used to be the end of the second piece, and so on. Note that these shifts will cause (relatively small) gaps between the intervals in \mathcal{L}_i for $i > 0$.

We let I denote an interval in \mathcal{L} . For each $j \in I$, we let A_j denote the event whose probability is bounded in (26) and \bar{A}_j denote its complement. We let P_I denote the set of points $j \in I$ for which A_j occurs. We let E_I denote the event that $|P_I| > \varepsilon_3|I|$. We now show:

Lemma 23. *Almost surely, only finitely many of the events E_I occur (for $I \in \mathcal{L}$).*

Proof. We consider an $I \in \mathcal{L}$ of size s^k . Then, for each $j \in I$,

$$\mathbb{P}[j \in P_I] \leq \exp(-ck^{1-\gamma}),$$

by Lemma 22. Hence, $\mathbb{E}[|P_I|] \leq s^k \exp(-ck^{1-\gamma})$. By Markov's inequality,

$$\mathbb{P}[E_I] \leq \frac{\mathbb{E}[|P_I|]}{\varepsilon_3 s^k} \leq \frac{1}{\varepsilon_3} \exp(-ck^{1-\gamma}).$$

We note that for each k , there are at most C intervals \mathcal{L} of size s^k . Therefore:

$$\sum_{I \in \mathcal{L}} \mathbb{P}[E_I] \leq \frac{C}{\varepsilon_3} \sum_{k=0}^{\infty} \exp(-ck^{1-\gamma}) < \infty.$$

Hence, by the Borel–Cantelli lemma, with probability one only finitely many of the events E_I occur. □

Finally, we will prove Theorem 20 by considering a sufficiently large N and defining the following partition of $[N]$. We begin by picking out a useful set of disjoint intervals in \mathcal{L} . We will denote this set by S , and we consider it initially to be empty. There is a unique positive integer n_0 such that $1 + s + \dots + s^{n_0} \leq N < 1 + s + \dots + s^{n_0+1}$. There also exists an i_0 such that:

$$0 \leq N - \left(1 + s + \dots + s^{n_0} + \frac{i_0 s^{n_0+1}}{C} \right) \leq \frac{s^{n_0+1}}{C}.$$

We add to S the interval of size s^{n_0} in \mathcal{L} which ends at $1 + s + \dots + s^{n_0} + \frac{i_0 s^{n_0+1}}{C}$, and we call this interval I_1 . Now, there exists some n_1 and i_1 such that the left endpoint of I_1 falls in the range:

$$\left[1 + s + \dots + s^{n_1} + \frac{i_1 s^{n_1+1}}{C}, 1 + s + \dots + s^{n_1} + \frac{(i_1 + 1) s^{n_1+1}}{C} \right).$$

We then define I_2 to be the interval of size s^{n_1} in \mathcal{L} which ends at $1 + s + \dots + s^{n_1} + \frac{i_1 s^{n_1+1}}{C}$. We then consider the left endpoint of I_2 , and we define I_3 analogously, continuing on until we reach a point where the next interval we would like to use does not exist.

When we are finished, S is a finite set of disjoint intervals covering most of the length from 1 to N . To see this, note that each time we add an interval of size s^{n_ℓ} , we insert a gap of size at most $s^{n_\ell+1}/C$. For C chosen sufficiently large with respect to s , the gap will be only a small proportion of the length of the interval being added. Thus, we lose only a small fraction of the length of N to these gaps. Additionally, we can afford to ignore the length from 1 up to $1 + s + \dots + s^{n_0/2}$, since this is $o(s^{n_0})$ and hence $o(N)$. Thus, at least a length of

$$N \left(1 - \frac{s}{C} \right) - (1 + s + \dots + s^{n_0/2}) \tag{28}$$

is contained in intervals in \mathcal{L} of size at least $s^{n_0/2}$.

We now define our partition of N , choosing our endpoints iteratively. We start at 0. (Our current position will always be the last element of the interval we just added to the partition, i.e. the last endpoint we choose.) When we are in a gap between intervals in S , we choose the next endpoint in our partition to be the end of the gap and we move to the first point of the next interval in S . When we are inside an interval in S , we let j denote our current position, and $size_j$ denote the size of the interval in S . If the event A_j occurs, we add an interval of size 1 to our partition and move to $j + 1$. If \bar{A}_j occurs and $j + size_j^{1-\alpha}$ is still within the current interval of S , we choose the i_1 and i_2 which maximize $(S_{i_1+j} - S_j)^2 + (S_{i_2+j} - S_{i_1+j})^2$ for $1 \leq i_1 \leq i_2 \leq size_j^{1-\alpha}$ as the next two endpoints and move to the point $i_2 + j$. If $j + size_j^{1-\alpha}$ lies outside the current interval of S , we choose the next endpoint to be the beginning of the next interval in S . We continue in this way until we reach N .

Now we must prove a lower bound (almost surely) for the sum of the squared partial sums over the intervals in this partition. We will ignore the gaps (which contribute an amount ≥ 0), and consider the contributions of the partition pieces which lie inside intervals of S . Almost surely, only finitely many of the intervals I in S have $|P_I| > \varepsilon_3 |I|$. We assume that N is large enough so that all of the intervals $I \in S$ of size at least $s^{n_0/2}$ have $|P_I| \leq \varepsilon_3 |I|$. Thus, the pieces of the partition in each of these intervals I contributes at least

$$2(1 - \varepsilon_3)^2 (|I| - |I|^{1-\alpha}) \ln \ln |I|$$

to the sum of the squared partial sums over the partition. For $|I| \geq s^{n_0/2}$, $\ln \ln |I| \sim \ln \ln N$. Since these intervals cover a length that is at least (28), we can choose ε_3 small enough with respect to ε and C large enough with respect to s to obtain a contribution that is $\geq (2 - \varepsilon)N \ln \ln N$. This completes our proof of Theorem 20.

4. A result for convergence in probability

We now prove the exact asymptotic holds for convergence in probability when the variables X_i only satisfy $\text{Var}[X_i] = 1$ and not any higher moment condition.

Theorem 24. *We let X_1, X_2, \dots denote independent, identically distributed random variables with mean zero and variance equal to one. Then for every $\varepsilon, \delta > 0$,*

$$\mathbb{P}[\|\{X_i\}^N\|_{V^2}^2 > 2(1 + \varepsilon)N \ln \ln N] < \delta$$

for all sufficiently large N , and similarly,

$$\mathbb{P}[\|\{X_i\}^N\|_{V^2}^2 < 2(1 - \varepsilon)N \ln \ln N] < \delta$$

for all sufficiently large N .

Proof. We will use the results of [14] along with truncation, writing $X_i = \bar{X}_i + Z_i$, where $\bar{X}_i := \mathbb{I}_{|X_i| \leq M}$ and M is chosen sufficiently large so that $\text{Var}[\bar{X}_i]$ is close to 1, and $\mathbb{E}[Z_i^2] = \varepsilon' < \frac{1}{c'}\varepsilon_1$, where c' is the value from Theorem 4 and $\varepsilon_1, \varepsilon_2 > 0$ are chosen so that $\sqrt{\varepsilon_1} + \sqrt{1 + \varepsilon_2} \leq \sqrt{1 + \varepsilon}$. Then, applying Theorem 4 to the mean zero variables $Z_i - \mathbb{E}[Z_i]$, we have (for all sufficiently large N):

$$\mathbb{P}[\|\{Z_i - \mathbb{E}[Z_i]\}^N\|_{V^2}^2 > \varepsilon_1(2N \ln \ln N)] < \frac{\delta}{2}.$$

By Egorov's theorem, convergence almost surely implies convergence in probability, so we can apply Theorem 1 to the bounded variables $\bar{X}_i - \mathbb{E}[\bar{X}_i]$ to obtain (for all sufficiently large N):

$$\mathbb{P}[\|\{\bar{X}_i - \mathbb{E}[\bar{X}_i]\}^N\|_{V^2}^2 > (1 + \varepsilon_2)2N \ln \ln N] < \frac{\delta}{2}.$$

Since $\sqrt{\varepsilon_1} + \sqrt{1 + \varepsilon_2} \leq \sqrt{1 + \varepsilon}$, the triangle inequality implies that:

$$\mathbb{P}[\|\{X_i\}^N\|_{V^2}^2 > 2(1 + \varepsilon)N \ln \ln N] < \delta.$$

The second bound can be proven by an analogous argument applying Theorem 1 to \bar{X}_i , Theorem 4 to Z_i , and employing $\|\{X_i\}^N\|_{V^2}^2 \geq \|\{\bar{X}_i - \mathbb{E}[\bar{X}_i]\}^N\|_{V^2}^2 - \|\{Z_i - \mathbb{E}[Z_i]\}^N\|_{V^2}^2$. \square

Appendix: Proof of Lemma 17

We first note Doob's inequality:

Lemma 25 (Doob's inequality [3]). *Let $\{M_i\}_{i=1}^L$ be a submartingale taking non-negative real values. Then*

$$\mathbb{P}\left[\sup_{0 \leq \ell \leq L} M_\ell \geq t\right] \leq \frac{\mathbb{E}[M_L]}{t}. \tag{29}$$

We now prove Lemma 17. Let $g(y) = 2 \sum_{l=2}^\infty \frac{y^{l-2}}{l!} = \frac{2(e^y - 1 - y)}{y^2}$. Now

$$\begin{aligned} \mathbb{E}[e^{h \sum_{i=1}^L X_i}] &= \prod_{i=1}^L \mathbb{E}[e^{hX_i}] = \prod_{i=1}^L \mathbb{E}\left[\sum_{k=0}^\infty \frac{h^k X_i^k}{k!}\right] \\ &= \prod_{i=1}^L \mathbb{E}\left[1 + h(X_i) + \frac{1}{2}h^2 X_i^2 g(hX_i)\right] \leq \prod_{i=1}^L \left(1 + h\mathbb{E}[X_i] + \frac{1}{2}h^2\mathbb{E}[X_i^2]g(hM)\right) \\ &\leq e^{h \sum_{i=1}^L \mathbb{E}[X_i] + (1/2)h^2 \sum_{i=1}^L \mathbb{E}[X_i^2]g(hM)} = e^{(1/2)h^2 g(hM) \sum_{i=1}^L \mathbb{E}[X_i^2]}. \end{aligned}$$

We have:

$$\mathbb{P}\left[\sum_{i=1}^L X_i \geq t\right] = \mathbb{P}\left[e^{h \sum_{i=1}^L X_i} \geq e^{ht}\right]$$

and, more generally:

$$\mathbb{P}\left[\max_{1 \leq \ell \leq L} \sum_{i=1}^{\ell} X_i \geq t\right] = \mathbb{P}\left[\max_{1 \leq \ell \leq L} e^{h \sum_{i=1}^{\ell} X_i} \geq e^{ht}\right].$$

Since $M_\ell := e^{h \sum_{i=1}^{\ell} X_i}$ forms a submartingale sequence, Doob's inequality yields:

$$\mathbb{P}\left[\max_{1 \leq \ell \leq L} \sum_{i=1}^{\ell} X_i \geq t\right] \leq e^{-ht} \mathbb{E}\left[e^{h \sum_{i=1}^L X_i}\right] \leq e^{-ht} e^{(1/2)h^2 g(hM) \sum_{i=1}^L \mathbb{E}[X_i^2]}.$$

Using the fact that $g(y) \leq \frac{1}{1-y/3}$ for $y < 3$ (which follows from the Taylor expansion given above), we have

$$\mathbb{P}\left[\max_{1 \leq \ell \leq L} \sum_{i=1}^{\ell} X_i \geq t\right] \leq e^{(1/2)h^2 \sum_{i=1}^L \mathbb{E}[X_i^2]/(1-hM/3)-ht}$$

when $hM < 3$.

Taking $h = \frac{t}{\sum_{i=1}^L \mathbb{E}[X_i^2] + Mt/3}$ (which satisfies $hM < 3$), we have:

$$\begin{aligned} & \exp\left(\frac{(1/2)h^2 \sum_{i=1}^L \mathbb{E}[X_i^2]}{1-hM/3} - ht\right) \\ &= \exp\left(\frac{(1/2)t^2 \sum_{i=1}^L \mathbb{E}[X_i^2]}{(\sum_{i=1}^L \mathbb{E}[X_i^2] + Mt/3)^2 (1-tM/(3(\sum_{i=1}^L \mathbb{E}[X_i^2] + Mt/3)))} - \frac{t^2}{\sum_{i=1}^L \mathbb{E}[X_i^2] + Mt/3}\right) \\ &= \exp\left(\frac{(1/2)t^2 \sum_{i=1}^L \mathbb{E}[X_i^2]}{(\sum_{i=1}^L \mathbb{E}[X_i^2] + Mt/3) \sum_{i=1}^L \mathbb{E}[X_i^2]} - \frac{t^2}{\sum_{i=1}^L \mathbb{E}[X_i^2] + Mt/3}\right) \\ &= \exp\left(\frac{-(1/2)t^2}{\sum_{i=1}^L \mathbb{E}[X_i^2] + Mt/3}\right). \end{aligned}$$

This establishes that

$$\mathbb{P}\left[\max_{1 \leq \ell \leq L} \sum_{i=1}^{\ell} X_i > t\right] \leq e^{-t^2/2/(\sum_{i=1}^L \mathbb{E}[X_i^2] + Mt/3)}.$$

Applying this result to the random variables $\{Y_i\}$ where $Y_i = -X_i$, we have

$$\mathbb{P}\left[\max_{1 \leq \ell \leq L} \sum_{i=1}^{\ell} X_i < -t\right] \leq e^{-t^2/2/(\sum_{i=1}^L \mathbb{E}[X_i^2] + Mt/3)}.$$

Combining these, we obtain the desired estimate

$$\mathbb{P}\left[\max_{1 \leq \ell \leq L} \left|\sum_{i=1}^{\ell} X_i\right| > t\right] \leq 2e^{-t^2/2/(\sum_{i=1}^L \mathbb{E}[X_i^2] + Mt/3)}.$$

Acknowledgement

We would like to thank the referee for a careful reading of the manuscript. The referee has pointed out that there is a number of ways to replace Corollary 18 (and thus avoid the need for Lemma 17) with existing results from the literature. For instance, Theorem B.2 of [15] would be a sufficient substitute. Since Lemma 17 is perhaps of independent interest we decided not to implement this simplification.

References

- [1] J. Bretagnolle. p -variation de fonctions aléatoires. II. Processus à accroissements indépendants. In *Séminaire de Probabilités VI (Univ. Strasbourg, Année Universitaire 1970–1971). Lecture Notes in Math.* **258** 64–71. Springer, Berlin, 1972. [MR0373020](#)
- [2] Y. S. Chow and H. Teicher. *Probability Theory: Independence, Interchangeability, Martingales*, 3rd edition. Springer, Berlin, 1997. [MR1476912](#)
- [3] J. L. Doob. *Stochastic Processes*. Wiley, New York, 1953. [MR0058896](#)
- [4] R. M. Dudley. Empirical processes and p -variation. In *Festschrift for Lucien Le Cam* 219–233. Springer, Berlin, 1997. [MR1462947](#)
- [5] N. Etemadi. On some classical results in probability theory. *Sankhyā Ser. A* **47** (1985) 215–221. [MR0844022](#)
- [6] J. Galambos. *Advanced Probability Theory*, 2nd edition. Dekker, New York, 1995. [MR1350792](#)
- [7] P. Hartman and A. Wintner. On the law of the iterated logarithm. *Amer. J. Math.* **63** (1941) 169–176. [MR0003497](#)
- [8] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** (301) (1963) 13–30. [MR0144363](#)
- [9] R. Jones and G. Wang. Variation inequalities for the Fejér and Poisson kernels. *Trans. Amer. Math. Soc.* **356** (11) (2004) 4493–4518. [MR2067131](#)
- [10] A. Lewko and M. Lewko. Estimates for the square variation of partial sums of Fourier series and their rearrangements. *J. Funct. Anal.* **262** (6) (2012) 2561–2607. [MR2885959](#)
- [11] A. Lewko and M. Lewko. Orthonormal systems in linear spans. *Anal. PDE* **7** (2014) 97–115. [MR3219501](#)
- [12] A. Lewko and M. Lewko. A variational Barban–Davenport–Halberstam theorem. *J. Number Theory* **132** (9) (2012) 2020–2045. [MR2925860](#)
- [13] A. Lewko and M. Lewko. The square variation of rearranged Fourier series. *Amer. J. Math.* To appear, 2015. Available at [arXiv:1212.1988](#).
- [14] J. Qian. The p -variation of partial sum processes and the empirical process. *Ann. Probab.* **26** (3) (1998) 1370–1383. [MR1640349](#)
- [15] E. Rio. *Théorie asymptotique des processus aléatoires faiblement dépendants. Mathématiques and Applications* **31**. Springer, Berlin, 1999. [MR2117923](#)
- [16] H. P. Rosenthal. On the subspaces of L^p ($p > 2$) spanned by sequences of independent random variables. *Israel J. Math.* **8** (1970) 273–303. [MR0271721](#)
- [17] S. J. Taylor. Exact asymptotic estimates of Brownian path variation. *Duke Math. J.* **39** (1972) 219–241. [MR0295434](#)