The Annals of Probability
2002, Vol. 30, No. 4, 1854–1892

MAXIMA OF PARTIAL SUMS INDEXED BY GEOMETRICAL STRUCTURES

BY TIEFENG JIANG

University of Minnesota

The maxima of partial sums indexed by squares and rectangles over lattice points and random cubes are studied in this paper. For some of these problems, the dimension ($d = 1, d = 2$ and $d \geq 3$) significantly affects the limit behavior of the maxima. However, for other problems, the maxima behave almost the same as their one-dimensional counterparts. The tools for proving these results are large deviations, the Chen–Stein method, number theory and inequalities of empirical processes.

1. Introduction. Motivated by comparisons of protein structures with three dimensional foldings (see [14] and [18] for further details), we study maxima of partial sums of i.i.d. random variables with indices on d-dimensional lattices points ($d \geq 2$) as well as positive random numbers generated by the uniform distribution over the three-dimensional cube.

Before stating our main results, let us recall a result for the one-dimensional case.

In our context, the random variable X is typically assumed to satisfy the following condition:

$$\begin{aligned} &X \text{ is nonlattice, } E(X) < 0,\ P(X > 0) > 0 \text{ and} \\ &E\exp(tX) < \infty \text{ for all } t \in \mathbb{R}. \end{aligned} \tag{1.1}$$

Under condition (1.1), there is a unique constant $\theta > 0$ so that

$$E\exp(\theta X) = 1. \tag{1.2}$$

The following lemma was probably first proved by Spitzer (E4 on page 217 of [22]). See also (5.13) in [11].

LEMMA A.1. *Let $\{X, X_n; n \geq 1\}$ be a sequence of i.i.d. random variables with X satisfying* (1.1). *Set $S_k = \sum_{i=1}^{k} X_i,\ k \geq 1$. Then*

$$K := \lim_{t \to +\infty} e^{\theta t} P\left(\max_{k \geq 1} S_k > t\right) = C/\theta,$$

Received April 2000; revised December 2001.

AMS 2000 subject classifications. 60F10, 28C15, 60B10.

Key words and phrases. Maxima, Chen–Stein method, number theory, large deviations, inequalities of empirical processes.

where

$$C = \frac{A \exp\{-2\sum_{k=1}^{\infty} \frac{1}{k}(E[\exp(\theta S_k); S_k < 0] + P(S_k \geq 0))\}}{E[X \exp(\theta X)]}$$

and $A = \exp\{\sum_{k=1}^{\infty} P(S_k \geq 0)/k\}$. *The above expression for C follows from i.i.d. fluctuation sum identities* ([11, *Chapter* 12]).

Lemma A.1 is important for i.i.d. partial sums. Iglehart [13] used this result in the continuous i.i.d. case in the course of characterizing the asymptotic maximal waiting time among the first n customers in a standard $GI/G/1$ queue. One such result is as follows:

THEOREM A.1. *Let* $\{X, X_n; n \geq 1\}$, S_k *and* K *be as in Lemma* A.1. *Set* $S_0 = 0$. *Then*

$$P\left(\max_{0 \leq i < j \leq n}(S_j - S_i) \leq \frac{\log n}{\theta} + x\right) \to e^{-Ke^{-\theta x}} \qquad \textit{as } n \to \infty$$

for any $x \in \mathbb{R}$.

For more information on oscillation phenomena for partial sums of i.i.d. random variables see [5, 17, 19] and the references therein.

In this paper, we study counterparts of Theorem A.1 for two or higher dimensional cases. Due to the complexity of higher dimensional spaces, a discrete version and a continuous version are studied separately.

Now we state our results for the first part.

Denote the set of all positive integers by $\mathbb{N}$ and d-fold Cartesian product of $\mathbb{N}$ by $\mathbb{N}^d$, namely, $\mathbb{N}^d = \{I = (i_1, i_2, \ldots, i_d);\ i_k \in \mathbb{N},\ k = 1, 2, \ldots, d\}$. For any $n \geq 1$, define the set of all subcubes and that of rectangles in $\{1, 2, \ldots, n\}^d$ by $\mathcal{O}_n$ and $\mathcal{R}_n$, respectively. Precisely, for any $\Delta \in \mathcal{O}_n$ and $\Delta' \in \mathcal{R}_n$, there exist $\{j_k\}_{k=1}^d$, $\{l_k\}_{k=1}^d \in \mathbb{N}^d$ and $m \in \mathbb{N}$ such that

$$\Delta = \{(i_1, \ldots, i_d) \in \mathbb{N}^d;\ 1 \leq j_k \leq i_k \leq j_k + m \leq n,\ k = 1, 2, \ldots, d\}$$

and

$$\Delta' = \{(i_1, \ldots, i_d) \in \mathbb{N}^d;\ 1 \leq j_k \leq i_k \leq l_k \leq n,\ k = 1, 2, \ldots, d\}.$$

Assuming that $\{X, X_I;\ I \in \mathbb{N}^d\}$ are i.i.d. random variables, let $S_\Delta = \sum_{I \in \Delta} X_I$,

$$W_n = \max_{\Delta \in \mathcal{O}_n} S_\Delta \quad \text{and} \quad U_n = \max_{\Delta \in \mathcal{R}_n} S_\Delta.$$

We focus mainly on these two statistics in the first part of the paper. Strong laws and limiting distributions of them are derived.

For the random variable X mentioned in (1.1), the corresponding log of the moment generating function and its conjugate which is also called the rate function are

$$\Lambda_X(t) = \log E\exp(tX), \qquad \Lambda_X^*(x) := \sup_{t\in\mathbb{R}}\{tx - \Lambda_X(t)\}.$$

When there is no confusion, we may for the sake of convenience, write $\Lambda(t)$ for $\Lambda_X(t)$ and $\Lambda^*(x)$ for $\Lambda_X^*(x)$, respectively. To understand the limiting distributions of W_n and U_n, we need the following local properties of partial sums corresponding to that of Lemma A.1 in one-dimensional case. Recall $S_k = \sum_{i=1}^k X_i,\ k \ge 1$, are the partial sums of a sequence of i.i.d. random variables $\{X, X_i; i \ge 1\}$.

THEOREM 1. *Suppose condition* (1.1) *holds. For* $z > 0$, *let* $\gamma(z) = (z/\Lambda'(\theta))^{1/2}$ *and* $\delta(z) = \sum_{i=-\infty}^{+\infty}\exp\{-\beta(i+z)^2\}$, *where* $\beta = 2\Lambda'(\theta)^2/\Lambda''(\theta)$. *Then*

$$\lim_{z\to\infty}\frac{\sqrt{z}e^{\theta z}}{\delta(\gamma(z))}P\left(\max_{k\ge 1} S_{k^2} \ge z\right) = \frac{1}{\theta}\sqrt{\frac{\Lambda'(\theta)}{2\pi\Lambda''(\theta)}}.$$

Although we assume X is nonlattice in all of our results here and later, the lattice cases can be treated similarly. We omit them in this paper.

When $d \ge 3$, we have the following result.

THEOREM 2. *Suppose condition* (1.1) *holds. Let* $G_n(z) = \exp(-n^d\Lambda^*(z/n^d))$, *where* $n := [(z/\Lambda'(\theta))^{1/d}]$ (*recall* $[x]$ *is the biggest integer no larger than* x). *Then for any integer* $d \ge 3$,

$$\lim_{z\to\infty}\sqrt{z}\big(G_n(z) + G_{n+1}(z)\big)^{-1}P\left(\max_{k\ge 1} S_{k^d} \ge z\right) = \frac{1}{\theta}\sqrt{\frac{\Lambda'(\theta)}{2\pi\Lambda''(\theta)}}.$$

Why are the results for $d = 1,\ d = 2$ and $d \ge 3$ so different? For ease of discussion, assume that X is bounded. Compared to a given large number $z,\ S_k$ is very small both when k is small (because X is bounded) and when k is large (because of the negative mean). Let Γ_d be the set of integers k which essentially contribute to $\max_{k\ge 1} S_{k^d}$.

By a computation given later in this paper, we know that Γ_1 is the set of integers in $(z - \sqrt{z\log z}, z + \sqrt{z\log z})$. The size of Γ_d is therefore roughly equal to

$$(z + \sqrt{z\log z})^{1/d} - (z - \sqrt{z\log z})^{1/d} \simeq \frac{\sqrt{\log z}}{z^{1/2-1/d}}.$$

Obviously, $d = 2$ is the critical value in which the size of Γ_d is roughly $\sqrt{\log z}$ in contrast to $\sqrt{z\log z}$ when $d = 1$. When $d \ge 3,\ \Gamma_d$ consists of at most some fixed set of integers. The real calculation shows that the size of Γ_d in this case is at most two.

Now we turn to another case of local properties of partial sums. Let $\{X, X_{i,j}; i \geq 1, j \geq 1\}$ be i.i.d. random variables and $S_{p,q} = \sum_{i=1}^{p} \sum_{j=1}^{q} X_{i,j}$.

THEOREM 3. *Let $U = \max_{p\geq 1, q\geq 1} S_{p,q}$. Suppose condition* (1.1) *holds*, *then*

$$\lim_{z\to+\infty} e^{\theta z} (\log z)^{-1} P(U \geq z) = \frac{1}{\theta\sqrt{\Lambda'(\theta)}}.$$

For the one-dimensional case, the proof of Lemma A.1 depends on classical fluctuation theory. The proofs of the above theorems are totally irrelevant to that.

The following are strong laws for W_n and U_n.

THEOREM 4. *Suppose condition* (1.1) *holds*, *then for any $d \geq 2$*,

(i) $\lim_{n\to\infty} \frac{W_n}{\log n} \to \frac{d}{\theta}$ *a.s.*,

(ii) $\lim_{n\to\infty} \frac{U_n}{\log n} \to \frac{d}{\theta}$ *a.s.*

The following are limit laws of W_n and U_n. As usual, $\log_2 n = \log(\log n)$, $\log_3 n = \log(\log_2 n)$.

THEOREM 5. *Suppose that $d = 2$ and condition* (1.1) *holds. Let $t_n = \log \delta(\gamma(2\log n/\theta))$, where the functions $\delta(\cdot)$ and $\gamma(\cdot)$ are as in Theorem* 1. *Define $\log_2 n = \log(\log n)$. Then*

$$\lim_{n\to\infty} P\left(W_n \leq \frac{1}{\theta}\left\{2\log n - \frac{1}{2}\log_2 n + t_n\right\} + x\right) = e^{-K_1 e^{-\theta x}}$$

for all $x \in \mathbb{R}$, where $K_1 = 2^{-1}\sqrt{\Lambda'(\theta)/(\pi\theta\Lambda''(\theta))}$.

THEOREM 6. *Suppose $d \geq 3$ and condition* (1.1) *holds. Let $k_n = \inf\{k \in \mathbb{N}; (\log k)/2 + \alpha k^d \geq \log n\}$, where $\alpha = \theta\Lambda'(\theta)/d$, and $r_n = \exp\{d\log n - d(\log k_n)/2 - k_n^d \theta\Lambda'(\theta)\}$. Then*

$$P\big(W_n \leq \Lambda'(\theta)k_n^d + x\big) - e^{-K_2 r_n e^{-\theta x}} \to 0,$$

where $K_2 = (\theta\sqrt{2\pi\Lambda''(\theta)})^{-1}$.

It is easy to see that r_n of Theorem 6 does not converge. Also $P(W_n \leq \Lambda'(\theta)k_n^d + x)$ does not converge, but Theorem 6 gives a first order of approximation for the probability.

THEOREM 7. *Suppose condition* (1.1) *holds*, *for* $d = 2$, *we have that*

$$P\left(U_n \le \frac{2\log n}{\theta} + \frac{\log_3 n}{\theta} + x\right) \to e^{-K_3 e^{-\theta x}} \qquad \forall x \in \mathbb{R},$$

where $K_3 = 1/\theta\sqrt{\Lambda'(\theta)}$.

From Theorem 4, we know that both U_n and W_n have the same scale. But evidently, $U_n \ge W_n$. Theorem 5 tells us, loosely speaking, that $W_n \sim (2\log n - (1/2)\log_2 n)/\theta$ when $d = 2$. The above theorem says roughly that $U_n \sim (2\log n + \log_3 n)/\theta$ when $d = 2$. The difference between them is obvious.

The analogue of Theorem 7 for the high-dimensional case is not derived in this paper because a related number theoretic problem is unsolved. In fact, one of the key steps in proving Theorem 7 is to show that

$$\sum_{k \in I_y} q(k) e^{-(k-y)^2/y} \sim \alpha\sqrt{y}\log y \qquad \text{as } y \to +\infty \tag{1.3}$$

for some constant $\alpha > 0$, where $q(k) = \#\{(r,s) \in \mathbb{N}^2; rs = k\}$ and I_y is an interval depends on y. To solve the analogue of Theorem 7 for the high-dimensional case, a calculation similar to (1.3) must be done. See Remark 5.6 for further details.

The above results can also be thought as natural extensions of the classical Erdös–Rényi law (see [10]) and its followups such as [5] to a higher dimensional setting.

In the second part of this paper, results in the "continuous" setting are obtained. They are actually motivated by a procedure given by Karlin and Zhu [15], which studied clusters of charged residues in protein structures. To focus on mathematics, we omit any details of biology throughout this paper.

Assume that $\{Y, Y_i; i \ge 1\}$ is a sequence of i.i.d. random variables with uniform distribution on $[0,1]^3$. For any $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, $\|x\| = \max\{|x_1|, |x_2|, |x_3|\}$ is the maximum norm. A ball centered at x and with radius r under this norm is denoted as $B(x, r)$. We denote by $\mathcal{F}$ the set of all subcubes inside of $[0,1]^3$ such that their six faces are parallel to those of $[0,1]^3$. Specifically,

$$\mathcal{F} = \{B(x,r) \subset [0,1]^3;\ x \in [0,1]^3,\ 0 < r < 1/2\}. \tag{1.4}$$

Let $\{X, X_n; n \ge 1\}$ be a sequence of i.i.d. random variables. For any $B \subset [0,1]^3$, define $S_n(B) = \sum_{i=1}^n X_i I\{Y_i \in B\}$. We consider the following two statistics:

$$\tilde{W}_n = \max_{B \in \mathcal{B}} S_n(B) \quad \text{and} \quad \tilde{U}_n = \max_{B \in \mathcal{F}} S_n(B), \tag{1.5}$$

where $\mathcal{B} := \{B = B(Y_i, r) \subset [0,1]^3;\ 1 \le i \le n,\ 0 < r < 1/2\}$.

THEOREM 8. *Suppose condition* (1.1) *holds*, *then*

$$\lim_{n\to\infty} \frac{\tilde{W}_n}{\log n} = \frac{1}{\theta} \quad a.s. \quad \textit{and} \quad \lim_{n\to\infty} \frac{\tilde{U}_n}{\log n} = \frac{1}{\theta} \quad a.s.$$

THEOREM 9. *Suppose condition* (1.1) *holds*, *then*

$$\lim_{n\to\infty} P\left(\tilde{W}_n \le \frac{\log n}{\theta} + x\right) = e^{-Ke^{-\theta x}}$$

for any $x \in \mathbb{R}$, *where* K *is a constant as in Lemma* A.1.

The method of proof of this result is a combination of the classical fluctuation theory in the one-dimensional case and the "diffuse" property of high dimensional spaces. We may think of Theorem 9 as one of the results of scan statistics see [20] and [21] and the literature therein.

One application of our results is the following change point problem. Suppose we have independent observation on two dimensional lattice points:

−1.4	−3.3	−1.8	−2.8	−0.2	−2.3	−3.0
−2.4	−3.1	−1.2	−2.5	−2.3	−2.7	−1.6
−0.6	−1.1	−0.3	−4.1	−0.9	−1.5	−0.5
−2.8	−1.9	−3.0	−0.7	−2.8	−1.2	−1.5
−1.2	−1.4	−2.6	1.2	1.4	1.3	−0.7
−1.8	−1.9	−2.5	1.6	1.3	1.4	−4.2
−1.5	−1.6	−1.1	−1.5	−0.1	−2.9	−1.2

FIG. 1.1.

One immediately notices that there is some zone where the data are significantly different from those in the other parts. [The above data are actually sampled from the distribution $N(-2, 1)$, and the data in the area enclosed by the fifth and sixth rows and the fourth and sixth columns are later changed manually to the current ones.] This is a typical setting in change point problems. The goal is to detect whether there is a zone from which the data are different from the data in other zones. Siegmund and Yakir in [21] studied this problem recently by using the likelihood ratio test. Our Theorems 5, 6 and 7 provide another way to study such a problem in which data are assumed from a population with negative mean and essential positive part. So far we do not know which method is more efficient.

Finally, let us give the outline of this paper. We will prove results on maxima on squares, rectangles and random cubes in Sections 2, 3 and 4, respectively. We give some concluding remarks in the last section.

2. Proofs of Theorems 1, 2, 4, 5 and 6.

2.1. *Notation and some auxiliary lemmas.* Throughout this paper, we use the following notation:

$\mathbb{N}$: The set of all positive integers.

$\mathbb{R} := (-\infty, +\infty)$.

$[a]$: The integer part of a.

$|A|$ or $\#A$: the cardinality of a set A.

I_A or $I(A)$ or 1_A or $1(A)$ are the same function of $x := 1$ if $x \in A$, $= 0$ otherwise.

$a_n \sim b_n : a_n/b_n \to 1$ as $n \to \infty$.

$a_n = O(b_n)$: $\limsup_{n\to\infty} |a_n/b_n| < \infty$.

$a_n = o(b_n)$: $\lim_{n\to\infty} a_n/b_n = 0$.

$a \vee b := \max\{a, b\}$ and $a \wedge b := \min\{a, b\}$.

$E^A f(\xi_1, \xi_2, \dots, \xi_n)$: suppose $\xi_1, \xi_2, \dots, \xi_n$ are random variables. For $f(x_1, x_2, \dots, x_n)$, a real-valued function on $\mathbb{R}^n$, denote by $E^A f(\xi_1, \xi_2, \dots, \xi_n)$ the conditional expectation $E(f(\xi_1, \xi_2, \dots, \xi_n)|\mathcal{B})$, where $\mathcal{B}$ is the σ-algebra generated by $\{\xi_k, k \notin A\}$ if $A \subset \{1, 2, \dots, n\}$ or by $\{\xi_1, \dots, \xi_n\} \setminus A$ if A is a subset of $\{\xi_k, 1 \le k \le n\}$. The same interpretation applies to P^A too.

Before proving the main theorems, we collect some tools which will be frequently used in this and later sections. Some of those tools are quoted directly from the literature. They will be denoted by Lemmas A.2 and A.3, etc. as in the introduction. We use the numbering such as Lemmas 2.1 and 3.2 to denote those results which need proofs.

The following inequality provides us with bounds for tails of sums of independent and bounded random variables; see Exercise 14 on page 111 in [4] or page 193 in [16].

LEMMA A.2 (Bernstein's inequality). *Let $\{X_i;\ 1 \le i \le n\}$ be a sequence of independent random variables with $EX_i = 0,\ EX_i^2 = \sigma_i^2$ and $|X_i| \le 1$. Denote $s_n^2 = \sum_{i=1}^n \sigma_i^2$. Then*

$$P(S_n > x) \le \exp\left\{-\frac{x^2}{2(s_n^2 + x)}\right\}, \qquad x > 0.$$

LEMMA 2.1. *Let $\{X_\alpha; \alpha \in \Delta\}$ be a collection of i.i.d. random variables, where Δ is a finite set. Suppose that A, A_1, B and B_1 are subsets of Δ satisfying $|A| = |A_1|,\ |B| = |B_1|$ and $|A \cap B| \le\ |A_1 \cap B_1|$. Let $\Phi(x)$ be a measurable function on $\mathbb{R}$. If $|A| = |B|$ or Φ is monotone, then*

$$E\Phi(S_A)\Phi(S_B) \le E\Phi(S_{A_1})\Phi(S_{B_1}),$$

where $S_C = \sum_{\alpha\in C} X_\alpha$ for any set C.

PROOF. We distinguish two cases.

(i) Suppose $|A| = |B|$. As mentioned earlier, $E^A(Y) := E(Y|X_\alpha, \alpha \in \Delta \setminus A)$ for any random variable Y. Take a subset $D \subset A_1 \cap B_1$ for which $|D| = |A_1 \cap B_1| -$

$|A \cap B|$. Then from the invariance property of the joint distribution of $\{X_\alpha; \alpha \in \Delta\}$, it follows that

$$\begin{aligned} E\Phi(S_{A_1})\Phi(S_{B_1}) &= E^{A_1\cap B_1}\big(E^{A_1\setminus B_1}\Phi(S_{A_1})\big)^2 = E^{(A_1\cap B_1)\setminus D}E^D\big(E^{A_1\setminus B_1}\Phi(S_{A_1})\big)^2 \\ &\geq E^{(A_1\cap B_1)\setminus D}\big(E^{(A_1\setminus B_1)\cup D}\Phi(S_{A_1})\big)^2 = E^{A\cap B}\big(E^{A\setminus B}\Phi(S_A)\big)^2 \\ &= E\Phi(S_A)\Phi(S_B) \end{aligned}$$

where the only "$\geq$" appearing in the above argument is by virtue of the Cauchy–Schwarz inequality.

(ii) Suppose that Φ is monotone. As above, take a subset $D \subset A_1 \cap B_1$ so that $|D| = |A_1 \cap B_1| - |A \cap B|$. It follows that

$$\begin{aligned} E\Phi(S_{A_1})\Phi(S_{B_1}) &= E^{A_1\cap B_1}\big\{E^{A_1\setminus B_1}\Phi(S_{A_1})E^{B_1\setminus A_1}\Phi(S_{B_1})\big\} \\ &= E^{(A_1\cap B_1)\setminus D}E^D\big\{E^{A_1\setminus B_1}\Phi(S_{A_1})E^{B_1\setminus A_1}\Phi(S_{B_1})\big\} \\ &\geq E^{(A_1\cap B_1)\setminus D}\big\{E^{(A_1\setminus B_1)\cup D}\Phi(S_{A_1})E^{(B_1\setminus A_1)\cup D}\Phi(S_{B_1})\big\} \\ &= E\Phi(S_A)\Phi(S_B) \end{aligned}$$

where we use the easy fact that $Ef(Y)g(Y) \geq Ef(Y)Eg(Y)$ for any two increasing functions $f,\ g$ and a random variable Y in the only inequality appearing above. $\square$

The following Poisson approximation theorem is a straightforward application of Theorem 1 in [1], which is a special case of the Chen–Stein method. The lemma is used quite often in analyzing maxima of random variables.

LEMMA 2.2. *Let Ω be a finite set and $\mathcal{A}$ be a collection of some subsets of Ω. Suppose that $\{X_\alpha,\ \alpha \in \Omega\}$ is a collection of random variables. Write $S_A = \sum_{\alpha\in A} X_\alpha$ and $\lambda = \sum_{A\in\mathcal{A}} P(S_A > t)$ for a fixed $t \in \mathbb{R}$. Then*

$$\left|P\Big(\max_{A\in\mathcal{A}} S_A \leq t\Big) - e^{-\lambda}\right| \leq (1 \wedge \lambda^{-1})(b_1 + b_2 + b_3),$$

where

$$b_1 = \sum_{A\in\mathcal{A}}\ \sum_{B: B\cap A\neq\varnothing} P(S_A > t)P(S_B > t),$$

$$b_2 = \sum_{A\in\mathcal{A}}\ \sum_{B: B\cap A\neq\varnothing} P(S_A > t,\ S_B > t),$$

$$b_3 = \sum_{A\in\mathcal{A}} E\big|P(S_A > t|\sigma\{S_B; B\cap A = \varnothing\}) - P(S_A > t)\big|,$$

where $\sigma\{S_B; B \cap A = \varnothing\}$ is the σ-algebra generated by the collection of random variables $\{S_B; B \cap A = \varnothing\}$. In particular, if $\{X_\alpha,\ \alpha \in \Omega\}$ is a set of independent random variables, then $b_3 = 0$.

PROOF. Let $Y_A = 1(S_A > t)$, and $\mathscr{B}_A$ be the set of all B such that $B \cap A \neq \varnothing$. Then the result follows from Theorem 1 in [1]. $\square$

The next lemma, whose proof may be found in Remark (a) below Theorem 3.7.4 in [8], collects useful properties of $\Lambda(t)$ and $\Lambda^*(x)$. Their definitions are given in the introduction.

LEMMA A.3. *Suppose $E\exp(tX) < \infty$ for all $t \in \mathbb{R}$ and X is nondegenerate. Then*:

(i) $\Lambda''(t) > 0$ *for all* $t \in \mathbb{R}$.

(ii) $\Lambda^*(x)$ *is infinitely differentiable in the interior of the convex hull of the support of* X.

(iii) *If condition* (1.1) *holds, then* $\Lambda^*(\Lambda'(\theta)) = \theta\Lambda'(\theta)$, $(\Lambda^*)'(\Lambda'(\theta)) = \theta$ *and* $(\Lambda^*)''(\Lambda'(\theta)) = 1/\Lambda''(\theta)$, *where* θ *is as in* (1.2).

LEMMA 2.3. *Suppose condition* (1.1) *holds. Let* $\mathscr{D}_{\Lambda^*} = \{x \in \mathbb{R}; \Lambda^*(x) < \infty\}$. *Then* $I(x) := \Lambda^*(x)/x$ *is strictly decreasing on* $(0, \Lambda'(\theta)]$ *and strictly increasing on* $[\Lambda'(\theta), +\infty) \cap \mathscr{D}_{\Lambda^*}$.

PROOF. Obviously, the condition (1.1) implies that $[0, +\infty) \subset \{\Lambda(t);\ t \in \mathbb{R}\}$. Moreover, $\Lambda(t)$ is a strictly convex function. Thus, for any $x_2 > x_1 \geq \Lambda'(\theta)$ such that $\Lambda^*(x_i) < \infty,\ i = 1, 2,$ there exist $t_2 > t_1 \geq \theta$ such that $x_i = \Lambda'(t_i),\ i = 1, 2$. It follows from $t_i \geq \theta$ that $\Lambda(t_i) \geq 0$. It is easy to see that $\Lambda^*(\Lambda'(t)) = t\Lambda'(t) - \Lambda(t)$. Consequently,

$$\frac{\Lambda^*(x_1)}{x_1} = t_1 - \frac{\Lambda(t_1)}{x_1} < t_1 - \frac{\Lambda(t_1)}{x_2} \leq \frac{\Lambda^*(x_2)}{x_2}. \tag{2.1}$$

If $0 < x_1 < x_2 < \Lambda'(\theta)$, then there exist $0 < t_1 < t_2 < \theta$ such that $x_i = \Lambda'(t_i)$ and $\Lambda(t_i) < 0$, $i = 1, 2$. By using the same argument as (2.1), we have $\Lambda^*(x_1)/x_1 > \Lambda^*(x_2)/x_2$. $\square$

Let $\{X, X_n; n \geq 1\}$ be a sequence of i.i.d. random variables with mean μ. Recall that $S_n = \sum_{i=1}^n X_i$ are partial sums. We always assume that X is nondegenerate. The following proposition, which is slightly stronger than the usual Bahadur–Rao theorem (see [2]), provides us with uniform estimates of tail probabilities. It is a pivotal tool in our proofs.

PROPOSITION 2.1. *Suppose X is nonlattice and $\Lambda(t) < \infty$ for all $t \in \mathbb{R}$. Then*

$$\sup_{a \leq \eta \leq b} \left|C_n(\eta)P\big(S_n \geq n\Lambda'(\eta)\big) - 1\right| \to 0 \qquad \textit{as } n \to \infty$$

for any two constants $b > a > 0$, where $C_n(\eta) = \eta\sqrt{2\pi n\Lambda''(\eta)}\, e^{n\Lambda^*(\Lambda'(\eta))}$.

PROOF. Obviously, $\Lambda^*(\Lambda'(\eta)) = \sup_{|t|<a+b}\{t\Lambda'(\eta) - \Lambda(t)\}$ for any $\eta \in [a,b]$. Since X is nonlattice, the random variable $Z := e^{tX}/Ee^{tX}$ is also nonlattice. Denote the characteristic function of Z by $\phi_Z(s)$. Then $\phi_Z(s) = Ee^{(t+is)X}/Ee^{tX}$ and $|\phi_z(s)| < 1$ for any $s \neq 0$. By continuity, we have that

$$\sup_{\substack{\delta_1 \le |s| \le \delta_2 \\ |t|<a+b}} \left|\frac{Ee^{(t+is)X}}{Ee^{tX}}\right| < 1$$

for any constants $\delta_j > 0$, $j = 1, 2$. It follows from Theorem 3.3 and Remark 3.6 in [3] that $P(S_n \ge n\Lambda'(\eta_n)) \sim C_n(\eta_n)^{-1}$ as $n \to \infty$ for any $\{\eta_n; n \ge 1\} \subset [a,b]$, which implies our desired result. $\square$

A consequence of the above proposition follows. It will be used as a tool for finding accurate estimates.

COROLLARY 2.1. *Suppose the condition in Proposition* 2.1 *holds. Then, for any given* $\delta > 0$,

$$\sup_{a\le\eta\le b} \sup_{|x|\le\delta\sqrt{n\log n}} \bigl|C_n(x,\eta)P\bigl(S_n \ge n\Lambda'(\eta) + x\bigr) - 1\bigr| \to 0 \qquad \textit{as } n \to \infty,$$

where $C_n(x,\eta) = C_n(\eta)\exp\{\eta x + (x^2/2\Lambda''(\eta)n)\}$, *and* $C_n(\eta)$ *is as in Proposition* 2.1.

PROOF. For any $|x| \le \delta\sqrt{n\log n}$, $a \le \eta \le b$ and sufficiently large n, there always exists an unique $\eta_{n,x}$ for which $\Lambda'(\eta_{n,x}) = \Lambda'(\eta) + x/n$. This is because $x/n \to 0$ and $\Lambda'(\cdot)$ is strictly increasing. By the same reason, there exist a_1 and b_1 satisfying $a_1 < a < b < b_1$and $\{\eta_{n,x};\ |x| \le \delta\sqrt{n\log n}\} \subset (a_1, b_1)$ for n large enough. Therefore

$$\begin{aligned} C_n(x,\eta)P\bigl(S_n \ge n\Lambda'(\eta) + x\bigr) &= C_n(x,\eta)P\bigl(S_n \ge n\Lambda'(\eta_{n,x})\bigr) \\ &= \frac{C_n(x,\eta)}{C_n(\eta_{n,x})}\bigl[C_n(\eta_{n,x})\bigl(P\bigl(S_n \ge n\Lambda'(\eta_{n,x})\bigr)\bigr)\bigr]. \end{aligned}$$

By Proposition 2.1, it remains to show that

$$\frac{C_n(x,\eta)}{C_n(\eta_{n,x})} \to 1 \tag{2.2}$$

uniformly in x and η. By Taylor's expansion and Lemma A.3,

$$\frac{C_n(x,\eta)}{C_n(\eta_{n,x})} = \frac{\eta\sqrt{\Lambda''(\eta)}}{\eta_{n,x}\sqrt{\Lambda''(\eta_{n,x})}}\exp\biggl[-\frac{x^3}{6n^2}(\Lambda^*)^{(3)}(\xi_\eta)\biggr], \tag{2.3}$$

where ξ_η is between $\Lambda'(\eta)$ and $\Lambda'(\eta) + n^{-1}x$. Obviously, we have that $|(x^3/6n^2) \times (\Lambda^*)^{(3)}(\xi_\eta)| \le n^{-1/2}(\log n)^{3/2}\sup_{a_1\le x\le b_1}|\Lambda^{*(3)}(x)|$. On the other hand, let

$h = \inf_{\eta\in[a_1,b_1]} \Lambda''(\eta)$. Then $h \in (0,\infty)$ and by the mean value theorem $\delta\sqrt{\log n/n} \ge |\Lambda'(\eta) - \Lambda'(\eta_{n,x})| \ge h|\eta - \eta_{n,x}|$. Apply the mean value theorem again to the function $x\sqrt{\Lambda''(x)}$ to obtain from (2.3)

$$\left|\frac{C_n(x,\eta)}{C_n(\eta_{n,x})} - 1\right| \le \frac{\delta}{(|a_1|+|b_1|)h} \sup_{a_1\le\eta\le b_1} \bigl|\bigl(\eta\sqrt{\Lambda''(\eta)}\bigr)'\bigr| \sqrt{\frac{\log n}{n}}.$$

Therefore, (2.2) is true. $\square$

The following easy fact is called Chernoff's bound (see, e.g., page 31 in [8]). It is weaker than Proposition 2.1, but it is a simple and nonasymptotic bound.

LEMMA A.4 (Chernoff's bound). *For any $x > EX$,*

$$P(S_n/n \ge x) \le \exp\bigl(-n\Lambda^*(x)\bigr) \qquad \forall n \ge 1.$$

The following lemma is frequently applied when proving theorems on maxima of partial sums via the Chen–Stein method.

LEMMA 2.4. *Suppose A, B and C are disjoint sets of indices. Let $\{X, X_\alpha; \alpha \in A\cup B\cup C\}$ be i.i.d. random variables with X satisfying condition* (1.1) *and $\mu := EX$. For any subset $D \subset A\cup B\cup C$, we use the notation $S_D := \sum_{\alpha\in D} X_\alpha$. Then,*

$$P(S_{A\cup B} \ge z, S_{B\cup C} \ge z) \le 2e^{-\theta z - m_1\zeta} \le 2e^{-\theta z - m_2\zeta},$$

where $\zeta = \sup_{\mu<x<0}\{\Lambda^*(x)\wedge\theta|x|\} > 0$, $m_1 = |A|\vee|C|$ and $m_2 = |A\cup C|/2$.

PROOF. Assume, without loss of generality, $|A| \ge |C|$. Then

$$P(S_{A\cup B} \ge z, S_{B\cup C} \ge z) \le P(S_A \ge x|A|)P(S_{B\cup C} \ge z) + P(S_B \ge z - x|A|)$$

for any $x > \mu$. By the Chernoff's bound, $P(S_A \ge x|A|) \le \exp(-|A|\Lambda^*(x))$. Also, by using that $E\exp(\theta X) = 1$ and the Markov's inequality, we have $P(S_{B\cup C} \ge z) \le e^{-\theta z}$ and $P(S_B \ge z - x|A|) \le \exp(-\theta z - \theta|x||A|)$. Therefore,

$$P(S_{A\cup B} \ge z, S_{B\cup C} \ge z) \le 2\exp\bigl\{-\theta z - |A|\bigl(\Lambda^*(x)\wedge\theta|x|\bigr)\bigr\}.$$

The lemma follows by choosing the smallest bound over $x \in (\mu, 0)$. $\square$

LEMMA 2.5. *Assume X satisfies condition* (1.1). *Let $d \in \mathbb{N}$ be a constant and $S_n = \sum_{i=1}^n X_i$. Then there exist constants $r > 1$ and $t_0 > \theta$ such that*

$$\sum_{k\ge rz^{1/d}} P(S_{k^d} \ge z) + \sum_{k\le r^{-1}z^{1/d}} P(S_{k^d} \ge z) = o(e^{-t_0 z}) \qquad \text{as } z\to\infty.$$

PROOF. Let $r_1 = (3\Lambda'(\theta)/2)^{1/d}$. Recall the definition of $I(x)$ in Lemma 2.3. It follows that $\lambda =: \inf_{x \geq 3\Lambda'(\theta)/2} I(x) > \theta$. Then by the Chernoff's bound and Lemma 2.3,

$$\sum_{k \leq r_1^{-1} z^{1/d}} P(S_{k^d} \geq z) \leq \sum_{k \leq r_1^{-1} z^{1/d}} e^{-k^d \Lambda^*(z/k^d)} \leq r_1^{-1} z^{1/d} e^{-\lambda z} = o(e^{-\lambda_1 z})$$

as $z \to \infty$, where $\lambda_1 = (\theta + \lambda)/2$. On the other hand, for any $c > 0$

$$\sum_{k \geq c z^{1/d}} P(S_{k^d} \geq z) \leq \sum_{k \geq c z^{1/d}} e^{-k^d \Lambda^*(0)} \leq \frac{e^{-c^d \Lambda^*(0) z}}{1 - e^{-\Lambda^*(0)}} = o(e^{-\lambda_1 z})$$

for any given $c > r_2 = (\lambda_1/\Lambda^*(0))^{1/d}$. Take $r = \max\{r_1, r_2 + 1\}$ to conclude the proof. $\square$

LEMMA 2.6. *Suppose condition* (1.1) *holds. For any two positive functions* $a(z)$ *and* $b(z)$ *such that* $(a(z) + b(z))/z^{1/d} \to 0$, *and two positive numbers* r, s *such that* $s < c_0 < r$, *where* $c_0 = (\Lambda'(\theta))^{-1/d}$, *we have that*

$$z^{1/2-1/d} e^{\theta z} \sum_{k \in \Gamma_z} P(S_{k^d} \geq z) = O\big(e^{-c(z)^2 z^{1-2/d}}\big) \qquad \text{as } z \to \infty,$$

where $\Gamma_z = \{k \in \mathbb{N};\ s z^{1/d} \leq k \leq c_0 z^{1/d} - b(z)$ *or* $c_0 z^{1/d} + a(z) \leq k \leq r z^{1/d}\}$ *and* $c(z) = a(z) \wedge b(z),\ z > 0$.

PROOF. Let $\Gamma_z' = \{k \in \mathbb{N}; c_0 z^{1/d} + a(z) \leq k \leq r z^{1/d}\}$. Then, by Proposition 2.1 and Lemma 2.3,

$$\begin{aligned}\sum_{k \in \Gamma_z'} P(S_{k^d} \geq z) &\leq C \sum_{k \in \Gamma_z'} \frac{1}{\sqrt{k^d}} \exp\left\{-k^d \Lambda^*\left(\frac{z}{k^d}\right)\right\} \\ &\leq C z^{-1/2+1/d} \exp\left\{-h(z) \Lambda^*\left(\frac{z}{h(z)}\right)\right\},\end{aligned}$$

where $h(z) = (c_0 z^{1/d} + a(z))^d$. Here the constant C, which depends on $\Lambda(\cdot)$ and d, may vary from line to line. Write $\Delta = z/h(z) - \Lambda'(\theta)$. By Taylor expansion $\Delta = -d c_0^{-1} a(z) z^{-1/d} + O(a(z)^2 z^{-2/d})$. By the Taylor expansion again and Lemma A.3,

$$\Lambda^*\big(z h(z)^{-1}\big) = \theta \Lambda'(\theta) + \Delta\theta + \frac{1}{2}\Delta^2\big(\Lambda''(\theta) + o(1)\big) = \frac{\theta z}{h(z)} + O\big(a(z)^2 z^{-2/d}\big).$$

Therefore, $h(z)\Lambda^*(z/h(z)) = \theta z + O(a(z)^2 z^{1-2/d})$. Hence,

$$z^{1/2-1/d} e^{\theta z} \sum_{k \in \Gamma_z'} P(S_{k^d} \geq z) = O\big(e^{-a(z)^2 z^{1-2/d}}\big).$$

By the same arguments, the above estimate is also true if Γ_z' and $a(z)$ are replaced by $\Gamma_z \setminus \Gamma_z'$ and $b(z)$, respectively. This completes the proof. $\square$

2.2. *Proofs.* Recall $\mathbb{N}$ is the set of all positive integers. $\mathbb{N}^d$ is the d-fold Cartesian product. Capital letters such as I, J, L, etc. will be used to denote points in $\mathbb{N}^d$. The notation $(i_1, i_2, \dots, i_d) = I \le J = (j_1, j_2, \dots, j_d)$ means that $i_l \le j_l$ for all $l = 1, 2, \dots, d$, and $I < J$ when all inequalities are strict. Also, as convention, $I + J = (i_1 + j_1, i_2 + j_2, \dots, i_d + j_d)$ and $m = (m, m, \dots, m) \in \mathbb{N}^d$. Let $\Delta = \Delta(I, J) = \{L \in \mathbb{N}^d;\ I \le L \le J\}$, $\mathcal{R}_n = \{\Delta = \Delta(I, J);\ 1 \le I \le J \le n\}$. We now turn to the proofs of Theorems 1, 2, 4, 5 and 6. To prove Theorem 4, we need the following two lemmas.

LEMMA 2.7. *Define $q_d(k) = \#\{(i_1, \dots, i_d) \in \mathbb{N}^d;\ i_1 i_2 \cdots i_d = k\}$. Then*

$$\sum_{k=1}^{m} q_d(k) \le m\big(\log(em)\big)^{d-1}$$

for $m \ge 2$ and $d \ge 2$.

PROOF. We prove the lemma by induction. When $d = 2$, it is easy to see that

$$\sum_{i=1}^{m} q_2(i) = \sum_{i=1}^{m} \left[\frac{m}{i}\right].$$

Note that

$$\sum_{k=p+1}^{q} \frac{1}{k} \le \log \frac{q}{p} \le \sum_{k=p}^{q-1} \frac{1}{k}$$

for any two positive integers $p < q$. Thus

$$\sum_{i=1}^{m} q_2(i) \le m \sum_{i=1}^{m} \frac{1}{i} \le m \log(em) \tag{2.4}$$

for all $m \ge 2$. So the lemma is true for $d = 2$. Observe that

$$\sum_{i=1}^{m} q_d(i) = \sum_{i=1}^{m} \sum_{k=1}^{[m/i]} q_{d-1}(k).$$

Suppose the lemma is true for $d = l \ge 2$. Then it is easy to check that

$$\begin{aligned}\sum_{k=1}^{m} q_{l+1}(k) &= \sum_{i=1}^{m} \sum_{k=1}^{[m/i]} q_l(k) \le \sum_{i=1}^{m} \left[\frac{m}{i}\right]\left(\log e\left[\frac{m}{i}\right]\right)^{l-1} \\ &\le m\big(\log(em)\big)^{l-1} \sum_{i=1}^{m} \frac{1}{i} \le m\big(\log(em)\big)^{l},\end{aligned}$$

where the last inequality is from (2.4). The proof is complete. $\square$

LEMMA 2.8. *Suppose that condition* (1.1) *holds*. *For any given* $\varepsilon > 0$, *there exists* $c > 0$ *such that* $c < \Lambda'(\theta) < 1/c$ *and*

$$\sum_{k \notin (cz, z/c)} q_d(k) P(S_k \geq z) = O\big(e^{-\varepsilon^{-1} z}\big)$$

as $z \to +\infty$.

PROOF. First, by Chebyshev's inequality and Lemmas 2.3 and 2.7, for any $c < \Lambda'(\theta)$, we have

$$\begin{aligned}\sum_{k \leq cz} q_d(k) P(S_k \geq z) &\leq \Big(\max_{k \leq cz} e^{-k\Lambda^*(z/k)}\Big) \sum_{k \leq cz} q_d(k) \leq cz(4\log cz)^{d-1} e^{-I(1/c)z} \\ &= O(e^{-\varepsilon z})\end{aligned}$$

for sufficiently large $c > 0$. By Lemma 2.7, we have that $q_d(k) \leq k(4\log k)^{d-1} \leq e^{\Lambda^*(0)k/2}$ for all k large enough. It follows that for given $c > 0$,

$$\sum_{k \geq z/c} q_d(k) P(S_k \geq z) \leq \sum_{k \geq z/c} q_d(k) e^{-k\Lambda^*(0)} \leq \frac{e^{-\Lambda^*(0)z/(2c)}}{1 - e^{-\Lambda^*(0)/2}}.$$

The result follows by choosing c appropriately. $\square$

PROOF OF THEOREM 1. Let $a(z) = \eta\sqrt{\log z}$, $\eta > 1$ and $A_z = \{k \in \mathbb{N};\ |k - \gamma(z)| \leq a(z)\}$. Then by Lemmas 2.5 and 2.6, we know that

$$\sqrt{z} e^{\theta z} \sum_{k \in A_z^c} P(S_{k^2} \geq z) = O\big(z^{1/2 - \eta^2}\big).$$

Therefore, to prove this theorem, we just need to prove the following two asymptotic formulas:

$$\sqrt{z} e^{\theta z} \delta(\gamma(z))^{-1} \sum_{k \in A_z} P(S_{k^2} \geq z) \sim \frac{1}{\theta}\sqrt{\frac{\Lambda'(\theta)}{2\pi\Lambda''(\theta)}}, \tag{2.5}$$

$$\sqrt{z} e^{\theta z} P\Big(\max_{k \in A_z} S_{k^2} \geq z\Big) \sim \sqrt{z} e^{\theta z} \sum_{k \in A_z} P(S_{k^2} \geq z). \tag{2.6}$$

By Corollary 2.1,

$$P(S_{k^2} \geq z) \sim \frac{1}{\theta\sqrt{2\pi k^2 \Lambda''(\theta)}} \exp\Big\{-\theta z - \frac{(z - \Lambda'(\theta)k^2)^2}{2\Lambda''(\theta)k^2}\Big\}$$

uniformly for all $k \in A_z$. Note that for these k's we also have that $1/k^2 = 1/c_0^2 z + O(a(z)/z^{3/2})$, $1/\sqrt{k^2} = 1/\sqrt{c_0^2 z} + O(a(z)/z)$ and $(z - \Lambda'(\theta)k^2)^2 = O(z^{-1/2}a(z)^2)$. Thus, it follows that

$$\begin{aligned} &\sqrt{z}e^{\theta z} \sum_{k \in A_z} P(S_{k^2} \ge z) \\ &\sim \frac{1}{\theta}\sqrt{\frac{\Lambda'(\theta)}{2\pi \Lambda''(\theta)}} \sum_{k \in A_z} \exp\biggl\{-\biggl(\frac{\Lambda'(\theta)}{2\Lambda''(\theta)}\biggr)\frac{(z - \Lambda'(\theta)k^2)^2}{z}\biggr\} + o\biggl(\frac{a(z)^2}{z}\biggr). \end{aligned} \tag{2.7}$$

By the definition of $\gamma(z)$, we have that

$$\begin{aligned} \frac{(z - \Lambda'(\theta)k^2)^2}{z} &= \Lambda'(\theta)^2 \frac{(k + \gamma(z))^2}{z}\bigl(k - \gamma(z)\bigr)^2 \\ &= 4\Lambda'(\theta)\bigl(k - \gamma(z)\bigr)^2 + O\bigl((\log z)^{3/2}/z\bigr) \end{aligned}$$

uniformly for all $k \in A_z$. Therefore

$$\sum_{k \in A_z} \exp\biggl\{-\biggl(\frac{\Lambda'(\theta)}{2\Lambda''(\theta)}\biggr)\frac{(z - \Lambda'(\theta)k^2)^2}{z}\biggr\} \sim \sum_{k \in A_z} e^{-\beta(k-\gamma(z))^2} + O\biggl(\frac{(\log z)^2}{z}\biggr).$$

Obviously, $\sum_{|k| \ge a(z)} e^{-\beta k^2} \le 2\sum_{k \ge a(z)} e^{-\beta a(z) k} = O(z^{-\beta\eta^2})$, which together with the above equality and (2.8), yields (2.5).

Now, we prove (2.6).

For any $(i, j) \in \Delta_z := \{(i, j) \in \mathbb{N}^2 : c_0\sqrt{z} - a(z) \le i < j \le c_0\sqrt{z} + a(z)\}$, set $C_{i,j} = (j^2 - i^2)|\mu|/2$. Then, $\min_{i,j \in \Delta_z} C_{i,j} \sim c_0|\mu|\sqrt{z}$ as $z \to \infty$. By Lemma 2.4 there exists a constant $C > 0$, so that

$$P(S_{i^2} \ge z, S_{j^2} \ge z) \le Cz^{-1/2}\exp\bigl\{-\theta z - C^{-1}\sqrt{\log z}\bigr\} \tag{2.8}$$

uniformly for all $i, j \in \Delta_z$. Therefore, by the inclusion-exclusion formula,

$$\begin{aligned} &\sqrt{z}e^{\theta z}\biggl(\sum_{k \in A_z} P(S_{k^2} \ge z) - P\Bigl(\max_{k \in A_z} S_{k^2} \ge z\Bigr)\biggr) \\ &\quad \le \sqrt{z}e^{\theta z} \sum_{(i,j) \in \Delta_z} P(S_{i^2} \ge z, S_{j^2} \ge z) \\ &\quad = O\bigl(e^{-C\sqrt{\log z}/2}\bigr). \end{aligned}$$

Then, (2.6) follows. $\square$

LEMMA 2.9. *Suppose condition* (1.1) *holds. Let* $g(t) = t\Lambda^*(z/t)$, $t > 0$. *Recall* $n = [(z/\Lambda'(\theta))^{1/d}]$, $d \ge 3$, *and* $G_k(z) = \exp(-g(k^d))$, $k \ge 1$, *as in Theorem* 2. *We have that*:

(i) $\liminf_{z\to+\infty} e^{\theta z+Cz^{1-2/d}}(G_n(z)+G_{n+1}(z))>0$ *for some constant* $C>0$.
(ii) $\liminf_{z\to+\infty} z^{-1+2/d}(g((n-1)^d)-g(n^d))>0$ *and* $\liminf_{z\to+\infty} z^{-1+2/d} \times(g((n+2)^d)-g((n+1)^d))>0$.

PROOF. (i) It is easy to check that

$$\frac{dg(t)}{dt}=\Lambda^*(z/t)-(z/t)(\Lambda^*)'(z/t), \qquad \frac{d^2g(t)}{dt^2}=(z^2/t^3)(\Lambda^*)''(z/t). \tag{2.9}$$

Let m be n or $n+1$ so that $|m-(z/\Lambda'(\theta))^{1/d}|\ge 1/2$. By Taylor's expansion and Lemma A.3,

$$g(m^d)=\theta z+\frac{1}{2}(\Lambda^*)''\bigl(\Lambda'(\theta)+o(1)\bigr)m^d\left(\frac{z}{m^d}-\Lambda'(\theta)\right)^2.$$

Then (i) follows from that $(x^d-y^d)/(x-y)>dy^{d-1}$ for any $x>y>0$.

(ii) We only need to prove the first $\liminf$ inequality. The second one is proved similarly. By (2.9), Taylor's expansion and Lemma A.3, $g(m^d)=\theta z+(C_X/z)(m^d-z/\Lambda'(\theta))^2(1+o(1))$ for some constant $C_X>0$ and $m=n-1,n$. Note that $(m^d-z/\Lambda'(\theta))^2(1/z)=O(z^{1-2/d})$ for $m=n-1,n$. Thus

$$g\bigl((n-1)^d\bigr)-g(n^d)=\frac{C_X}{z}\left\{\left((n-1)^d-\frac{z}{\Lambda'(\theta)}\right)^2-\left(n^d-\frac{z}{\Lambda'(\theta)}\right)^2\right\}+o(z^{1-2/d}).$$

Consequently, the desired result follows by using the formula that $a^2-b^2=(a+b)(a-b)$ and $n^d\le z/\Lambda'(\theta)<(n+1)^d$. $\square$

PROOF OF THEOREM 2. By Lemma 2.5, there exists $r>0$ and $t_0>\theta$ such that

$$\left|P\Bigl(\max_{k\ge1}S_{k^d}\ge z\Bigr)-P\Bigl(\max_{r^{-1}z^{1/d}\le k\le rz^{1/d}}S_{k^d}\ge z\Bigr)\right|\le e^{-t_0z}. \tag{2.10}$$

By Chernoff's bound and (iii) of Lemma A.3, there is a constant C_X such that

$$\begin{aligned}&\left|P\Bigl(\max_{r^{-1}z^{1/d}\le k\le rz^{1/d}}S_{k^d}\ge z\Bigr)-P\bigl(\max\{S_{n^d},S_{(n+1)^d}\}\ge z\bigr)\right|\\ &\qquad\le\Biggl(\sum_{r^{-1}z^{1/d}\le k\le n-1}+\sum_{n+2\le k\le rz^{1/d}}\Biggr)e^{-g(k^d)}\\ &\qquad\le C_Xz^{1/d}\bigl(e^{-g((n-1)^d)}+e^{-g((n+2)^d)}\bigr).\end{aligned} \tag{2.11}$$

By the same arguments as are used to obtain (2.8), we have that $P(S_{n^d}\ge z, S_{(n+1)^d}\ge z)=O(\exp(-\theta z-Cz^{1-1/d}))$ for some $C>0$. Therefore

$$\begin{aligned}&P(S_{n^d}\ge z)+P(S_{(n+1)^d}\ge z)-P\bigl(\max\{S_{n^d},S_{(n+1)^d}\}\ge z\bigr)\\ &\qquad=O\bigl(e^{-\theta z-Cz^{1-1/d}}\bigr).\end{aligned} \tag{2.12}$$

Also, by Proposition 2.1, $P(S_{m^d} \ge z) \sim \theta^{-1}\sqrt{\Lambda'(\theta)/2\pi\Lambda''(\theta)}\, z^{-1/2}G_m(z),\ m = n, n+1$. Collecting (2.10), (2.11) and (2.12), we obtain from Lemma 2.9 that

$$\sqrt{z}\big(G_n(z)+G_{n+1}(z)\big)^{-1}P\Big(\max_{k\ge 1} S_{k^d} \ge z\Big) \sim \frac{1}{\theta}\sqrt{\frac{\Lambda'(\theta)}{2\pi\Lambda''(\theta)}}. \qquad \square$$

PROOF OF THEOREM 4. To prove (i) and (ii) at the same time, it suffices to prove the following two inequalities:

$$\limsup_{n\to\infty}\frac{U_n}{\log n} \le \frac{d}{\theta} \quad \text{a.s.,} \qquad \liminf_{n\to\infty}\frac{W_n}{\log n} \ge \frac{d}{\theta} \quad \text{a.s.} \tag{2.13}$$

We first prove the lim sup inequality in (2.13).

Given $\eta > 0$, set $z_n = (1+\eta)d(\log n)/\theta$. Choose ε in Lemma 2.8 small enough so that $n^d e^{-\varepsilon^{-1}z_n} \le n^{-2}$. Note that the number of rectangles with the same upper-left corner and area k is at most $q_d(k)$. By Lemma 2.8, there exists $c > 0$ so that

$$\begin{aligned} P\Big(U_n \ge \frac{(1+\eta)d\log n}{\theta}\Big) &\le \frac{1}{n^2} + n^d \sum_{cz_n \le k \le z_n/c} q_d(k)P(S_k \ge z_n) \\ &\le \frac{1}{n^2} + n^d e^{-\theta z_n}\sum_{cz_n \le k \le z_n/c} q_d(k) = O\Big(\frac{(\log n)^2}{n^{\eta d}}\Big), \end{aligned}$$

where we use $E\exp(\theta X) = 1$ in the second inequality and Lemma 2.7 in the only equality above. Put $l_n = [n^{2/\eta d}]$. Then the above inequality implies that

$$P\Big(U_{l_n} \ge \frac{(1+\eta)d\log l_n}{\theta}\Big) = O\Big(\frac{(\log n)^2}{n^2}\Big). \tag{2.14}$$

The Borel–Cantelli Lemma implies that $\limsup_n U_{l_n}/\log l_n \le (1+\eta)d/\theta$ a.s. for any $\eta > 0$. Observe that U_n is increasing in n, and $l_{n+1}/l_n \to 1$. The lim sup inequality in (2.13) then follows.

Now, we turn to the proof of the lim inf inequality in (2.13). Set $k_n = [(c_1 d\log n)^{1/d}]$ and $m_n = [n/k_n]^d$, where $c_1 = (\theta\Lambda'(\theta))^{-1}$. Let $\{Y_i; 1 \le i \le m_n\}$ be i.i.d. random variables with the same law as that of $S_{k_n^d}$. We break the cube $\{1, 2, \dots, n\}^d$ into m_n many disjoint subcubes. Then, the partial sums of X_i's over these disjoint subcubes are i.i.d. Therefore, for any given $\eta \in (0, 1)$,

$$\begin{aligned} P\Big(W_n \le \frac{(1-\eta)d\log n}{\theta}\Big) &\le P\Big(\max_{1\le i\le m_n} Y_i \le \frac{(1-\eta)d\log n}{\theta}\Big) \\ &\le \exp\big(-m_n P\big(S_{k_n^d} \ge t_n\big)\big), \end{aligned}$$

where $t_n = (1-\eta)d\log n/\theta$. For any $\eta < 1/2$, we find $\delta > 0$ such that $\Lambda'(\delta) = (1-\eta/2)\Lambda'(\theta)$. Note that $t_n/k_n^d \to (1-\eta)\Lambda'(\theta)$. Then, by Proposition 2.1,

$$P(S_{k_n^d} \ge t_n) \ge P\Big(\frac{S_{k_n^d}}{k_n^d} \ge \Lambda'(\delta)\Big) \sim \frac{C(\delta)}{\sqrt{\log n}}e^{-k_n^d\Lambda^*(\Lambda'(\delta))}.$$

Now $\Lambda^*(x)$ is strictly increasing on $[0, \Lambda'(\theta)]$, hence $\Lambda^*(\Lambda'(\delta)) < \Lambda^*(\Lambda'(\theta)) = \theta\Lambda'(\theta)$, which implies that $m_n P(S_{k_n^d} \geq t_n) > n^{\eta'}$ for some $\eta' = \eta(\delta) > 0$ and all n large enough. Combining all the above inequalities, we obtain

$$P\left(W_n \leq \frac{(1-\eta)d\log n}{\theta}\right) \leq e^{-n^{\eta'}},$$

for all n large enough. It follows from the Borel–Cantelli Lemma that

$$\liminf_{n\to\infty} \frac{W_n}{\log n} \geq (1-\eta)\frac{d}{\theta} \qquad \text{a.s.}$$

for any $\eta > 0$ small enough. Then the lim inf inequality in (2.13) is proved. $\square$

PROOF OF THEOREM 5. Let $a_n = 2h(\log_2 n)^{1/2}, h > 1$, and $f_0 = (\theta\Lambda'(\theta)/2)^{-1/2}$. Denote by E_n the set of all subsquares in $\{1, 2, \dots, n\}^2$ with side lengths between $f_0\sqrt{\log n} - a_n$ and $f_0\sqrt{\log n} + a_n$. More precisely, $E_n = \{\Delta \in \mathcal{O}_n; |\sqrt{|\Delta|} - f_0\sqrt{\log n}| \leq a_n\}$. Define

$$\overline{W}_n = \max_{\Delta\in E_n} S_\Delta \quad \text{and} \quad z_n = \frac{1}{\theta}\left(2\log n - \frac{1}{2}\log_2 n + t_n\right) + x.$$

Throughout the paper, when we do computations with $\overline{W}_n$ or its counterparts, we always view it as two iterated maxima. The first maximum is that of S_Δ over all subcubes Δ with fixed upper-left corner, and then the second maximum is the maximum of the former ones over all n^2 corners. Based on this observation, by Lemmas 2.5 and 2.6,

$$P(W_n > z_n) - P(\overline{W}_n > z_n) = e^{-\theta x} O\big((\log n)^{-3}\big). \tag{2.15}$$

Now we use Lemma 2.2 to get the asymptotic distribution. First we need a lemma, as follows.

LEMMA 2.10. *$\delta((t + O(\log t))^{1/2})/\delta(t^{1/2}) \to 1$ as $t \to \infty$.*

PROOF. Just note that $\delta(t)$ is a positive, continuous and periodic function with a period 1. Also, $\inf_{t\in\mathbb{R}} \delta(t) > 0$. Obviously, $(t + O(\log t))^{1/2} = t^{1/2} + O((\log t)t^{-1/2})$. Then the conclusion follows from the uniform continuity of $\delta(t)$. $\square$

Let us continue the proof of Theorem 5. Let $\Omega_n = \{k \in \mathbb{N};\ |k - f_0\sqrt{\log n}| < a_n\}$. By (2.5) and Lemma 2.10, we have that

$$\begin{aligned}\lambda &:= \sum_{\Delta\in E_n} P(S_\Delta \geq z_n) = \sum_{k\in\Omega_n} (n-k+1)^2 P(S_{k^2} \geq z_n) \\ &= n^2 \sum_{k\in\Omega_n} P(S_{k^2} \geq z_n) + o(1) \qquad (2.16) \\ &\sim n^2 z_n^{-1/2} e^{-\theta z_n} \delta(\gamma(z_n))(1/\theta)\sqrt{\Lambda'(\theta)/2\pi\Lambda''(\theta)} \sim K_1 e^{-\theta x},\end{aligned}$$

where $K_1 = (1/2)\sqrt{\Lambda'(\theta)/\pi\theta\Lambda''(\theta)}$. By Lemma 2.2, to complete the proof, we just need to show that the corresponding b_1 and b_2 in the lemma go to zero. Actually, for any particular $\Delta \in E_n$, the number of squares which intersect Δ is at most $8|\Delta|a_n$. Moreover, $P(S_{k^2} \ge z_n) = e^{-\theta x} O(n^{-2}\sqrt{\log n})$ by using the fact $E\exp(\theta X) = 1$. Consequently,

$$b_1 \le \underbrace{n^2 \cdot 8|\Delta|a_n|\Omega_n|^2}_{O(n^2(\log n)^2)} \max_{(i,j)\in(\Omega_n)^2} P(S_{i^2} \ge z_n)P(S_{j^2} \ge z_n) = O\big(n^{-2}(\log n)^3\big).$$

Similarly,

$$b_2 \le O\big(n^2(\log n)^2\big) \max_{\Delta_1 \ne \Delta_2, \Delta_1, \Delta_2 \in E_n} P(S_{\Delta_1} \ge z_n, S_{\Delta_2} \ge z_n).$$

For any two $\Delta_1, \Delta_2 \in E_n$, $\Delta_1 \ne \Delta_2$, their symmetric difference, that is, $(\Delta_1 \backslash \Delta_2) \cup (\Delta_2 \backslash \Delta_1)$ is at least $f_0\sqrt{\log n} - a_n$ (this is the key observation in handling such type of high-dimensional problems in this paper). By Lemma 2.4, $P(S_{\Delta_1} \ge z_n, S_{\Delta_2} \ge z_n) \le n^{-2}\exp(-(f_0/2)\sqrt{\log n})$ for all n large enough. Thus, $b_2 = O((\log n)^2 \exp(-(f_0/2)\sqrt{\log n}))$. $\square$

PROOF OF THEOREM 6. Obviously, k_n, depends on n, is either $[((\log n)/\alpha)^{1/d}]$ or $[((\log n)/\alpha)^{1/d}]+1$. Define E_n as the set of all subcubes in $\{1, 2, \dots, n\}^d$ with side length k_n, that is, $E_n = \{\Delta \in \mathcal{O}_n; |\Delta|^{1/d} = k_n\}$. Also,

$$\overline{W}_n = \max_{\Delta \in E_n} S_\Delta.$$

Let $z_n := \Lambda'(\theta)k_n^d + x$. By Lemma 2.9 and a proof similar to those of (2.10) and (2.11) [replacing $\max\{S_{n^d}, S_{(n+1)^d}\}$ in (2.11) by $S_{k_n^d}$], we have that

$$P(W_n \ge z_n) - P(\overline{W}_n \ge z_n) \to 0$$

as $n \to \infty$. Hence, to prove our theorem, it is enough to show that

$$P\big(\overline{W}_n \le \Lambda'(\theta)k_n^d + x\big) - \exp(-K_2 r_n e^{-\theta x}) \to 0. \tag{2.17}$$

Note that

$$\lambda_n := \sum_{\Delta \in E_n} P(S_{k^d} > z_n) = (n - k_n + 1)^d P(S_{k_n^d} \ge z_n).$$

But, by Corollary 2.1,

$$P(S_{k_n^d} \ge z_n) \sim \frac{e^{-\theta x}}{\theta\sqrt{2\pi\Lambda''(\theta)}} \exp\bigg\{-\frac{d}{2}\log k_n - k_n^d \theta\Lambda'(\theta)\bigg\} = K_2 r_n n^{-d} e^{-\theta x}.$$

Now we prove the theorem by Lemma 2.2, that is, the Chen–Stein method. It is easy to check

$$b_1 \le n^d(2k_n)^d P(S_{k_n^d} > z_n)^2 = O(\log n/n^d).$$

For any two different overlapping subcubes with side lengths k_n, their symmetric difference is at least $2k_n^{d-1}$. By Lemma 2.4,

$$b_2 \le n^d (2k_n)^d \max_{\Delta_1, \Delta_2 \in E_n, \Delta_1 \ne \Delta_2} P(S_{\Delta_1} > z_n,\ S_{\Delta_2} > z_n)$$
$$= O\big((nk_n)^d \exp\{-\theta \Lambda'(\theta) k_n^d - \zeta k_n^{d-1}\}\big)$$

for some constant $\zeta > 0$. By definition, $\theta \Lambda'(\theta) k_n^d \ge d \log n - (d/2) \log k_n$. It follows that,

$$b_2 = O\big((\log n)^{d/2}\big) e^{-\zeta (\log n)^{1-1/d}}.$$

Therefore, (2.17) follows. $\square$

3. Proofs of Theorems 3 and 7. For any constant α, denote $E_z = \{(p,q) \in \mathbb{N}^2;\ |pq - \Lambda'(\theta)^{-1} z| < \alpha \sqrt{z \log z}\,\}$ and

$$U_z^1 = \max_{(p,q) \in E_z} S_{p,q}.$$

The definition of $S_{p,q}$ is given before the statement of Theorem 3 in the Introduction.

LEMMA 3.1. *Suppose condition* (1.1) *holds. Then*

$$P(U > z) - P(U_z^1 > z) = O(1/z), \qquad z \to \infty$$

for large α.

PROOF. Recall $q(k) = \#\{(r,s) \in \mathbb{N}^2;\ \ rs = k\}$. Obviously, $q(k) \le \sqrt{k}$. Therefore

$$P(U > z) - P(U_z^1 > z) \le \sum_{(p,q) \notin E_z} P(S_{pq} \ge z) \le \sum_{k \notin \Omega_z} \sqrt{k} P(S_k \ge z),$$

where $\Omega_z = \{k \in \mathbb{N};\ \ |k - \Lambda'(\theta)^{-1} z| < \alpha \sqrt{z \log z}\}$ and $S_k = \sum_{i=1}^n X_{1,i}$. By comparing $\sum_{n \notin \Omega_z} \sqrt{n} P(S_n \ge z)$ with similar expressions in Lemmas 2.5 and 2.6 ($d = 1$), we see that the only difference between them is that $\sqrt{n}$ appears in the former term. But this term does not dominate the sum. So by checking the proofs of Lemmas 2.5 and 2.6 ($d = 1$), we have that

$$\sum_{n \notin \Omega_z} \sqrt{n} P(S_n \ge z) \le \frac{C}{z^{\alpha^2}} + \frac{1}{z}$$

for some constant $C > 0$. $\square$

LEMMA 3.2. *Suppose condition* (1.1) *holds. Recall* $E_z = \{(p,q) \in \mathbb{N}^2;\ |pq - \Lambda'(\theta)^{-1}z| < \alpha\sqrt{z\log z}\ \}$. *Then for any* $\alpha > 0$,

$$\lim_{z\to+\infty} e^{\theta z}(\log z)^{-1} \sum_{(p,q)\in E_z} P(S_{p,q} \ge z) = \frac{1}{\theta\sqrt{\Lambda'(\theta)}}.$$

PROOF. We write $z = \Lambda'(\theta)pq + \Lambda'(\theta)(z/\Lambda'(\theta) - pq)$. Then by Corollary 2.1, we have

$$P(S_{p,q} > z) \sim \frac{e^{-\theta z}}{\theta\sqrt{2\pi\Lambda''(\theta)pq}} \exp\left(-\frac{\Lambda'(\theta)^2}{2pq\Lambda''(\theta)}\left(\frac{z}{\Lambda'(\theta)} - pq\right)^2\right) \tag{3.1}$$

uniformly for all $p, q \in E_z$. For simplicity, set $y = z/\Lambda'(\theta)$. Then $E_z = \{(p,q) \in \mathbb{N}^2;\ |pq - y| < \alpha\sqrt{z\log z}\ \}$. Thus,

$$\begin{aligned} &\theta\sqrt{2\pi\Lambda''(\theta)/\Lambda'(\theta)}e^{\theta z}(\log z)^{-1} \sum_{(p,q)\in E_z} P(S_{p,q} \ge z) \\ &\quad\sim \frac{1}{\sqrt{z}\log z} \sum_{(p,q)\in E_z} \exp\left(-\tilde{K}\frac{(pq-y)^2}{y}\right) \\ &\quad= \frac{1}{\sqrt{z}\log z} \sum_{k\in\Omega_z} q(k)\exp\left(-\tilde{K}\frac{(k-y)^2}{y}\right), \end{aligned}$$

where $\tilde{K} = \Lambda'(\theta)^2/2\Lambda''(\theta)$ and $\Omega_z = \{k \in \mathbb{N};\ |k - y| \le \alpha\sqrt{z\log z}\}$. To complete the proof, we need to show that

$$\frac{1}{\sqrt{z}\log z} \sum_{k\in\Omega_z} q(k)\exp\left(-\tilde{K}\frac{(k-y)^2}{y}\right) \to \sqrt{\frac{\pi}{\tilde{K}\Lambda'(\theta)}}. \tag{3.2}$$

Given any $\gamma \in (0,1)$, let $\Delta = \gamma\sqrt{y/\log y}$ and

$$A_i = \{k \in \mathbb{N};\ y + i\Delta < k \le y + (i+1)\Delta\}, \qquad i = -i_z, -i_z + 1, \ldots, i_z,$$

where $i_z = [\alpha\sqrt{z\log z}/\Delta] \sim \alpha\gamma^{-1}\sqrt{\Lambda'(\theta)}\log y$. Since $\max_{k\in\Omega_z} q(k) = O(\sqrt{z})$,

$$\begin{aligned} &\sum_{k\in\Omega_z} q(k)\exp\left(-\tilde{K}\frac{(k-y)^2}{y}\right) \\ &\quad= \sum_{i=-i_z}^{i_z}\sum_{k\in A_i} q(k)\exp\left(-\tilde{K}\frac{(k-y)^2}{y}\right) + O(\sqrt{z}). \end{aligned} \tag{3.3}$$

Now we estimate the part $\sum_{k\in A_i}$ in (3.3). Note that for any $k \in A_i$,

$$e^{-\tilde{K}(k-y)^2/y} - e^{-\tilde{K}(i\Delta)^2/y} = e^{-\tilde{K}(i\Delta)^2/y}(e^{\phi_k(y)} - 1),$$

where $\phi_k(y) := \tilde{K}(k - y + i\Delta)(k - y - i\Delta)/y$. It is easy to check that $\rho_i := \max_{k \in A_i} |\phi_k(y)| \le \tilde{K}(2|i| + 1)\Delta^2/y \le C\gamma$ for some constant C depending on X and α only. Therefore, since $|e^x - 1| \le |x|e^{|x|}$ for any $x \in \mathbb{R}$, we have that

$$\sum_{k \in A_i} q(k) e^{-\tilde{K}(k-y)^2/y} = (1 + \rho_i') e^{-\tilde{K}(i\Delta)^2/y} \sum_{k \in A_i} q(k), \tag{3.4}$$

where $\rho_i' := e^{\phi_k(y)} - 1$. By Theorem 320 on page 264 of [12]

$$\sum_{i=1}^{n} q(i) = n \log n + cn + O(\sqrt{n}),$$

where c here is a universal positive constant. Therefore, for any $m = m_z \sim n$,

$$\sum_{i=m+1}^{n} q(k) = (n - m) \log n + (c + 1)(n - m) + O(\sqrt{n}). \tag{3.5}$$

As a consequence,

$$\sum_{k \in A_i} q(k) = \sum_{[y+i\Delta]+1}^{[y+(i+1)\Delta]} q(k) = \gamma \sqrt{y \log y} + O(\sqrt{y}). \tag{3.6}$$

By (3.3), (3.4) and (3.6), we obtain

$$\frac{1}{\sqrt{z} \log z} \sum_{k \in \Omega_z} q(k) e^{-\tilde{K}(k-y)^2/y} \sim \frac{\gamma}{\sqrt{\Lambda'(\theta) \log y}} \sum_{-i_z}^{i_z} (1 + \rho_i'') e^{-\tilde{K}(i\gamma)^2/\log y}, \tag{3.7}$$

where $\rho_i'' = \rho' + O(1/\log z) \le C\gamma$. Because of the monotonicity of e^{-x^2} on $(0, +\infty)$, it is not difficult to see that

$$\frac{\gamma}{\sqrt{\log y}} \sum_{-i_z}^{i_z} e^{-\tilde{K}(i\gamma)^2/\log y} \to \int_{-\infty}^{\infty} e^{-\tilde{K}t^2}\, dt = \sqrt{\frac{\pi}{\tilde{K}}} \qquad \forall \gamma \in (0, \infty). \tag{3.8}$$

It follows that

$$\limsup_{y \to +\infty} \frac{\gamma}{\sqrt{\log y}} \sum_{-i_z}^{i_z - 1} \rho_i' e^{-\tilde{K}(i\gamma)^2/\log y} \le C\gamma. \tag{3.9}$$

Thus, combining (3.7), (3.8) and (3.9), we obtain that

$$\limsup_{y \to +\infty} \left| \frac{1}{\sqrt{z} \log z} \sum_{k \in \Omega_z} q(k) e^{-\tilde{K}(k-y)^2/y} - \sqrt{\frac{\pi}{\tilde{K}\Lambda'(\theta)}} \right| \le C\gamma$$

for arbitrary given $\gamma > 0$. Let $\gamma \downarrow 0$, then (3.2) follows. $\square$

LEMMA 3.3. *Suppose condition* (1.1) *holds. Given any* $\alpha > 0$ *and any function* $p_z > 0$, *let*

$$E_z^5 = \{(p,q) \in \mathbb{N}^2;\ |pq - \Lambda'(\theta)^{-1}z| \le \alpha\sqrt{z\log z},\ p \wedge q \ge p_z\}$$

and E_z *be the same as in Lemma* 3.2. *Then*

$$e^{\theta z}(\log z)^{-1} \sum_{(p,q)\in E_z \setminus E_z^5} P(S_{p,q} \ge z) = O\left(\frac{\log p_z}{\sqrt{\log z}}\right).$$

PROOF. Set $\beta_z^{\pm} = z/\Lambda'(\theta) \pm \alpha\sqrt{z\log z}$. Then by Corollary 2.1, we know that $P(S_{p,q} \ge z) \le C_\theta z^{-1/2}e^{-\theta z}$, where C_θ is a constant depending on θ. Set $B_{z,p} = \{q \in \mathbb{N};\ \beta_z^- \le pq \le \beta_z^+\}$. Then

$$\begin{aligned}\sum_{(p,q)\in E_z \setminus E_z^5} P(S_{p,q} \ge z) &\le 2 \sum_{\substack{1\le p\le p_z,\\ (p,q)\in E_z}} P(S_{p,q} \ge z) \le 2C_\theta z^{-1/2}e^{-\theta z} \sum_{p=1}^{p_z} \sum_{q\in B_{z,p}} 1 \\ &\le C_\theta z^{-1/2}e^{-\theta z}(\beta_z^+ - \beta_z^-) \sum_{p=1}^{p_z} \frac{1}{p} \\ &= O(\sqrt{\log z}\, e^{-\theta z} \log p_z). \qquad \square\end{aligned}$$

PROOF OF THEOREM 3. Combining Lemmas 3.1, 3.2 and 3.3 we have that for any $\varepsilon > 0$ there is $\alpha > 0$ such that

$$\limsup_{z\to\infty} e^{\theta z}(\log z)^{-1}\big(P(U \ge z) - P(U_z^5 \ge z)\big) \le \varepsilon, \tag{3.10}$$

where $U_z^5 = \max_{(p,q)\in E_z^5} S_{p,q}$ and E_z^5 is as in Lemma 3.3 with $p_z = \exp(\varepsilon\sqrt{\log z})$. We claim that

$$\lim_{z\to\infty} e^{\theta z}(\log z)^{-1}\left(P(U_z^5 \ge z) - \sum_{(p,q)\in E_z^5} P(S_{p,q} \ge z)\right) \to 0 \qquad \forall\, \alpha > 0. \tag{3.11}$$

If the claim is true, then by Lemmas 3.2, 3.3 and (3.10),

$$\limsup_{z\to\infty}\left|P(U > z) - \frac{1}{\theta\sqrt{\Lambda'(\theta)}}\right| \le 2\varepsilon$$

for any $\varepsilon > 0$. Therefore the theorem follows by letting $\varepsilon \downarrow 0$. Now we prove the claim. Observe that

$$P\left(\max_{\Gamma\in E_z^5} S_\Gamma \ge z\right) \ge \sum_{\Gamma\in E_z^5} P(S_\Gamma \ge z) - \sum_{\Gamma_1,\Gamma_2\in E_z^5,\Gamma_1\ne\Gamma_2} P(S_{\Gamma_1} \ge z, S_{\Gamma_2} \ge z).$$

To prove the claim (3.11), it suffices to show that

$$e^{\theta z}(\log z)^{-1} \sum_{\Gamma_1,\Gamma_2\in E_z^5,\Gamma_1\neq\Gamma_2} P(S_{\Gamma_1}\geq z, S_{\Gamma_2}\geq z)\to 0 \tag{3.12}$$

as $z\to+\infty$. By the definition of E_z^5, the size of the symmetric difference of any Γ_1 and Γ_2, that is, $|\Gamma_1\Delta\Gamma_2|$ is at least $\exp(\varepsilon\sqrt{\log z})$. Thus, by Lemma 2.4, there is $\zeta>0$ such that

$$P(S_{\Gamma_1}\geq z, S_{\Gamma_2}\geq z)\leq 2\exp\Big(-\theta z-\zeta e^{\varepsilon\sqrt{\log z}}\Big), \tag{3.13}$$

as z is sufficiently large. Therefore

$$\sum_{\Gamma_1,\Gamma_2\in E_z^5,\Gamma_1\neq\Gamma_2} P(S_{\Gamma_1}\geq z, S_{\Gamma_2}\geq z)\leq 2|E_z|^2\exp\Big(-\theta z-\delta e^{\varepsilon\sqrt{\log z}}\Big),$$

where $E_z=\{(p,q)\in\mathbb{N}^2;\ |pq-\Lambda'(\theta)^{-1}z|<\alpha\sqrt{z\log z}\ \}$ is as before. By (3.5), we have that

$$|E_z|\leq \sum_{\beta_z^-\leq i\leq\beta_z^+} q(i)=O\Big(\sqrt{z}(\log z)^{3/2}\Big),$$

where β_z^- and β_z^+ are as in the proof of Lemma 3.3. It follows that

$$\sum_{\Gamma_1,\Gamma_2\in E_z^5,\Gamma_1\neq\Gamma_2} P(S_{\Gamma_1}\geq z, S_{\Gamma_2}\geq z)=O\Big(z(\log z)^3\exp\Big(-\theta z-\delta e^{\varepsilon\sqrt{\log z}}\Big)\Big),$$

which implies (3.12). □

PROOF OF THEOREM 7. Denote $z_n=(2\log n+\log_3 n)/\theta+x$. Take $p_{z_n}=e^{(\log_2 z_n)^{1/4}}$ in the definition of E_z^5 in Lemma 3.3. Of course $E_{z_n}^5$ is a subsequence of E_z^5. Let $\mathcal{R}_n^1$ be the set of all the rectangles in $\{1,2,\dots,n\}^2$ whose length and width, say, p,q, satisfy $(p,q)\in E_{z_n}^5$. Accordingly, $W_n^1:=\max_{\Delta\in\mathcal{R}_n^1} S_\Delta$. By Lemmas 3.1, 3.2 and 3.3, there is $\alpha>0$ such that

$$e^{\theta z}(\log z)^{-1}\sum_{(p,q)\in\mathbb{N}^2\setminus E_z^5} P(S_{p,q}\geq z)=O\big((\log z)^{-1/4}\big) \tag{3.14}$$

for large z, where E_z^5 is as in Lemma 3.3 corresponding to $p_z=\exp((\log z)^{1/4})$. As before, we view $\mathcal{R}_n$ as the union of rectangles with fixed upper-left corners for all such possible corners. It follows by (3.14) that

$$\sum_{\Delta\in\mathcal{R}_n\setminus\mathcal{R}_n^1} P(S_\Delta\geq z_n)=O\big(n^2e^{-\theta z_n}(\log z_n)^{3/4}\big)=O\big((\log_2 n)^{-1/4}\big).$$

Thus, to complete the proof, we just need to prove that

$$P(W_n^1\geq z_n)\to 1-e^{-K_3e^{-\theta x}}. \tag{3.15}$$

First, it is easy to see that

$$(n-L_n)^2 \sum_{(p,q)\in E_{z_n}} P(S_{p,q} \ge z_n) \tag{3.16}$$
$$\le \sum_{\Delta\in\mathcal{R}_n^1} P(S_\Delta \ge z_n) \le n^2 \sum_{(p,q)\in E_{z_n}} P(S_{p,q} \ge z_n),$$

where $L_n := \max\{p \vee q;\ \ (p,q) \in E^5_{z_n}\}$. Obviously, $L_n \le \Lambda'(\theta)^{-1} z_n + \alpha\sqrt{z_n \log z_n} = O(\log n)$. By Lemmas 3.2 and 3.3,

$$n^2 \sum_{(p,q)\in E_{z_n}} P(S_{p,q} \ge z_n) \sim \frac{n^2 e^{-\theta z_n} \log z_n}{\theta\sqrt{\Lambda'(\theta)}} \sim \frac{e^{-\theta x}}{\theta\sqrt{\Lambda'(\theta)}}. \tag{3.17}$$

Thus, (3.16) and (3.17) imply that $\lambda_n := \sum_{\Delta\in\mathcal{R}_n^1} P(S_\Delta \ge z_n) \to e^{-\theta x}/\theta\sqrt{\Lambda'(\theta)}$ as $n \to \infty$.

Now we use the Chen–Stein method to complete the proof.

For any $\Delta \in \mathcal{R}_n^1$, define $\mathcal{A}_\Delta = \{\Delta' \in \mathcal{R}_n^1;\ \ \Delta' \cap \Delta \ne \varnothing\}$. It is easy to see that $|\mathcal{A}_\Delta| = O((\log n)\log_2 n)$. By Lemma 2.2, to prove (3.15), we need to verify that b_1 and b_2 in the lemma go to zero. Recall $P(S_\Delta \ge z_n) \le e^{-\theta z_n}$ for all $\Delta \in \mathcal{R}_n^1$. Then

$$b_1 \le n^2 |\mathcal{A}_\Delta| \max_{\Delta'\in\mathcal{A}_\Delta} P\,(S_{\Delta'} \ge z_n)^2 = O\biggl(\frac{(\log n)\log_2 n}{n^2}\biggr).$$

By (3.13),

$$P\,(S_\Delta \ge z_n,\ S_{\Delta'} \ge z_n) = O\bigl(n^{-2}(\log_2 n)^{-1}\exp\bigl(-\zeta e^{(\log_2 n)^{1/4}}\bigr)\bigr)$$

for some $\zeta > 0$ uniformly for all $\Delta,\ \Delta' \in \mathcal{R}_n^1$. Since $|\mathcal{R}_n^1| = O(n^2(\log n)\log_2 n)$, it follows that

$$b_2 \le |\mathcal{R}_n^1||\mathcal{A}_\Delta| \max_{\Delta,\Delta'\in\mathcal{R}_n^1} P(S_\Delta \ge z_n,\ S_{\Delta'} \ge z_n)$$
$$= O\Bigl((\log n)^2(\log_2 n)^2 \exp\bigl(-\zeta e^{(\log_2 n)^{1/4}}\bigr)\Bigr). \qquad \square$$

4. Proofs of Theorems 8 and 9. We first prove an inequality on empirical processes which will be used later. We review some basic definitions and facts about empirical processes. Let S be a set and $\mathcal{G}$ a class of subsets of S. Define $\Delta^{\mathcal{G}}(s_1,\dots,s_n) = \#\{G \cap \{s_1,\dots,s_n\};\ G \in \mathcal{G}\}$ for any $\{s_1,\dots,s_n\} \subset S$. Also, let

$$m^{\mathcal{G}}(n) = \max\{\Delta^{\mathcal{G}}(s_1,\dots,s_n);\ s_i \in S, i = 1,2,\dots,n\}$$

and

$$V(\mathcal{G}) = \inf\{n : m^{\mathcal{G}}(n) < 2^n\}.$$

Of course, by convention, $V(\mathcal{G}) = +\infty$ if $m^{\mathcal{G}}(n) = 2^n$ for all $n \geq 1$. Dudley [9] calls $\mathcal{G}$ a Vapnik–Červonenkis (VC for short) class of sets if $V(\mathcal{G}) < \infty$. The quantity $V(\mathcal{G})$ is called the exponent of the VC class $\mathcal{G}$. A result on VC class given by Vapnik–Červonenkis is the following inequality from [23]:

$$m^{\mathcal{G}}(n) \leq n^{V(\mathcal{G})} \qquad \forall n \geq 2. \tag{4.1}$$

LEMMA 4.1. *Let* $\mathcal{L}_d = \{\prod_{i=1}^d [a_i, b_i] \subset [0,1]^d;\ 0 \leq a_i \leq b_i \leq 1, i = 1, 2, \ldots, d\}$ *be a VC class for any* $d \geq 1$. *Then*, $\mathcal{F}$, *as in* (1.4) *and as a subset of* $\mathcal{L}_3$, *is a VC class with some exponent* v_3. *Therefore*, *for any* $(y_1, \ldots, y_n) \in \mathbb{R}^n$ *and* $n \geq 2$, $\#\{(y_1, y_2, \ldots, y_n) \cap F;\ F \in \mathcal{F}\} \leq n^{v_3}$.

PROOF. We just need to prove the first part of the lemma. We show it by induction. Obviously, $\mathcal{L}_1$ is a VC class with exponent 3.

Suppose the lemma is true for any $1, 2, \ldots, d-1$. We now prove that $\mathcal{L}_d$ is also a VC class. Define $v_k = V(\mathcal{L}_k)$, $k \geq 1$. For any $n_d := 2d v_{d-1} + 1$ distinctive points $y_1, \ldots, y_{n_d}$, let $\prod_{i=1}^d [a_i, b_i]$ be the smallest rectangular solid to contain those n_d points. If there is a point, say y_1, in the interior of $\prod_{i=1}^d [a_i, b_i]$, then no rectangular solid can contain $\{y_2, \ldots, y_{n_d}\}$ without y_1. If there is no such point, then $\{y_1, \ldots, y_{n_d}\}$ must be in the following $2d$ sets: $\{a_1\} \times [a_2, b_2] \times \cdots \times [a_d, b_d]$, $\{b_1\} \times [a_2, b_2] \times \cdots \times [a_d, b_d], \ldots, [a_1, b_1] \times \cdots \times [a_{d-1}, b_{d-1}] \times \{a_d\}$ and $[a_1, b_1] \times \cdots \times [a_{d-1}, b_{d-1}] \times \{b_d\}$. Then, there is such a set, say, $\{a_1\} \times [a_2, b_2] \times \cdots \times [a_d, b_d]$, containing at least $v_{d-1} + 1$ points of $\{y_1, \ldots, y_{n_d}\}$. By assumption, $\{a_1\} \times [a_2, b_2] \times \cdots \times [a_d, b_d]$ is a VC class with exponent no more than v_{d-1}, so no $(n-1)$-dimensional rectangular solid can separate these $v_{d-1} + 1$ points. This implies that $m^{\mathcal{L}_d}(n_d) < 2^{n_d}$. Actually, it is easily seen that $v_d \leq 2^d (d+1)!$. $\square$

Suppose that $\{Y, Y_i;\ i \geq 1\}$ is a sequence of i.i.d. random variables with the uniform distribution on $[0,1]^3$. For any $C > 0$, define

$$\mathcal{F}_{n,1} = \{B(x, \varepsilon) \in \mathcal{F},\ x \in (0,1)^3,\ 0 < \varepsilon \leq C(\log n/n)^{1/3}\}.$$

LEMMA 4.2. *For any class of subsets* $\mathcal{C}$, *define* $\Gamma_{\mathcal{C}} = \#\{\{Y_1, \ldots, Y_n\} \cap F;\ F \in \mathcal{C}\}$. *Then there is a constant* $v > 0$ *such that for any* $C > 0$, *there is a constant* $D > 0$ *for which*

$$P(\#\Gamma_{\mathcal{F}_{n,1}} \geq Dn(\log n)^v) = O(n^{-3}).$$

PROOF. Let $r_n = (\log n/n)^{1/3}$ and $\mathcal{G}_i = \{B(Y_i, Cr_n) \cap F; F \in \mathcal{F}\}$ for $i = 1, 2, \ldots, n$. Since $\mathcal{F}$ is a VC class with exponent no greater than $v = v_3$, so is $\mathcal{G}_i$. By (4.1),

$$\#\Gamma_{\mathcal{G}_1} \leq \left\{\sum_{i=1}^n I_{B(Y_1, Cr_n)}(Y_i)\right\}^v. \tag{4.2}$$

Therefore, for $t > 0$,

$$P(\#\Gamma_{\mathcal{F}_{n,1}} \geq t) \leq n P(\#\Gamma_{\mathcal{G}_1} \geq t/n) \leq n P\left(\sum_{i=1}^{n} I_{B(Y_1, Cr_n)}(Y_i) \geq (t/n)^{1/v}\right).$$

Substituting $D^v n(\log n)^v$ for t in the above inequality, we have that

$$P(\#\Gamma_{\mathcal{F}_{n,1}} \geq D^v n(\log n)^v) \leq n P\left(\sum_{i=2}^{n} I_{B(Y_1, Cr_n)}(Y_i) \geq D \log n - 1\right). \tag{4.3}$$

By Lemma A.2, for any $D > 20C^v$,

$$P\left(\sum_{i=2}^{n} I_{B(Y_1, Cr_n)}(Y_i) \geq D \log n - 1\right) \leq 2e^{-KD\log n}$$

for large n, where K is a constant depending only on C. The above inequality and (4.3) yield the desired inequality by choosing D sufficiently large. $\square$

Let $\{X, X_i;\ i \geq 1\}$ be a sequence of i.i.d. $\mathcal{X}$-valued random variables with law P, where $\mathcal{X}$ is a metric space. Let P_n be the empirical law of $\{X_n\}$, that is,

$$P_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}.$$

We assume that $\mathcal{H}$, a class of subsets of $\mathcal{X}$, is a VC class with exponent υ. Let $\{\mathcal{H}_n \subset \mathcal{H};\ n \geq 1\}$ be a sequence of subclasses of sets and $\mathcal{H}_n$ be countable for each $n \geq 1$. Set $\mu_n := \inf_{V \in \mathcal{H}_n} P(V)$ and

$$f(P_n) = \sup_{V \in \mathcal{H}_n} \frac{|P_n(V) - P(V)|}{\sqrt{P(V)(1 - P(V))/n}}.$$

The classical exponential inequality (see, e.g., page 16 of [16]) is

$$P\left(\sup_{V \in \mathcal{H}} \left|\sum_{i=1}^{n} P_n(V) - P(V)\right| \geq \varepsilon\right) \leq 8n^v e^{-n\varepsilon^2/32}. \tag{4.4}$$

It is not sharp enough in our later proofs because we need to take $\{P(V), V \in \mathcal{H}\}$ into account. The following inequality provides us with a result for this case.

LEMMA 4.3. *Suppose* $\sup_{V \in \mathcal{H}_n} P(V) < 1 - \delta_0$ *for some* $\delta_0 \in (0, 1)$, *then*

$$P(f(P_n) \geq t_n) \leq \frac{11n^{\upsilon+1}}{(1 - \delta_0)t_n^2 \mu_n} \exp\left\{-\frac{\delta_0 t_n^2}{32}\left(1 + \frac{2t_n}{\sqrt{n\mu_n}}\right)^{-1}\right\}$$

for any positive t_n *satisfying* $n \geq \mu_n t_n^2 + 2$.

PROOF. Let $\{\varepsilon_i;\ i \geq 1\}$ be a sequence of i.i.d. Bernoulli sequence. By the argument of (11) on page 15 of [16],

$$P\big(f(P_n) \geq t_n\big) \leq 4P\left(\sup_{V \in \mathcal{H}_n} \frac{|\sum_{i=1}^n \varepsilon_i I_V(X_i)|}{\sqrt{P(V)(1-P(V))/n}} > \frac{t_n}{4}\right) \tag{4.5}$$

for all $n \geq 1$, where $I_V(\cdot)$ is the indicator function of V. By Hoeffding's inequality (see, e.g., page 91 of [16]), $P(|\sum_{i=1}^m \varepsilon_i| \geq x) \leq 2\exp(-x^2/2m)$ for all $x > 0$. It follows that

$$P^{\varepsilon}\left(\frac{|\sum_{i=1}^n \varepsilon_i I_V(X_i)|}{\sqrt{P(V)(1-P(V))/n}} > \frac{t_n}{4}\right) \leq 2\exp\left\{-\frac{\delta_0 t_n^2}{32}\left(\frac{P(V)}{P_n(V)}\right)\right\}. \tag{4.6}$$

Set $A_n = \{f(P_n) \leq 2t_n\}$. Note that by (4.1), $\#\{\{X_1,\dots,X_n\} \cap V;\ V \in \mathcal{H}_n\} \leq n^{\upsilon}$ for $n \geq 2$. It follows from (4.5) and (4.6) that

$$\begin{aligned} P\big(f(P_n) \geq t_n\big) &\leq 4n^{\upsilon} E^X \sup_{V \in \mathcal{H}_n} P^{\varepsilon}\left(\frac{|\sum_{i=1}^n \varepsilon_i I_V(X_i)|}{\sqrt{P(V)(1-P(V))/n}} > \frac{t_n}{4}\right) I_{A_n} + 4P(A_n^c) \\ &\leq 8n^{\upsilon} \exp\left\{-\frac{\delta_0 t_n^2}{32}\left(1 + \frac{2t_n}{\sqrt{n\mu_n}}\right)^{-1}\right\} + 4P\big(f(P_n) > 2t_n\big) \end{aligned} \tag{4.7}$$

for $n \geq 2$. Repeat (4.7) to obtain

$$\begin{aligned} P(f(P_n) \geq t_n) \leq{}& \sum_{l=0}^{k} 8n^{\upsilon} \cdot 4^l \exp\left\{-\frac{4^l \delta_0 t_n^2}{32}\left(1 + \frac{2^l t_n}{\sqrt{n\mu_n}}\right)^{-1}\right\} \\ &+ 4^{k+1} P\big(f(P_n) > 2^{k+1} t_n\big). \end{aligned} \tag{4.8}$$

Note that $f(P_n) \leq \{n/(\mu_n(1-\delta_0))\}^{1/2}$ for all $n \geq 1$. Let $k_0 = [\log_4(n/(\mu_n t_n^2(1-\delta_0)))]$. Then $2^{k_0+1} t_n > f(P_n)$. Consequently, the probability in the right-hand side of (4.8) is zero. Since $x^2/(1+xy)$ is increasing in $x \in (0,\infty)$ for any $y > 0$ and $\sum_{l=0}^{k_0} 4^l \leq 4^{k_0+1}/3$, by (4.8),

$$P(f(P_n) \geq t_n) \leq \frac{11n^{\upsilon+1}}{(1-\delta_0)t_n^2 \mu_n} \exp\left\{-\frac{\delta_0 t_n^2}{32}\left(1 + \frac{2t_n}{\sqrt{n\mu_n}}\right)^{-1}\right\}$$

for all n such that $n \geq 2$ and $k_0 \geq 0$. The fact that $n \geq \mu_n t_n^2$ implies that $k_0 \geq 0$. $\square$

Before proving Theorem 8, we need the following lemma.

LEMMA 4.4. *Suppose condition* (1.1) *holds. Let $\mathcal{F}$ and $\mathcal{F}_{n,1}$ be same as in Lemmas* 4.1 *and* 4.2, *respectively. Then, there exists $C > 0$ such that*

$$P\left(\max_{B \in \mathcal{F} \setminus \mathcal{F}_{n,1}} S_n(B) \geq 0\right) = O(n^{-2}).$$

PROOF. Set $A_n = \{\sum_{i=1}^n I_B(Y_i) \geq 7nr^3\}$ for any $B = B(x, r) \in \mathcal{F}$. Since by Lemma 4.1, $\mathcal{F}$ is a VC class with exponent v_3,

$$P\left(\max_{B\in\mathcal{F}\setminus\mathcal{F}_{n,1}} S_n(B) \geq 0\right) \leq n^{v_3} E^Y\left\{\max_{B\in\mathcal{F}\setminus\mathcal{F}_{n,1}} P^X\big(S_n(B) \geq 0\big)(I_{A_n} + I_{A_n^c})\right\}.$$

For each $B(x, r) \in \mathcal{F} \setminus \mathcal{F}_{n,1},\ 7nr^3 \geq 7nr_n^3 \sim 7C^3 \log n$ as $n \to \infty$. It follows that

$$P^X(S_n(B) \geq 0) I_{A_n} \leq \max_{n \geq 4C^3 \log n} P(S_n \geq 0) \leq e^{-4C^3 \Lambda^*(0) \log n}$$

for large n. The last inequality is from Chernoff's bound. Therefore

$$n^{v_3} E^Y\left\{\max_{B\in\mathcal{F}\setminus\mathcal{F}_{n,1}} P^X\big(S_n(B) \geq 0\big) I_{A_n}\right\} \leq n^{v_3 - 4C^3\Lambda^*(0)} \leq 1/n^2 \tag{4.9}$$

if $C \geq ((v_3 + 2)/(4\Lambda^*(0)))^{1/3}$. On the other hand, note

$$E^Y\left\{\max_{B\in\mathcal{F}\setminus\mathcal{F}_{n,1}} P^X\big(S_n(B) \geq 0\big) I_{A_n^c}\right\} \leq E^Y\left\{\max_{B\in\mathcal{F}_{n,2}} I_{A_n^c}\right\} + E^Y\left\{\max_{B\in\mathcal{F}_{n,3}} I_{A_n^c}\right\},$$

where

$$\mathcal{F}_{n,2} = \{B(x, \varepsilon) \in \mathcal{F}, x \in (0, 1]^3, \varepsilon \in [1/3, 1/2)\}$$

and

$$\mathcal{F}_{n,3} = \{B(x, \varepsilon) \in \mathcal{F}, x \in (0, 1]^3, \varepsilon \in [r_n, 1/3]\}.$$

Apparently, $A_n^c \subset \{|(1/n)\sum_{i=1}^n 1_B(Y_i) - \mathrm{Vol}(B)| \geq r^3\}$. Therefore,

$$\begin{aligned} E^Y\left\{\max_{B\in\mathcal{F}_{n,2}} I_{A_n^c}\right\} &\leq P\left(\sup_{B\in\mathcal{F}_{n,2}} \left|\sum_{i=1}^n 1_B(Y_i) - n\mathrm{Vol}(B)\right| \geq \frac{n}{27}\right) \\ &\leq 8n^{v_3} e^{-n/2^{15}}, \end{aligned} \tag{4.10}$$

where the last inequality follows from (4.4) because $\mathcal{F}_{n,2}$ is a VC class with exponent no larger than v_3. Finally, by Lemma 4.3,

$$\begin{aligned} P\left(\bigcup_{B\in\mathcal{F}_{n,3}} A_n^c\right) &\leq P\left(\sup_{B\in\mathcal{F}_{n,3}} \frac{|(1/n)\sum_{i=1}^n I_B(Y_i) - \mathrm{Vol}(B)|}{\sqrt{P(B)(1-P(B))/n}} > \sqrt{\frac{nr_n^3}{8}}\right) \\ &\leq c_1 n^8 e^{-c_1 C^3 \log n} \end{aligned}$$

for sufficiently large n and some universal constant c_1 (there is no problem in applying Lemma 4.3 because the above "sup" equals that over all subcubes in $\mathcal{F}_{n,3}$ with rational radius almost surely). The above inequality together with (4.9) and (4.10) yields the desired result by choosing a large C. □

PROOF OF THEOREM 8. We only need to prove that

$$\limsup_{n\to\infty} \frac{\tilde{U}_n}{\log n} \le \frac{1}{\theta} \quad \text{a.s.} \quad \text{and} \quad \liminf_{n\to\infty} \frac{\tilde{W}_n}{\log n} \ge \frac{1}{\theta} \quad \text{a.s.} \tag{4.11}$$

By Lemma 4.4, choose C so that

$$P\Big(\max_{B\in\mathcal{F}\setminus\mathcal{F}_{n,1}} S_n(B) \ge 0\Big) = O(n^{-2}).$$

Therefore, by the Borel–Cantelli lemma, to prove the limsup inequality in (4.11), it is enough to show that

$$\limsup_{n\to\infty} \frac{\max_{B\in\mathcal{F}_{n,1}} S_n(B)}{\log n} \le \frac{1}{\theta}. \tag{4.12}$$

For any $\varepsilon > 0$, let $q = 2 + 2\varepsilon^{-1}$. Define

$$V_n = \max_{n^q \le k \le (n+1)^q} \max_{B\in\mathcal{F}_{k,1}} S_k(B), \qquad n = 1, 2, \dots.$$

By Lemma 4.2, choose D such that $\Xi_k := \{\#\Gamma_{\mathcal{F}_{k,1}} \ge Dk(\log k)^6\}$ has probability $O(k^{-2})$. It is easy to check that $\#\{(S_1(B), S_2(B), \dots, S_m(B)); B \in \mathcal{G}\} = \#\{S_m(B);\ B \in \mathcal{G}\}$ for any VC class $\mathcal{G}$. This is because the set $\{Y_1, Y_2, \dots, Y_m\} \cap B$ determines sets $\{Y_1, Y_2, \dots, Y_k\} \cap B$, $k = 1, 2, \dots, m-1$ for any B and m. Therefore, since $\mathcal{F}_{k,1}$ is decreasing,

$$\begin{aligned}
&P\big(V_n > \overbrace{(1+\varepsilon)(\log(n+1)^q)/\theta}^{x_n}\big)\\
&\quad\le E^Y\bigg\{P^X\Big(\max_{B\in\mathcal{F}_{n^q,1}} \max_{n^q\le k\le(n+1)^q} S_k(B) > x_n\Big) I_{\Xi^c_{n^q}}\bigg\} + O(n^{-2q})\\
&\quad\le 2Dq^6 n^q (\log n)^6 P\Big(\max_{1\le k\le(n+1)^q} S_k > x_n\Big) + O(n^{-2q}) = O(n^{-2}),
\end{aligned}$$

where we use the submatingale inequality and the fact $E\exp(\theta X) = 1$ in the last inequality. By the same arguments as in (2.14), we obtain (4.12).

Now we turn to prove the liminf inequality of (4.11). For any integer p (which will be chosen specifically later), let $s_{n,p} = (\log n^p/8n^p\theta\Lambda'(\theta))^{1/3}$ and

$$\begin{aligned}
L_n &= \bigcup_{i=n^p+1}^{(n+1)^p} B(Y_i, 2s_{n,p}) \quad \text{and}\\
J_n &= \{1 \le j \le n^p;\ Y_j \in [s_{n,p}, 1-s_{n,p}]^3 \setminus L_n\}.
\end{aligned} \tag{4.13}$$

Note that L_n may not be necessarily a subset of $[0,1]^3$ although with a large probability it is. Evidently,

$$\inf_{n^p<k\le(n+1)^p} \Big\{\max_{B\in\mathcal{B}} S_k(B)\Big\} \ge \max_{j\in J_n} S_{n^p}(B_j), \tag{4.14}$$

where $B_j := B(Y_j, s_{n,p})$. Define

$$N_n = \max\left\{k;\ \exists \text{ different } i_1, \dots, i_k \in J_n \text{ such that } \inf_{1\le s<t\le k} \|Y_{i_s} - Y_{i_t}\| > 2s_{n,p}\right\}.$$

It is not difficult to check that $N_n \le 8^{-1}s_{n,p}^{-3} = \theta\Lambda'(\theta)p^{-1}(n^p/\log n)$ deterministically. We claim that the reverse is almost true in the sense that there exists a constant $C' > 0$ so that

$$P\left(N_n \le \frac{C'n^p}{\log n}\right) = O(n^{-2}) \tag{4.15}$$

for large p. Indeed, list all subcubes $\prod_{i=1}^3[(3k_i+1)s_{n,p}, (3k_i+2)s_{n,p}]$, $0 \le k_i \le [(3s_{n,p})^{-1}] - 1$, $i = 1, 2, 3$, as $A_1, A_2, \dots, A_{m_n}$. It is easy to check that $\inf_{x\in A_i,\ y\in A_j} d(x,y) > 2s_{n,p}$ a.s. for all pairs $i \ne j$. Obviously, $m_n = 2C'n^p(\log n)^{-1} + O(1)$ for some constant $C' > 0$. Pick all those A_i such that $Y_l \notin A_i$ for all $l = n^p+1, \dots, (n+1)^p$, and list them again as $\tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_{l_n}$. Then $l_n \ge m_n - ((n+1)^p - n^p) = 2C'n^p(\log n)^{-1} + O(n^{p-1})$. So, by Chebyshev's inequality,

$$\begin{aligned} &P\left(\sum_{i=1}^{l_n} I\big(\text{at least one of } \{Y_j, 1\le j\le n^p\} \in \tilde{A}_i\big) \le \frac{C'n^p}{\log n}\right) \\ &\qquad \le \frac{l_n}{l_n - C'n^p(\log n)^{-1}}(1 - s_{n,p}^3)^{n^p} = O(n^{-2}) \end{aligned} \tag{4.16}$$

for $p > 16\theta\Lambda'(\theta)$. Thus claim (4.15) follows.

For any given $\varepsilon \in (0, 1/4)$, set $b_n = (1-\varepsilon)\log n^p/\theta$. Then, by symmetry and (4.15),

$$\begin{aligned} P\left(\max_{j\in J_n} S_{n^p}(B_j) \le b_n\right) &\le E^Y \prod_{j=1}^{N_n} P^X\big(S_{n^p}(B_j) \le b_n\big) I(Y_j \in [s_{n,p}, 1-s_{n,p}]^3) \\ &\le E^Y \exp\left\{-\sum_{j=1}^{N_n} P^X\big(S_{n^p}(B_j) > b_n\big) I(Y_j \in [s_{n,p}, 1-s_{n,p}]^3)\right\} \\ &\qquad \times I\left(N_n \ge \frac{C'n^p}{\log n}\right) + O(n^{-2}) \\ &\le \max_{C'n^p/\log n \le k \le n^p} \zeta_k + O(n^{-2}), \end{aligned}$$

where

$$\zeta_k = E^Y\left\{\exp\left(-\sum_{i=1}^k P^X\big(S_{n^p}(B_i) \ge b_n\big) I(Y_j \in [s_{n,p}, 1-s_{n,p}]^3)\right)\right\}.$$

Lemma 4.5 below will prove that $\max_{C'n^p/\log n \le k \le n^p} \zeta_k = O(n^{-2})$ for some $p = p_\varepsilon$. So by the Borel–Cantelli lemma, $\max_{j \in J_n} S_{n^p}(B_j)/\log n^p \ge (1-\varepsilon)/\theta$ eventually. Combine this with (4.14) to complete the proof of liminf inequality of (4.11). $\square$

LEMMA 4.5. *Suppose condition* (1.1) *holds. Set* $G_n = \{k \in \mathbb{N};\ C'n^p/\log n \le k \le n^p\}$. *For any given* $\varepsilon \in (0, 1/4)$, *there exists* $p > 0$ *such that* $\max_{k \in G_n} \zeta_k = O(n^{-2})$.

PROOF. Recall $B_i = B(Y_i, s_{n,p})$. Set $N_i = \#\{1 \le j \le n^p;\ Y_j \in B_i\}$, $1 \le i \le k$. By Bernstein's inequality (Lemma A.2), for any given $x \in (0, 1)$,

$$P\Bigg(N_i \notin \underbrace{\bigg(\frac{(1-x)\log n^p}{\theta\Lambda'(\theta)}, \frac{(1+x)\log n^p}{\theta\Lambda'(\theta)}\bigg)}_{O_x}\Bigg) \le \exp\{-(px^2/6\theta\Lambda'(\theta))\log n\} \tag{4.17}$$

uniformly for $i = 1, 2, \dots, k$ and for large n. It follows from Chebyshev's inequality similar to (4.16) that

$$\sup_{k \in G_n} P\Bigg(\sum_{i=1}^{k} I_{O_x}(N_i) \le \frac{C'n^p}{2\log n}\Bigg) = O(n^{-2}) \tag{4.18}$$

for $p \ge 12\theta\Lambda'(\theta)/x^2$. If $N_i \in O_x$, by Proposition 2.1, there is a constant C_x such that

$$P^X(S_{n^p}(B_i) > b_n) \ge \min_{l \in O_x} \frac{C_x e^{-l\Lambda_X^*(b_n/l)}}{\sqrt{l}} \tag{4.19}$$

for large n. Since $\Lambda^*(t)$ is increasing on $(0, +\infty)$, $\lim_{x \to 0^+}\{\max_{l \in O_x} l\Lambda_X^*(b_n/l)/\log n^p\} = \Lambda^*((1-\varepsilon)\Lambda'(\theta))/\theta\Lambda'(\theta) < 1$ uniformly for any $n \ge 2$ by Lemma A.3. So, for the given $\varepsilon \in (0, 1/4)$, choose a suitable $x_0 \in (0, 1)$ in the definition of O_x such that

$$\inf_{i;\, N_i \in O_{x_0}} P^X(S_{n^p}(B_i) > b_n) \ge C_{p,x_0} n^{\alpha_\varepsilon p}/\sqrt{\log n}$$

for some constant C_{p,x_0} and $\alpha_\varepsilon \in (0, 1)$ and large n. Consequently, by (4.18), we have that

$$\max_{k \in G_n} \zeta_k \le \exp\{-Cn^{(1-\alpha_\varepsilon)p}(\log n)^{-3/2}\} + O(n^{-2}).$$

The desired equality follows by taking p sufficiently large. $\square$

We need the following two lemmas to prove Theorem 9. Define

$$l_n^{\pm} = \frac{1}{2}\bigg(\frac{\log n}{\theta\Lambda'(\theta)n}\bigg)^{1/3}\Bigg(1 \pm \beta\sqrt{\frac{\log_2 n}{\log n}}\Bigg),$$

where the sign on the left-hand side matches the sign on the right-hand side. Let $\Omega_n = \{(r,i) \in [l_n^-, l_n^+] \times \{1,2,\dots,n\};\ B(Y_i, l_n^+) \subset [0,1]^3\}$ and

$$W_{n,1} = \max_{(r,i)\in\Omega_n} S_n\big(B(Y_i,r)\big).$$

LEMMA 4.6. *Suppose condition* (1.1) *holds. Let $z_n = \log n/\theta + x$. Then, for any $\alpha > 0$, there exists $\beta > 0$ such that*

$$P(W_n \geq z_n) - P(W_{n,1} \geq z_n) = O\big((\log n)^{-\alpha}\big).$$

PROOF. By Lemma 4.4, it is enough to show that

$$P\Big(\max_{(r,i)\in\Omega_n'} S_n\big(B(Y_i,r)\big) > z_n\Big) = O\big((\log n)^{-\alpha}\big)$$

for some $\beta > 0$, where $\Omega_n' = \{(r,i) \in ((0, Lr_n) \setminus (l_n^-, l_n^+)) \times \{1,2,\dots,n\};\ B(Y_i, l_n^+) \subset [0,1]^3\}$ for some constant C_L. By Lemma 4.2,

$$\begin{aligned} &E^Y\Big\{P^X\Big(\max_{(r,i)\in\Omega_n'} S_n\big(B(Y_i,r)\big) > z_n\Big)\Big\} \\ &\quad\leq n(\log n)^6 E^Y\Big\{\max_{(r,i)\in\Omega_n'} P^X\big(S_n\big(B(Y_i,r)\big) > z_n\big)\Big\} + O(n^{-2}). \end{aligned} \tag{4.20}$$

Set $c_0 = (\theta\Lambda'(\theta))^{-1/3}$. Note that $|n\,\mathrm{vol}(B(Y_i,r)) - c_0^3 \log n| \geq 2\beta c_0^3((\log n) \times \log_2 n)^{1/2}$ for any $(r,i) \in \Omega_n'$ when n is sufficiently large. Consequently, for such (r,i),

$$\begin{aligned} &\left\{\left|\sum_{j=1}^n I_{B(Y_i,r)}(Y_j) - n\,\mathrm{vol}\big(B(Y_i,r)\big)\right| \leq (\beta c_0^3)\sqrt{(\log n)\log_2 n}\right\} \\ &\qquad \subset \left\{\left|\sum_{j=1}^n I_{B(Y_i,r)}(Y_j) - c_0^3 \log n\right| \geq (\beta c_0^3)\sqrt{(\log n)\log_2 n}\right\} := \Delta_n. \end{aligned}$$

So, by Bernstein's inequality, $P(\Delta_n^c) \leq 2e^{-\beta C_L \log_2 n}$ for some constant C_L and large n. It follows that for any $(r,i) \in \Omega_n'$,

$$P^X\big(S_n\big(B(Y_i,r)\big) > z_n\big) \leq P^X\big(S_n\big(B(Y_i,r)\big) > z_n\big) I_{\Delta_n} + \frac{2e^{-x}}{n} I_{\Delta_n^c}, \tag{4.21}$$

where the second term above is obtained by using the Chernoff's bound for $P^X(S_n(B(Y_i,r)) > z_n)$. By Corollary 2.1, there exists a constant C_X such that

$$P^X\big(S_n\big(B(Y_i,r)\big) > z_n\big) I_{\Delta_n} \leq \frac{e^{-\theta x}}{n(\log n)^{C_X\beta^2+1/2}}$$

uniformly on $(r,i)\in\Omega'_n$. Combining this with (4.20) and (4.21), we finally get

$$P\Big(\max_{(r,i)\in\Omega'_n} S_n\big(B(Y_i,r)\big)>z_n\Big)\le\frac{e^{-\theta x}}{(\log n)^{C_X\beta^2-5.5}}+\frac{4e^{-x}}{(\log n)^{\beta C_L-6}}$$

for n large. The proof is completed by choosing β sufficiently large. $\square$

LEMMA 4.7. *Suppose condition* (1.1) *holds. Let* $T_n=\{1\le i\le n;\ \ Y_i\in[l_n^+,1-l_n^+]^3\}$. *Then*,

$$\lambda_n:=\sum_{i\in T_n}P^X\Big(\max_{l_n^-\le r\le l_n^+}S_n\big(B(Y_i,r)\big)>z_n\Big)\to Ke^{-\theta x}\qquad \textit{in probability},$$

where K is as in Lemma A.1.

PROOF. First, recall $S_n=\sum_{i=1}^n X_i$, by Lemma A.1,

$$\lambda_n\le nP\Big(\max_{i\ge1}S_n>z_n\Big)\to Ke^{-\theta x}\qquad\text{a.s.}\tag{4.22}$$

By Bernstein's inequality (Lemma A.2),

$$P\big(T_n\le n(1-2l_n^+)^3-\log n\big)\le n^{-\xi}\tag{4.23}$$

for some constant $\xi>0$. Define $h_n=(8n(l_n^+)^3\log_2 n)^{1/2}$ and

$$A_i=\Bigg\{\Bigg|\sum_{j=1}^n I(\|Y_j-Y_i\|\le l_n^+)-8n(l_n^+)^3\Bigg|\le h_n\Bigg\},\qquad i=1,2,\dots,n.$$

It is easy to check that

$$\max_{i\in T_n}P(A_i^c)\le2\exp\{-C_X\log_2 n\}$$

for some constant C_X and large n. Consequently, by Chebyshev's inequality,

$$P\Bigg(\sum_{i\in T_n}I(A_i)\le T_n-nP(A_1^c)^{1/2}\Bigg)=O\big((\log n)^{-C_X/2}\big)\qquad\text{a.s.}\tag{4.24}$$

Also, $8n\{(l_n^+)^3-(l_n^-)^3\}\sim(6\beta/\theta\Lambda'(\theta))\sqrt{(\log n)\log_2 n}$. Set $h'_n=((\log n)\log_2 n)^{0.3}$ and

$$L_i=\Bigg\{\Bigg|\sum_{j=1}^n I(l_n^-\le\|Y_j-Y_i\|\le l_n^+)-8n\big((l_n^+)^3-(l_n^-)^3\big)\Bigg|\le h'_n\Bigg\},$$
$$i=1,2,\dots,n.$$

By Bernstein's inequality again,

$$\max_{i\in T_n}P(L_i^c)\le2\exp\big\{-C_X(\log n)^{0.1}\big\}\qquad\text{a.s.}\tag{4.25}$$

for n large. Thus, by an argument similar to that used in establishing (4.24), we have

$$P\left(\sum_{i\in T_n} I(L_i) \le T_n - nP(L_1^c)^{1/2}\right) \tag{4.26}$$
$$= O\big(\exp(-(C_X/2)(\log n)^{0.1})\big) \qquad \text{a.s.}$$

From (4.23), (4.24) and (4.26), it follows that with probability approaching 1, at least $n - o(n)$ of $\{Y_i; 1 \le i \le n\}$ satisfy the following conditions:

(a) fall into $[l_n^+, 1 - l_n^+]^3$;

(b) every box centered at every such Y_i and with radius l_n^+ contains at least $8n(l_n^+)^3 - h_n \sim (\theta\Lambda'(\theta))^{-1}(\log n + 3\beta\sqrt{(\log n)\log_2 n})$ elements of $\{Y_1, Y_2, \dots, Y_n\}$;

(c) For every such Y_i, $B(Y_i, l_n^+) \setminus B(Y_i, l_n^-)$ contains at least $(6/\theta\Lambda'(\theta)) \times \beta\sqrt{(\log n)\log_2 n}$ elements of $\{Y_1, Y_2, \dots, Y_n\}$.

By Lemmas A.1, 2.5 and 2.6, there exists $\gamma > 0$ for which

$$nP\left(\max_{k\in Q_n} S_k \ge z_n\right) \to Ke^{-\theta x}, \tag{4.27}$$

where $Q_n = \mathbb{N} \cap \{k;\ |k - \log n/\theta\Lambda'(\theta)| \le \gamma\sqrt{(\log n)\log_2 n}\}$. Therefore, by (a), (b), (c) and the definition of λ_n,

$$P^Y\big(\lambda_n \ge \big(n - o(n)\big)(1/n)\{Ke^{-\theta x} + o(1)\}\big) \to 1$$

for sufficiently large β, which together with (4.22) proves the lemma. $\square$

PROOF OF THEOREM 9. We continue to use the notation of Lemmas 4.6 and 4.7. Define $V_{n,i} = \max_{l_n^- \le r \le l_n^+} S_n(B(Y_i, r))$, $i = 1, 2, \dots, n$. Then $W_{n,1} = \max_{i\in T_n} V_{n,i}$. By Lemma 4.6, to prove the theorem, it is enough to show that

$$P\left(\max_{i\in T_n} V_{n,i} > z_n\right) \to 1 - e^{-Ke^{-\theta x}}. \tag{4.28}$$

By Lemma 2.2, we have

$$\left|P\left(\max_{i\in T_n} V_{n,i} > z_n\right) - 1 + E^Y e^{-\lambda_n}\right| \le b_1' + b_2',$$

where

$$b_1' = E^Y\left\{\sum_{i\in T_n}\sum_{j\in T_n} P^X(V_{n,j} > z_n)P^X(V_{n,i} > z_n)I\big(d(Y_j, Y_i) \le 2l_n^+\big)\right\}$$

and

$$b_2' = E^Y\left\{\sum_{i\in T_n}\sum_{j\in T_n} P^X(V_{n,j} > z_n, V_{n,i} > z_n) I\big(d(Y_j, Y_i) \le 2l_n^+\big)\right\}.$$

Evidently, $E^Y e^{-\lambda_n} \to \exp(-Ke^{-\theta x})$ by Lemma 4.7. Also,

$$b_1' \le \frac{e^{-2\theta x}}{n^2} n^2 P\big(d(Y_1, Y_2) \le 2l_n^+\big) = e^{-2\theta x} O(\log n/n)$$

since $P^X(V_{n,j} > z_n) \le e^{-\theta x}/n$. Moreover,

$$b_2' \le n^2 E^Y\{P^X(V_{n,1} > z_n, V_{n,2} > z_n)(I_{\Psi_n} + I_{\Psi_n'})\},$$

where

$$\Psi_n = \{n^{-1}(\log n)^{-\delta} \le d(Y_1, Y_2) \le 2l_n^+,\ Y_1, Y_2 \in [l_n^+, 1 - l_n^+]^3\}$$

and

$$\Psi_n' = \{d(Y_1, Y_2) < n^{-1/3}(\log n)^{-\delta},\ Y_1, Y_2 \in [l_n^+, 1 - l_n^+]^3\}$$

for some $\delta \in (0, 1/6)$. Obviously,

$$n^2 E^Y\{P^X(V_{n,1} > z_n, V_{n,2} > z_n) I_{\Psi_n'}\} \le 8e^{-\theta x}(\log n)^{-3\delta}. \tag{4.29}$$

Define $\kappa_n =$ the volume of $B(Y_2, l_n^-) \setminus B(Y_1, l_n^+)$. It is easy to check that on Ψ_n,

$$\kappa_n \ge (2l_n^-)^2\{(n^{-1/3}(\log n)^{-\delta} - (l_n^+ - l_n^-)\} \sim Cn^{-1}(\log n)^{2/3-\delta} \tag{4.30}$$

for some constant $C > 0$ because $l_n^+ - l_n^- = o(n^{-1/3}(\log n)^{-\delta})$ with $\delta \in (0, 1/6)$. By Bernstein's inequality, conditionally on $Y_1, Y_2 \in [l_n^+, 1 - l_n^+]^3$,

$$\begin{aligned} P\Bigg(&\underbrace{\sum_{i=1}^n I_{B(Y_2, l_n^-)\setminus B(Y_1, l_n^+)}(Y_i) \le n\kappa_n - (\log n)^{7/12-\delta/2}}_{E_n}\Bigg) \\ &\le \exp\big(-C(\log n)^{1/6-\delta}\big). \end{aligned} \tag{4.31}$$

Recall the definition of L_i appearing between (4.24) and (4.25). We have, by (4.25),

$$\begin{aligned} &n^2 E^Y\{P^X(V_{n,1} > z_n, V_{n,2} > z_n) I_{\Psi_n}\} \\ &\quad \le C_{X,\beta} n^2 (\log n) \\ &\qquad \times \log_2 n E^Y\left\{\max_{l_n^- \le r_1, r_2 \le l_n^+} P^X\big(S_n(B(Y_i, r_i)) > z_n,\ i = 1, 2\big) I_{\Psi_n \cap L_1 \cap L_2}\right\} \\ &\qquad + \underbrace{n^2 P^{Y_1, Y_2}(\Psi_n)(n^{-1}e^{-\theta x}) \cdot 4e^{-C_X(\log n)^{0.1}}}_{O((\log n)e^{-(\log n)^{0.1}})}. \end{aligned} \tag{4.32}$$

For any $r_1, r_2 \in (l_n^-, l_n^+)$, on Ψ_n, $B(Y_2, r_1) \setminus B(Y_1, r_2) \supset B(Y_2, l_n^-) \setminus B(Y_1, l_n^+)$. Therefore, by Lemma 2.4, (4.30) and (4.31), there exists a positive ζ depending on X and β such that

$$P^X\big(S_n(B(Y_i, r_i)) > z_n,\ i = 1, 2\big) I_{E_n^c} \le n^{-1} e^{-\zeta(\log n)^{2/3-\delta}}$$

for any fixed $Y_1, Y_2 \in [l_n^+, 1 - l_n^+]^3$. Consequently, the first term of (4.32) is less than

$$\begin{aligned} &Cn^2(\log n)(\log_2 n) P^{Y_1,Y_2}(\Psi_n)\Big(n^{-1}e^{-\zeta(\log n)^{2/3-\delta}} + n^{-1}e^{-\theta x}e^{-C(\log n)^{1/6-\delta}}\Big) \\ &\qquad = O\Big((\log n)^3 e^{-\zeta(\log n)^{1/6-\delta}}\Big). \end{aligned}$$

Combining the above equality with (4.32) and then (4.29), we conclude that $b_2' \to 0$. $\square$

5. Concluding remarks. In this short section, we comment on some results obtained in this paper and list an open problem.

REMARK 5.1. The one-dimensional setting of Theorem 5 originally arose from studying $GI/G/1$ queue in [13] and was later applied to the CUSUM method and the BLAST program. It would be interesting to know any possible applications of Theorems 5 and 6 to queuing theory.

REMARK 5.2. Let $\{X_i; i \ge 1\}$ be a sequence of i.i.d. random variables and $S_k = \sum_{i=1}^k X_i$. Let $f : \mathbb{N} \to \mathbb{N}$ be an increasing function. We studied in Theorems 1 and 2 the asymptotic behavior of $\max_{k\ge1} S_{f(k)}$ when $f(x) = x^p$, $p = 2, 3, \dots,$ and the case $f(x) = x$ is treated in Lemma A.1. It would be interesting to see what happens for general $f(x)$, particularly the case when $f(x)/x \to \infty$ but the fluctuation theory still works.

REMARK 5.3. We impose the condition $Ee^{tX} < \infty$ for any $t \in \mathbb{R}$ in almost every one of our theorems. It would be interesting to see what happens in the case that the moment generating function does not exist, for example, X is a α-stable variable $(0 < \alpha < 1)$.

REMARK 5.4. One of basic assumptions in this paper is that random variables $\{X_I\}$ are i.i.d. It would be interesting to know what happens if random variables $\{X_I\}$ are not independent, but instead are related in a Gaussian manner. Also, X_I and X_J become asymptotically independent as the distance between I and J goes to infinity.

REMARK 5.5. We dealt with the maximum indexed by squares and rectangles in Theorems 5 and 7. One should not have much difficulty in handling general convex sets by understanding the local behavior as in Theorems 1 and 3 and then using the Chen–Stein method globally.

REMARK 5.6. Theorem 7 is a result for $d = 2$. It is interesting to ask what happens when $d \geq 3$. The key is to address the following number theoretic question: set $q_d(k) = \#\{(p_1, \dots, p_d) \in \mathbb{N}^d;\ p_1, \dots, p_d = k\}$. What is the asymptotic behavior of $\sum_{k=1}^{n} q_d(k)$? If there exist two constants c_1 and c_2 such that

$$\sum_{k=1}^{n} q_d(k) = c_1 n(\log n)^{d-1} + c_2 n(\log n)^{d-2}\big(1 + o(1)\big)$$

as $n \to \infty$, then the following result can be proved: under condition (1.1), there exist a constant $K > 0$, an integer $m > 0$, and coefficients $a_1, a_2, \dots, a_m$ such that

$$P\left(U_n \leq \sum_{k=1}^{m} a_k \log_k(n) + x\right) \to e^{-Ke^{-\theta x}} \qquad \forall x \in \mathbb{R},$$

as $n \to \infty$, where $\log_k(n) := \log(\log(\cdots(\log n)))$ with k iterated natural logs.

OPEN PROBLEM. Suppose condition (1.1) holds. Do there exist a constant $K' > 0$ and a sequence of numbers $\{a_n; n \geq 1\}$ such that $0 \leq a_n = o(\log n)$ and

$$\lim_{n\to\infty} P\left(\tilde{U}_n \geq \frac{\log n}{\theta} + a_n + x\right) = 1 - e^{-K'e^{-\theta x}}$$

for all $x \in \mathbb{R}$?

Acknowledgments. This is a part of the author's dissertation in the Department of Statistics, Stanford University. The author would like to thank his advisor Amir Dembo for his stimulating supervision. The author also thanks David Siegmund for his suggestion to study the problems in Section 3, and Charles Stein for pointing out a flaw in Proposition 2.1 in an earlier draft of the author's dissertation. Finally, the author thanks an anonymous referee for his/her suggestions on revising the paper.

REFERENCES

[1] ARRATIA, R., GOLDSTEIN, L. and GORDON, L. (1989). Two moments suffice for Poisson approximation: The Chen–Stein method. *Ann. Probab.* **17** 9–25.

[2] BAHADUR, R. R. and RANGA RAO, R. (1960). On deviations of sample mean. *Ann. Math. Statist.* **31** 1015–1027.

[3] CHAGANTY, N. R. and SETHURAMAN, J. (1993). Strong large deviations and local limit theorems. *Ann. Probab.* **21** 1671–1690.

[4] CHOW, Y. S. and TEICHER, H. (1988). *Probability Theory, Independence, Interchangeability, Martingales*, 2nd ed. Springer, New York.

[5] DEHEUVELS, P. and DEVROYE, L. (1987). Limit law of Erdös–Rényi–Shepp type. *Ann. Probab.* **15** 1363–1386.

[6] DEMBO, A., KARLIN, S. and ZEITOUNI, O. (1994). Critical Phenomena for sequence matching with scoring. *Ann. Probab.* **22** 1993–2021.

[7] DEMBO, A., KARLIN, S. and ZEITOUNI, O. (1994). Limit distribution of maximal non-aligned two-sequences segmental score. *Ann. Probab.* **22** 2022–2039.
[8] DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. Springer, New York.
[9] DUDLEY, R. M. (1978). Central limit theorems for empirical measures. *Ann. Probab.* **6** 899–929.
[10] ERDÖS. P. and RÉYNI, A. (1970). On a new law of large numbers. *J. Anal. Math.* **22** 103–111.
[11] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, 2nd ed. Wiley, New York.
[12] HARDY, G. H. and WRIGHT, E. M. (1988). *An Introduction to the Theory of Numbers*, 5th ed. Oxford Univ. Press.
[13] IGLEHART, D. (1972). Extreme values in the GI/G/1 queue. *Ann. Math. Statist.* **43** 627–635.
[14] JIANG, T. (2000). A comparison of scores of two protein structures with foldings. Preprint.
[15] KARLIN, S. and ZHU, Z. Y. (1996). Clusters of charged residues in protein structures. *Proc. Natl. Acad. Sci. USA* **93** 8350–8355.
[16] POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.
[17] ROOTZEN, H. (1988). Maxima and exceedances of stationary Markov chains. *Adv. in Appl. Probab.* **20** 371–390.
[18] ŠALI, A., SHAKHNOVICH, E. and KARPLUS, M. (1994). How does a protein fold? *Nature* **369** 248–251.
[19] SIEGMUND, D. (1988). Approximate tail probabilities for the maxima of some random fields. *Ann. Probab.* **16** 487–501.
[20] SIEGMUND, D. and VENKATRAMAN, E. S. (1995). Using the generalized likelihood ratio statistic for sequential detection of a change point. *Ann. Probab.* **23** 255–271.
[21] SIEGMUND, D. and YAKIR, B. (2000). Tail probabilities for the null distribution of scanning statistics. *Bernoulli* **6** 191–213.
[22] SPITZER, F. (1964). *Principles of Random Walk*. Van Nostrand, Princeton.
[23] VAPNIK, V. N. and ČERVONENKIS, A. JA. (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab. Appl.* **16** 264–280.

SCHOOL OF STATISTICS
UNIVERSITY OF MINNESOTA
313 FORD HALL
224 CHURCH STREET S.E.
MINNEAPOLIS, MN 55455
E-MAIL: tjiang@stat.umn.edu