Maxima of Partial Sums Indexed by Geometrical Structures

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF STATISTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Tiefeng Jiang

August 1999

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Amir Dembo
(Principal Adviser)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Tze Leung Lai

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Charles Stein

Approved for the University Committee on Graduate Studies:

Dedicated to God

The LORD is my shepherd, I shall not be in want. He makes me lie down in green pastures, he leads me beside quiet waters, he restores my soul. He guides me in paths of righteousness for his name's sake. Even though I walk through the valley of the shadow of death, I will fear no evil, for you are with me; your rod and your staff, they comfort me. You prepare a table before me in the presence of my enemies. You anoint my head with oil; my cup overflows. Surely goodness and love will follow me all the days of my life, and I will dwell in the house of the LORD forever. — Psalm 23

耶和华是我的牧者。我必不至缺乏。
他使我躺卧在青草地上，领我在可安歇的水边。
他使我的灵魂苏醒，为自己的名引导我走义路。
我虽然行过死阴的幽谷，也不怕遭害。因为你与我同
在。你的杖，你的竿，都安慰我。在我敌人面前，你为我
摆设筵席。你用油膏了我的头，使我的福杯满溢。我一
生一世必有恩惠慈爱随着我。我且要住在耶和华
的殿中，直到永远。 --诗篇23篇

Abstract

Motivated by DNA and protein matching problems, the author studies maxima of partial sums indexed by binary trees, squares and rectangles over lattice points and random cubes in the three dimensional unit cube. Results on comparisons of two protein structures and two collections of data sitting at lattice points are also obtained. For some of these problems, the dimension ($d = 1, d = 2$ and $d \geq 3$) significantly affects the limit behavior of the maxima (because of the so called "curse of dimensionality"). However, for some other problems, the maxima behave almost the same as their one dimensional counterparts. The tools of proving these results are refined large deviations, the Chen-Stein method, number theory and inequalities of empirical processes.

Acknowledgement

The author thanks his advisor Amir Dembo for his excellent supervision during his four years study in Stanford University and personal help. By random chance, the author got in touch with him and became his student before long. His sharp intuition and quick and deep thoughts enabled the author to finish his dissertation much more easily. The author is greatly indebted to Persi Diaconis and Ingram Olkin for their excellent classes as well as their inspired coaching of his academic career. Many thanks to his teachers in and outside classes: Jerry Friedman, Samuel Karlin, Tze Leung Lai, Joe Romano and David Siegmund.

Many thanks to the spiritual support of Chinese Agape Vineyard Christian Fellowship, particularly pastor Timothy Su. With their help, the author has enjoyed the best quality of life in any sense so far in the Bay Area.

The author would like to thank his classmates Tongwei Liu, Sandra McBride and Julia Tung. Sandra and Julia constantly help improve the author's English and learn about America's culture. Gifted Tongwei often helps the author check his computations and proofs. Of course, the author has to thank some Chinese restaurants in the Bay area which provided delicious food for their meetings.

Finally, the author thanks his wife Li Sun for her patience and care. Without that, he would not get his Ph.D. degree so smoothly. Thanks to baby Ruth Jiang, who changed the author from a man to a busier but much happier father.

Contents

Chapter 1

Introduction

1.1 Motivation

In the 80's to 90's, Arratia and Waterman in [5], Arratia, Morris and Waterman in [4], Karlin and Ost in [20], Arratia, Gordon and Waterman in [2] and [3], Karlin and Altschul in [19], Karlin in [18], Dembo, Karlin and Zeitouni in [10] and [11] have studied bimolecular comparison problems. Based on these works, the BLAST program was established and has been used widely in the industry of bioinformatics. To see the motivation of our current work clearly, we next review a piece of bimolecular background.

A protein is a polymer with a linear single chain called backbone composed of peptide bonds. Amino acids are the building blocks of protein. An amino acid has three functional ends: an amino end, a carboxyl end, and a side chain. There are 20 different amino acids and they differ only in their side chain composition in either charge, hydrophobic or chemical properties. The amino backbone of one amino acid links to the carboxyl backbone end of another amino acid to form the peptide bond that is the backbone of a protein chain. This single chain protein folds into a stable complex three dimensional structure in solution. Although certain secondary structures (alpha helixes and beta sheets) can be predicted from the primary linear sequence, the overall three dimensional structure is still beyond the ability of even the best structural prediction algorithms available today.

Suppose the letters of amino acids in a primary linear sequence are $X_1, X_2, \cdots, X_n$. The positions of these amino acids in the three dimensional space are $U_1, U_2, \cdots, U_n$. We usually call $\{U_i; i = 1, 2 \cdots, n\}$ the folding of the protein chain. The letter-position pairs of another protein chain are $\{(Y_i, V_i); i = 1, 2 \cdots, m\}$. Biologists want to know if there are some local similarities between the two chains, namely, if there exist two neighborhoods B_U and B_V such that the alphabets of those amino acids with positions U_i in B_U and V_i in B_V are similar. For example, they may be completely the same or partially matched.

All previous work mentioned so far essentially use the following paradigm. For a given real score function $F(x, y)$, which primarily has negative mean and an essential positive part, they constructed a statistic by taking the maximum of all $\sum_{l=1}^{\Delta} F(X_{i+l}, Y_{j+l})$ over all possible i and j running in sequences X and Y, respectively. The asymptotic distribution of the statistic is derived. In their work, however, one important feature is lost, which is the foldings of protein structures. Some obvious problems immediately arise. For instance, for two amino acid chains which are very close(far) to each other in space, their corresponding letters appearing in the sequence $X_1, X_2, \cdots, X_n$, could be far(close) from each other. So the statistic portrayed above is not accurate in this sense, but people have not taken this into account because the folding is very complicated.

In our present work, a statistic involved with folding is studied. In addition, we work on the distribution of clusters of charged residues in protein structures. Moreover, the method used in the analysis of the above two problems enables us to study change point problems.

1.2 Problems we solved

The first problem we deal with is a change point problem. We acquired results about the maxima of partial sums indexed by squares and rectangles, which enable us to know that there is a change point area. Siegmund and Yakir in [26] studied this problem recently by using likelihood ratio test. We briefly review our change point problem as follows. 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

Figure 1.1

One immediately notices that there is some zone where data are significantly different from those in the other parts(The above data are actually sampled from a $N(-2,1)$ distribution, and 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 an exaggerated typical setting in change point problems. The goal is to detect whether there is a zone from which data are different from others. See more details in Sections 2.5 and 2.4.

We will use the same method of pursuing limiting distribution of statistics in the above problem to study the maxima of partial sums indexed by binary trees proposed by Karlin. The problem comes as follows. Because the folding of protein structure is complicated, we consider a simple folding, i.e., all the amino acids in a protein structure are sitting on the roots of binary trees:

$$X_{1,1}$$

$$X_{2,1} \qquad X_{2,2}$$

$$X_{3,1},\ X_{3,2},\ X_{3,3},\ X_{3,4},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$X_{n,1}\ X_{n,2}, \cdots, \qquad X_{n,2^{n-1}-1},\ X_{n,2^{n-1}}$$

Figure 1.2

Biologists are interested in finding the biggest cluster of charged residues emerging in the tree. In terms of the binary tree in Figure 1.2, the problem is studying the maxima of partial sums indexed by subtrees.

Motivated by Karlin and Zhu [21], which studied clusters of charged residues in protein structures, we constructed in section 3.2 a model as follows: Let $\{X_i;\ i = 1, 2, \cdots\}$ and $\{U_i;\ i = 1, 2, \cdots\}$ be two sequences of i.i.d. random variables. X_1 is bounded with negative mean and has an essential positive part. Our statistics is maximum of $\sum_{i=1}^{n} X_i I\{U_i \in$ subcube$\}$ over a certain class of sub-cubes of $[0, 1]^3$.

For comparison problems, the first one we dealt with in Chapter 4 is the comparison of two collections of data sitting on lattice points. Statistically, the null hypothesis is that the two collections of data are independent. We construct a statistic and get its asymptotic distribution for the testing.

The last problem, which is also our initial motivation for this work, is the comparison of two protein structures stated in Section 1.1. We analyze the procedure(or statistics) proposed by Karlin. The procedure is briefly described as follows: We use the notation $\{X_i, U_i, i = 1, 2 \cdots,\}$ and $\{Y_i, V_i, i = 1, 2 \cdots,\}$ in Section 1.1 to represent the two protein structures with foldings. From the structure $\{X_i, U_i, i = 1, 2 \cdots,\}$, pick any U_{i_0}, and list U_i in an order such that the distance of U_i and U_{i_0} is increasing. Call the random variables $\{X_i\}$ ordered in this fashion $X_{1,1}, X_{1,2}, \cdots, X_{1,n}$. Then repeat this procedure to $X_{1,2}, \cdots$, etc. Also, apply this procedure to $\{Y_i\}$. We then finally obtain two collections of new data generated from X and Y, respectively:

$$
\begin{array}{ll}
X_{1,1},\ X_{1,2}, \cdots,\ X_{1,n} & \qquad Y_{1,1},\ Y_{1,2}, \cdots,\ Y_{1,n} \\
X_{2,1},\ X_{2,2}, \cdots,\ X_{2,n} & \qquad Y_{2,1},\ Y_{2,2}, \cdots,\ Y_{2,n} \\
\cdots\cdots\cdots & \qquad \cdots\cdots\cdots \\
X_{n,1},\ X_{n,2}, \cdots,\ X_{n,n} & \qquad Y_{n,1},\ Y_{n,2}, \cdots,\ Y_{n,n}
\end{array}
$$

Figure 1.3

By comparing each row of X-matrix with that of the Y-matrix, we end up with the statistic

$$\max_{1\leq i,j,m\leq n}\sum_{p=1}^{m}F(X_{i,p},Y_{j,p}).$$

The asymptotic distribution of the above statistics is given in Chapter 5. In the real protein structure, the amino acids are not uniformly distributed in some cube. However, the proof of our main result indicates that it does not depend on the specific geometry of a cube but will apply any regular geometric shape on which $\{U_i;\ i=1,2,\cdots\}$ are uniformly or close to being uniformly distributed. We did not make a statistic on such non-cubic geometry but rather used the above indication. The reason is obvious because we have to go to more rigorous mathematics once we make that kind of statistic. It may not be worth doing that.

1.3 Methodology of proofs

We found that the so-called curse of dimensionality phenomenon appeared critically in the analysis of our statistics. It is well known that the phenomenon generally occurs in two or higher dimensional spaces. Now we use a simple example to illustrate this.

Let X_1 be a bounded random variable with negative mean and an essential positive part. $S_k=\sum_{i=1}^{k}X_i$ are the partial sums. One part of our analysis is involved in investigating the asymptotic behavior of $\max_{k\geq 1}S_{k^d}, d\geq 1$. It is well known that when X_1 is non-lattice, one has

$$P(\max_{k\geq 1}S_k\geq z)\sim Ke^{-\theta z}\quad \text{as } z\to+\infty \tag{1.1}$$

for some positive constants K and θ (see Lemma 2.2.4). (1.1) is obtained by using renewal and fluctuation theories. In our problems, we are concerned with the case $d\geq 2$. We obtain that

$$P(\max_{k\geq 1}S_{k^2}\geq z)\sim f(z)e^{-\theta z}\quad \text{as } z\to+\infty, \tag{1.2}$$

where $f(z)$ is an explicit function with order roughly of $z^{-1/2}$ as $z\to+\infty$ (see Theorem

2.4). For the case $d \geq 3$, see Theorem 2.3. (1.2) is proved based on the observation that the distribution of the index set in (1.2) is so sparse compared to that in (1.1) that it allows us to examine (1.2) directly by using inclusion-exclusion formula and refined large deviations of S_k. The renewal and fluctuation theories used in the one-dimensional case are completely irrelevant to statistics on change point problems for the dimension $d \geq 2$. The key observation is as follows: The worst case of dependency of two different blocks of r.v.'s $X_i, X_{i+1}, \cdots, X_{i+\Delta}$ and $X_j, X_{j+1}, \cdots, X_{j+\Delta}$ is that they are almost a common sequence except two new variables either X_i and $X_{j+\Delta}$ or X_j and $X_{i+\Delta}$. However, in two dimensional cases, the worst dependency of two different squares of r.v.'s with size Δ^2 is that there are $2\Delta^2$ new variables. Such a difference in the two cases is reflected strongly in Lemma 2.2.8, which allows us to pursue the Chen-Stein Poisson approximation method. On the other hand, in the "continuous case," e.g. in charged residue model, we do use some results obtained from renewal and fluctuation theories for the local area, and the Chen-Stein method globally with the help of the curse of dimensionality phenomenon.

We use number theory to estimate certain parameters. For example, in Lemma 2.4.2, we estimate the order of the following quantity precisely:

$$\sum_{|k-y| \leq \sqrt{y} \log y} q(k) e^{-\frac{(k-y)^2}{y}} \sim \alpha \sqrt{y} \log y \quad \text{as } y \to +\infty$$

for some constant $\alpha > 0$, where $q(k) = \#\{(r,s) \in \mathbb{N}^2; rs = k\}$. In Section 2.4, an inequality on empirical processes is obtained, which seems to be of interest in itself.

Chapter 2

Maxima on Lattice Structures

2.1 Notations

Throughout this thesis, we use the following notations.

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

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

$[a]$: The biggest integer no larger than x.

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 $n \to \infty$.

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

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

2.2 Some auxiliary lemmas.

In this section, we accumulate some tools which will be frequently used in this and later chapters. Some of those tools are quoted directly from literatures, while the others need proofs.

The following inequality provides us with bounds for tails of sums of independent and bounded random variables, see Exercise 14 in [8] or page 193 in [23].

Lemma 2.2.1 *(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*

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

Let A be a finite set. Throughout, $|A|$ denotes the cardinality of A.

Lemma 2.2.2 *Let* $\{X_\alpha; \alpha \in \Delta\}$ *be a collection of i.i.d random variables, where* Δ *is a finite set.* $A_i, B_i,\ i = 1, 2$ *are subsets of* Δ *satisfying* $|A| = |A_1|,\ |B| = |B_1|$ *and* $|A \cap B| \le |A_1 \cap B_1|$. $\Phi(x)$ *is a measurable function on* R^1. *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 prove this lemma according to the two cases.
(i) Suppose $|A| = |B|$. Take a subset $D \subset A_1 \cap B_1$ for which $|D| = |A_1 \cap B_1| - |A \cap B|$. In what follows, we use notation E^A for expectations with respect to $\{X_\alpha; \alpha \in A\}$. Then from the invariance property of the joint distribution of $\{X_\alpha; \alpha \in \Delta\}$, it follows that

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

where the only "$\le$" appear in the above argument is the Cauchy-Schwartz inequality.
(ii) Suppose that Φ is monotone. Likewise, take a subset $D \subset A_1 \cap B_1$ so that $|D| = |A_1 \cap B_1| - |A \cap B|$. It follows that

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

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

The following Poission approximation theorem is a straightforwad 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 2.2.3 *Let Ω be a finite set and $\mathcal{A}$ is a collection of some subsets of Ω. $\{X_\alpha,\ \alpha \in \Omega\}$ is a collection of random variables. Denote $S_A = \sum_{\alpha \in A} X_\alpha$ and $\lambda = \sum_{A \in \mathcal{A}} P(S_A > t)$ for some $t \in R^1$. Then*

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

where

$$b_1 = \sum_{A \in \mathcal{A}} \sum_{B: B \cap A \ne \emptyset} P(S_A > t) P(S_B > t), \quad b_2 = \sum_{A \in \mathcal{A}} \sum_{B: B \cap A \ne \emptyset} P(S_A > t,\ S_B > t),$$
$$b_3 = \sum_{A \in \mathcal{A}} E|P(S_A > t | \sigma\{S_B; B \cap A = \emptyset\}) - P(S_A > t)|,$$

where $\sigma\{S_B; B \cap A = \emptyset\}$ is the σ-algebra generated by the collection of random variables $\{S_B; B \cap A = \emptyset\}$. Particularly, 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 $\mathcal{B}_A$, the neighborhood of A be the set of all B so that $B \cap A \ne \emptyset$. Then the result follows from Theorem 1 in [1]. ∎

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

$$E(Z) < 0,\ P(Z > 0) > 0 \text{ and } E\exp(tZ) < \infty \quad \forall t \in R^1. \tag{2.1}$$

Under condition (2.1), there is an unique constant $\theta > 0$ so that

$$E\exp(\theta X) = 1. \tag{2.2}$$

The following lemma was first proved probably by Spitzer(E4 in page 217 from [27]). See also (5.13) from [15].

Lemma 2.2.4 *If X is non-lattice random variable satisfying (2.1), then*

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

When X is a lattice variable of span δ, then

$$K = \lim_{n\to\infty} e^{n\theta} P(\max_{k\geq 1} S_k > n\delta) = \frac{\delta C}{e^{\theta\delta} - 1}.$$

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 expression of C above is due to i.i.d. fluctuation sum identities(Chapter 12, [15]).

Lemma 2.2.4 is important for i.i.d. partial sums. Iglehart([17]) 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 For more information on oscillation phenomena for partial sums of i.i.d. variables see Deheuvels and Devroye([9]), Rootzen([24]), Siegmund([25]) and the references therein.

Given a random variable X, the corresponding log moment generating function and its conjugate which is also called rate function are

$$\Lambda_X(t) = \log E\exp(tX), \qquad \Lambda_X^*(x) := \sup_{t\in R^1}\{tx - \Lambda_X(t)\}.$$

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.

The next lemma whose proof may be found in Remark (a) below Theorem 3.7.4 in [12], collects useful properties of $\Lambda(t)$ and $\Lambda^*(x)$.

Lemma 2.2.5 *Suppose $E\exp(tX) < \infty$ for all $t \in \mathbb{R}^1$ and X is non-degenerate. Then*
(i) $\Lambda''(t) > 0$ for all $t \in R$,
(ii) $\Lambda^*(x)$ is infinitely differentiable on the interior of the convex hull of the support of X.

(iii) If condition (2.1) holds, then $\Lambda^*(\Lambda'(\theta)) = \theta\Lambda'(\theta)$, $(\Lambda^*)'(\Lambda'(\theta)) = \theta$ *and* $(\Lambda^*)''(\Lambda'(\theta)) = 1/\Lambda''(\theta)$, *where* θ *is as in (2.2).*

Lemma 2.2.6 *Suppose condition (2.1) holds. Denote* $\mathcal{D}_{\Lambda^*} = \{x \in R; \Lambda^*(x) < \infty\}$. *Then* $I(x) := \Lambda^*(x)/x$ *is strictly decreasing on* $(0, \Lambda'(\theta)]$ *and strictly increasing on* $[\Lambda'(\theta), +\infty) \cap \mathcal{D}_{\Lambda^*}$.

Proof. Obviously, the condition (2.1) implies that $[0, +\infty) \subset \{\Lambda(t);\ t \in R^1\}$. Moreover, $\Lambda(t)$ is a strictly convex function. Thus, for any $x_2 > x_1 \geq \Lambda'(\theta)$ such that $\Lambda^*(x_i) < \infty$ for $i = 1, 2$, 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,

$$\frac{\Lambda^*(x_1)}{x_1} = t_1 - \frac{\Lambda(t_1)}{x_1} < t_1 - \frac{\Lambda(t_1)}{x_2} \leq \frac{\Lambda^*(x_2)}{x_2}. \tag{2.3}$$

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 (2.3), we have $\Lambda^*(x_1)/x_1 > \Lambda^*(x_2)/x_2$. ∎

Throughout the rest of this section, we assume $\{X, X_n; n \geq 1\}$ is a sequence of i.i.d. random variables with mean μ. $S_n = \sum_{i=1}^n X_i$ are partial sums. We always assume that X is non-degenerate. The following proposition, which is slightly stronger than the usual Bahadur-Rao theorem (see [6]), provides us uniform estimates of tail probabilities. It is a pivotal tool in our proofs.

Proposition 2.1 *Suppose* X *is non-lattice and* $\Lambda(t) < \infty$ *for all* $t \in \mathbb{R}^1$. *Then*

$$\sup_{a \leq \eta \leq b} |C_n(\eta) P(S_n \geq n\Lambda'(\eta)) - 1| \to 0 \quad \text{as } n \to \infty$$

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 non-lattice, the random variable $Z := e^{tX}/Ee^{tX}$ is also non-lattice. 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

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

for any constants $\delta_i > 0, i = 1, 2$. It follows from Theorem 3.3 and Remark 3.6 in [7] 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. ■

An immediate consequence of the above proposition is as follows, which is served as a tool for accurate estimates.

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

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

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 of Corollary 2.1. For any $|x| \le \delta\sqrt{n \log n}$, $a \le \eta \le b$, 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 $\mu < a_1 < a < b < b_1 < B$ and $\{\eta_{n,x};\ |x| \le \delta\sqrt{n \log n}\} \subset (a_1, b_1)$ for n large enough. Therefore

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

By Proposition 2.1, it remains to show that

$$\frac{C_n(x, \eta)}{C_n(\eta_{n,x})} \to 1 \tag{2.4}$$

uniformly in x and η. By Taylor expansion and Lemma 2.2.5,

$$\frac{C_n(x, \eta)}{C_n(\eta_{n,x})} = \frac{\eta\sqrt{\Lambda''(\eta)}}{\eta_{n,x}\sqrt{\Lambda''(\eta_{n,x})}} \exp\left[-\frac{x^3}{6n^2} (\Lambda^*)^{(3)}(\xi_\eta) \right] \tag{2.5}$$

where ξ_η is between $\Lambda'(\eta)$ and $\Lambda'(\eta)+n^{-1}x$. Obviously, we have that $|(x^3/6n^2)(\Lambda^*)^{(3)}(\xi_\eta)| \leq n^{-1/2}(\log n)^{3/2}\sup_{a_1\leq x\leq b_1}|\Lambda^{*(3)}(x)|$. 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 $\delta\sqrt{\log n/n} \geq |\Lambda'(\eta)-\Lambda'(\eta_{n,x})| \geq h|\eta-\eta_{n,x}|$. Apply the Mean-value theorem again to the function $x\sqrt{\Lambda''(x)}$, we obtain from (2.5)

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

Therefore, (2.4) is true. ■

The following easy fact is called Chernoff bound(see e.g. p.p. 31 in [12]). It is weaker than Proposition 2.1, but it is a simple and non-asymptotic bound.

Lemma 2.2.7 *(Chernoff bound) For any $x > EX$,*

$$P(S_n/n \geq x) \leq \exp(-n\Lambda^*(x)), \quad \forall n\geq 1.$$

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

Lemma 2.2.8 *Suppose A, B, and C are disjoint sets of indices. $\{X, X_\alpha; \alpha\in A\cup B\cup C\}$ are i.i.d. random variables with X satisfying condition (2.1) with $\mu = EX$. For any subset $D\subset A\cup B\cup C$, we use the notation $S_D := \sum_{\alpha\in D}X_\alpha$. Then,*

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

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 that, without loss of generality, $|A|\geq|C|$. Then

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

By Chernoff bound, $P(S_A\geq x|A|) \leq \exp(-|A|\Lambda^*(x))$. Also, by using that $E\exp(\theta X) = 1$ and Markov inequality, we have $P(S_{B\cup C}\geq z)\leq e^{-\theta z}$ and $P(S_B\geq z - x|A|) \leq \exp(-\theta z -$

$\theta|x||A|)$. Therefore,

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

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

Lemma 2.2.9 *Assume X satisfies condition (2.1). Let $d\geq 1$ be a constant and $S_n=\sum_{i=1}^n X_i$. Then there exist constants $r>1$ and $t_0>\theta$ such that*

$$\sum_{k\geq rz^{1/d}} P(S_{k^d}\geq z)+\sum_{k\leq r^{-1}z^{1/d}} P(S_{k^d}\geq z)=o(e^{-t_0 z})\quad as\ z\to\infty.$$

Proof. Let $r_1=(3\Lambda'(\theta)/2)^{1/d}$. Recall the definition of $I(x)$ in Lemma 2.2.6, it follows that $\lambda =: \inf_{x\geq 3\Lambda'(\theta)/2} I(x)>\theta$. Then by the Chernoff bound and Lemma 2.2.6,

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

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

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

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

Lemma 2.2.10 *Suppose condition (2.1) 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*

$$z^{\frac{1}{2}-\frac{1}{d}}e^{\theta z}\sum_{k\in\Gamma_z} P(S_{k^d}\geq z)=O\left(e^{-c(z)^2 z^{1-2/d}}\right)\quad z\to\infty,$$

where $\Gamma_z=\{k\in\mathbb{N};\ sz^{1/d}\leq k\leq c_0z^{1/d}-b(z)\ or\ c_0z^{1/d}+a(z)\leq k\leq rz^{1/d}\}$ and $c(z)=a(z)\wedge b(z),\ z>0$.

Proof. Let $\Gamma'_z = \{k \in \mathbb{N}; c_0 z^{1/d} + a(z) \le k \le r z^{1/d}\}$. Then, By Proposition 2.1 and Lemma 2.2.6

$$\sum_{k\in\Gamma'_z} P(S_{k^d} \ge z) \le C \sum_{k\in\Gamma'_z} \frac{1}{\sqrt{k^d}} \exp\{-k^d \Lambda^*(\frac{z}{k^d})\} \le C z^{-\frac{1}{2}+\frac{1}{d}} \exp\{-h(z)\Lambda^*(\frac{z}{h(z)})\},$$

where $h(z) = (c_0 z^{1/d} + a(z))^d$. Here constants C dependent on $\Lambda(\cdot)$ and d may vary from line to line. Denote $\Delta = z/h(z) - \Lambda'(\theta)$. By Taylor expansion $\Delta = -dc_0^{-1} a(z) z^{-1/d} + O(a(z)^2 z^{-2/d})$. By Taylor expansion again and Lemma 2.2.5

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

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

$$z^{\frac{1}{2}-\frac{1}{d}} e^{\theta z} \sum_{k\in\Gamma'_z} P(S_{k^d} \ge z) = O\left(e^{-a(z)^2 z^{1-2/d}}\right).$$

By the same arguments, the above estimate is also true if Γ'_z and $a(z)$ are replaced by $\Gamma_z \backslash \Gamma'_z$ and $b(z)$, respectively. ■

2.3 Maxima on squares.

As usual, $\mathbb{N}$ is the set of all positive integers. $\mathbb{N}^d$ is the d-fold Cartesian product, namely, $\mathbb{N}^d = \{I = (i_1, i_2, \cdots, i_d);\ i_k \in \mathbb{N}, k = 1, 2, \cdots, d\}$. So capital letters such as I, J, L, etc. are used to denote points in $\mathbb{N}^d$. The notation $(i_1, i_2, \cdots, i_d) = I \le J = (j_1, j_2, \cdots, j_d)$ means that $i_l \le j_l$ for all $l = 1, 2, \cdots, d$, and $I < J$ when all inequalities are strict. Also, as convention, $I + J = (i_1 + j_1, i_2 + j_2, \cdots, i_d + j_d)$ and $m = (m, m, \cdots, m) \in \mathbb{N}^d$. Let $\Delta = \Delta(I, J) = \{L \in \mathbb{N}^d;\ I \le L \le J\}$, $\mathcal{R}_n = \{\Delta = \Delta(I, J);\ 1 \le I \le J \le n\}$. Assuming that $\{X_I;\ I \in \mathbb{N}^d\}$ are i.i.d. non-lattice random variables, we denote $S_\Delta = \sum_{L\in\Delta} X_L$ and

$$U_n = \max_{\Delta\in\mathcal{R}_n} S_\Delta.$$

For the cube case, we denote $\mathcal{H}_n^d$ as the set of all the sub-cubes of $\{1, 2, \cdots, n\}^d$. More precisely, $\mathcal{H}_n^d = \{\Delta = \Delta(I, I+k) \subset \mathbb{N}^d;\ 1 \le I \le I+k \le n\}$. Also,

$$W_n = \max_{\Delta \in \mathcal{H}_n^d} S_\Delta.$$

For $\Delta = \Delta(I, J)$, where $I = (i_1, i_2, \cdots, i_d)$ and $J = (j_1, j_2, \cdots, j_d)$, we use $|\Delta| = (j_1 - i_1 + 1) \cdots (j_d - i_d + 1)$ to denote the number of points in Δ. Hereafter, we state main results first and then prove them later on. The following are strong laws for the two statistics W_n and U_n.

Theorem 2.1 *Suppose condition (2.1) holds, then for any $d \in \mathbb{N}$,*

$$\text{(i)}\ \lim_{n\to\infty} \frac{W_n}{\log n} \to \frac{d}{\theta}\ a.s. \quad \text{and} \quad \text{(ii)}\ \lim_{n\to\infty} \frac{U_n}{\log n} \to \frac{d}{\theta}\ a.s.$$

Although we give the strong law of maxima indexed by rectangles, we will not analyze this object in detail here. The strong laws of both the square and rectangle cases are proved at same time tightly. The asymptotic law of U_n will be given in the next section. To prove the following Theorems 2.4 and 2.5, we first study two one-dimensional problems on S_{k^d}.

Theorem 2.2 *Suppose condition (2.1) 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*

$$\lim_{z\to\infty} \frac{\sqrt{z}e^{\theta z}}{\delta(\gamma(z))} P(\max_{k\ge1} S_{k^2} \ge z) = \frac{1}{\theta}\sqrt{\frac{\Lambda'(\theta)}{2\pi\Lambda''(\theta)}}.$$

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

Theorem 2.3 *Suppose condition (2.1) 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 $d \ge 3$*

$$\lim_{z\to\infty} \sqrt{z}(G_n(z) + G_{n+1}(z))^{-1} P(\max_{k\ge1} S_{k^d} \ge z) = \frac{1}{\theta}\sqrt{\frac{\Lambda'(\theta)}{2\pi\Lambda''(\theta)}}\ .$$

With the help of above Theorems 2.2 and 2.3, we are now able to analyze the asymptotic behavior of W_n. The results are as follows.

Theorem 2.4 *Suppose that $d = 2$ and condition (2.1) holds. Let $t_n = \log\delta(\gamma(2\log n/\theta))$, where functions $\delta(\cdot)$ and $\gamma(\cdot)$ are as in Theorem 2.2. Define $\log_2 n = \log(\log n)$. Then*

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

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

Theorem 2.5 *Suppose $d \geq 3$ and condition (2.1) holds. Let $k_n = \inf\{k \in \mathcal{N}; (\log k)/2 + \alpha k^d \geq \log n\}$, where $\alpha = \theta\Lambda'(\theta)/d$, and $r_n = \exp\left\{d\log n - d(\log k_n)/2 - k_n^d\theta\Lambda'(\theta)\right\}$. Then*

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

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

It is easy to see that r_n of Theorem 2.5 does not converge. So also $P(W_n \leq \Lambda'(\theta)k_n^d + x)$ does not converge, but Theorem 2.5 gives the first degree of approximation for the latter probability.

We turn to proving the above theorems. To prove Theorem 2.1, we need the following two lemmas.

Lemma 2.3.1 *Define $q_d(k) = \#\{(i_1, \cdots i_d) \in \mathbb{N}^d;\ i_1 i_2 \cdots i_d = k\}$. Then*

$$\sum_{k=1}^{m} q_d(k) \leq m\left(\log(em)\right)^{d-1}.$$

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

Proof. 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

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

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

$$\sum_{i=1}^{m} q_2(i) \leq m \sum_{i=1}^{m} \frac{1}{i} \leq m \log(em) \tag{2.6}$$

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

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

Now we prove the lemma by induction. Suppose the lemma is true for $d = l \geq 2$. Then it is easy to check that

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

where the last inequality is from (2.6). The proof is done. ∎

Lemma 2.3.2 *Suppose that condition (2.1) holds. For any given $\epsilon > 0$, there exists $c > 0$ such that $c < \Lambda'(\theta) < 1/c$ and*

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

as $z \to +\infty$.

Proof. First, by Chebyshev's inequality and Lemmas 2.2.6 and 2.3.1, for any $c < \Lambda'(\theta)$, we have

$$\sum_{k \leq cz} q_d(k) P(S_k \geq z) \leq \left(\max_{k \leq cz} e^{-k\Lambda^*(z/k)}\right) \sum_{k \leq cz} q_d(k) \leq cz(4 \log cz)^{d-1} e^{-I(1/c)z}$$
$$= O(e^{-\epsilon z})$$

for sufficiently large $c > 0$. By Lemma 2.3.1, we have that $q_d(k) \le k(4\log k)^{d-1} \le e^{\Lambda^*(0)k/2}$ for all k large enough. It follows that for given $c > 0$

$$\sum_{k\ge z/c} q_d(k)P(S_k \ge z) \le \sum_{k\ge z/c} q_d(k)\exp(-k\Lambda^*(0)) \le e^{-\Lambda^*(0)z/(2c)}/(1-e^{-\Lambda^*(0)/2}).$$

Then choose c appropriately the result follows. ■

Proof of Theorem 2.1. To prove (i) and (ii) at same time, it suffices to prove the following two inequalities:

$$\limsup_{n\to\infty} \frac{U_n}{\log n} \le \frac{d}{\theta} \quad a.s. \qquad \liminf_{n\to\infty} \frac{W_n}{\log n} \ge \frac{d}{\theta} \quad a.s. \tag{2.7}$$

We first prove the $\limsup$ in (2.7)

Given $\eta > 0$, set $z_n = (1+\eta)d(\log n)/\theta$. Choose ϵ in Lemma 2.3.2 big enough so that $n^d e^{-\epsilon z_n} \le n^{-2}$. Note that the number of rectangles with the same upper-left corner and area k is at most $q_d(k)$. By Lemma 2.3.2, there exists $c > 0$ so that

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

where we use $E\exp(\theta X) = 1$ in the second inequality and Lemma 2.3.1 in the only equality above. Define $l_n = [n^{2/\eta d}]$, then the above inequality implies that

$$P\left(U_{l_n} \ge \frac{(1+\eta)d\log l_n}{\theta}\right) = O((\log n)^2/n^2). \tag{2.8}$$

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 on n, and $l_{n+1}/l_n \to 1$, then limsup in (2.7) follows.

Now, we turn to prove $\liminf$ in (2.7). Denote $k_n = [(c_1 d\log n)^{1/d}]$ 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 of $S_{k_n^d}$. We break the cube $\{1, 2, \cdots, n\}^d$ into m_n many disjoint sub-cubes. Then, the partial sums of X_i's over the disjoint sub-cubes are i.i.d. Therefore, for any given $\eta \in (0,1)$,

$$P\left(W_n \leq \frac{(1-\eta)d\log n}{\theta}\right) \leq P\left(\max_{1\leq i\leq m_n} Y_i \leq \frac{(1-\eta)d\log n}{\theta}\right) \leq \exp(-m_n P(S_{k_n^d} \geq t_n)),$$

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,

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

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

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

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

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

for any $\eta > 0$ small enough. Then the liminf in (2.7) is proved. ■

Proof of Theorem 2.2. 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 2.2.9 and 2.2.10, we know that

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

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

$$\sqrt{z}e^{\theta z}\delta(\gamma(z))^{-1} \sum_{k\in A_z} P(S_{k^2} \geq z) \sim \frac{1}{\theta}\sqrt{\frac{\Lambda'(\theta)}{2\pi\Lambda''(\theta)}}, \tag{2.9}$$

$$\sqrt{z}e^{\theta z} P(\max_{k\in A_z} S_{k^2} \geq z) \sim \sqrt{z}e^{\theta z} \sum_{k\in A_z} P(S_{k^2} \geq z). \tag{2.10}$$

By Corollary 2.1

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

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

$$\begin{aligned}&\sqrt{z}e^{\theta z} \sum_{k \in A_z} P(S_{k^2} \geq z) \\ &\sim \frac{1}{\theta}\sqrt{\frac{\Lambda'(\theta)}{2\pi\Lambda''(\theta)}} \sum_{k\in A_z} \exp\left\{-\left(\frac{\Lambda'(\theta)}{2\Lambda''(\theta)}\right)\frac{(z-\Lambda'(\theta)k^2)^2}{z}\right\} + o\left(\frac{a(z)^2}{z}\right). \qquad (2.11)\end{aligned}$$

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

$$\frac{(z-\Lambda'(\theta)k^2)^2}{z} = \Lambda'(\theta)^2 \frac{(k+\gamma(z))^2}{z}(k-\gamma(z))^2 = 4\Lambda'(\theta)(k-\gamma(z))^2 + O((\log z)^{3/2}/z)$$

uniformly for all $k \in A_z$. Therefore

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

Obviously, $\sum_{|k|\geq a(z)} e^{-\beta k^2} \leq 2\sum_{k\geq a(z)} e^{-\beta a(z)k} = O(z^{-\beta\eta^2})$, which together with the above equality and (2.11) yield (2.9).

Now, we turn to prove (2.10).

For any $(i,j) \in \Delta_z := \{(i,j) \in \mathbb{N}^2 : c_0\sqrt{z} - a(z) \leq i < j \leq c_0\sqrt{z} + a(z)\}$, 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 $z \to \infty$. Obviously,

$$P(S_{i^2} \geq z, S_{j^2} \geq z) \leq P(S_{i^2} \geq z + C_{i,j}) + P(S_{i^2} \geq z)P(S_{j^2-i^2} \geq -C_{i,j}).$$

Since $E\exp(\theta X) = 1$, by Markov inequality, $P(S_{i^2} \geq z + C_{i,j}) \leq \exp\{-\theta z - \theta C_{i,j}\}$. By Chernoff bound, $P(S_{j^2-i^2} \geq -C_{i,j}) \leq \exp\{-(j^2-i^2)\Lambda^*(\mu/2)\}$. Thus, there exists a constant

$C > 0$, so that

$$P(S_{i^2} \geq z, S_{j^2} \geq z) \leq Cz^{-1/2} \exp\{-\theta z - C^{-1}\sqrt{\log z}\} \tag{2.12}$$

uniformly for all $i, j \in \Delta_z$. Therefore, by union-intersection formula,

$$\begin{aligned}\sqrt{z}e^{\theta z}\left(\sum_{k\in A_z} P(S_{k^2} \geq z) - P(\max_{k\in A_z} S_{k^2} \geq z)\right) &\leq \sqrt{z}e^{\theta z} \sum_{(i,j)\in\Delta_z} P(S_{i^2} \geq z, S_{j^2} \geq z) \\ &= O(\exp(-C^{-1}\sqrt{\log z})).\end{aligned}$$

Then, (2.10) follows. ■

Proof of Theorem 2.3. Take $a(z) = b(z) = 1/3$, note that $\{k \in \mathbb{N}; |k - (z/\Lambda'(\theta))^{1/d}| \leq 1/3\} \subset \{n, n+1\}$. Then by Lemmas 2.2.9 and 2.2.10, we obtain

$$|P(\max_{k\geq 1} S_{k^d} \geq z) - P(\max\{S_{n^d}, S_{(n+1)^d}\} \geq z)| = O(e^{-\theta z - 9^{-1}z^{1-2/d}}). \tag{2.13}$$

By the same arguments as obtaining (2.12), we have that $P(S_{n^d} \geq z, S_{(n+1)^d} \geq z) = O(\exp(-\theta z - Cz^{1-1/d}))$ for some $C > 0$. Therefore

$$P(S_{n^d} \geq z) + P(S_{(n+1)^d} \geq z) - P(\max\{S_{n^d}, S_{(n+1)^d}\} \geq z) = O(e^{-\theta z - Cz^{1-1/d}}).$$

Also, by Proposition 2.1, $P(S_{m^d} \geq z) \sim \theta^{-1}\sqrt{\Lambda'(\theta)/2\pi\Lambda''(\theta)}\ z^{-1/2}G_m(z)$, $m = n, n+1$. Moreover, $G_m(z) \leq \theta z$ for $m = n, n+1$ by Lemma 2.2.6. It follows that

$$\sqrt{z}(G_n(z) + G_{n+1}(z))^{-1}P(\max_{k\geq 1} S_{k^d} \geq z) = \frac{1}{\theta}\sqrt{\frac{\Lambda'(\theta)}{2\pi\Lambda''(\theta)}} + O(\sqrt{z}e^{-Cz^{1-2/d}}). \quad ■$$

Proof of Theorem 2.4 Let $a_n = 2h(\log_2 n)^{1/2}, h > 1$ and $f_0 = (\theta\Lambda'(\theta)/2)^{-1/2}$. Denote E_n as the set of all the subsquares in $\{1, 2, \cdots, n\}^2$ with side length between $f_0\sqrt{\log n} - a_n$ and $f_0\sqrt{\log n} + a_n$. Precisely, $E_n = \{\Delta \in \mathcal{S}_n; |\sqrt{|\Delta|} - f_0\sqrt{\log n}| \leq a_n\}$. Define

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

Through all this thesis, when we do computation on $\overline{W}_n$ or its counterparts, we always view it as two maxima. The first maximum is that of S_Δ over all sub-cubes Δ with fixed up-left corner, then $\overline{W}_n$ is the maximum of the former ones over all n^2 corners. Based on this observation, by Lemmas 2.2.9 and 2.2.10

$$P(W_n > z_n) - P(\overline{W}_n > z_n) = e^{-\theta x} O((\log n)^{-3}). \tag{2.14}$$

Now we use Chen-Stein method Lemma 2.2.3 to get the asymptotic distribution. First, we need a lemma as follows.

Lemma 2.3.3 $\delta((t + O(\log t))^{1/2})/\delta(t^{1/2}) \to 1$ *as* $t \to \infty$.

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

Let us continue the proof of Theorem 2.4. Let $\Omega_n = \{k \in \mathbb{N};\ |k - f_0\sqrt{\log n}| < a_n\}$, by (2.9) and Lemma 2.3.3, we have that

$$\lambda := \sum_{\Delta\in E_n} P(S_\Delta \geq z_n) = \sum_{k\in\Omega_n} (n-k+1)^2 P(S_{k^2} \geq z_n) = n^2 \sum_{k\in\Omega_n} P(S_{k^2} \geq z_n) + o(1)$$
$$\sim n^2 z_n^{-1/2} e^{-\theta z_n} \delta(\gamma(z_n))(1/\theta)\sqrt{\Lambda'(\theta)/2\pi\Lambda''(\theta)} \sim K_1 e^{-\theta x}, \tag{2.15}$$

where $K_1 = (1/2)\sqrt{\Lambda'(\theta)/\pi\theta\Lambda''(\theta)}$. By Lemma 2.2.3, to complete the proof, we just need to show b_1 and b_2 go to zero. Actually, for any particular $\Delta \in E_n$, the number of other squares which intersect Δ is at most $4|\Delta|a_n$. Moreover, $P(S_{k^2} \geq z_n) = e^{-\theta x}O(n^{-2}\sqrt{\log n})$ by using the fact $E\exp(\theta X) = 1$. Consequently,

$$b_1 \leq \underbrace{n^2 \cdot 4|\Delta|a_n \cdot |\Omega_n|^2}_{O(n^2(\log n)^2)} \max_{(i,j)\in(\Omega_n)^2} P(S_{i^2} \geq z_n)P(S_{j^2} \geq z_n) = O(n^{-2}(\log n)^3).$$

Likewise,

$$b_2 \leq O(n^2(\log n)^2) \max_{\Delta_1\neq\Delta_2, \Delta_1,\Delta_2\in E_n} P(S_{\Delta_1} \geq z_n, S_{\Delta_2} \geq z_n).$$

For any two $\Delta_1, \Delta_2 \in E_n$, $\Delta_1 \neq \Delta_2$, the symmetric difference of them, i.e., $(\Delta_1 \backslash \Delta_2) \cup (\Delta_2 \backslash \Delta_1)$ is at least $f_0\sqrt{\log n} - a_n$. This is the key observation in handling such kind of high dimensional problems in this thesis. By Lemma 2.2.8, $P(S_{\Delta_1} \geq z_n, S_{\Delta_2} \geq z_n) \leq n^{-2}\exp(-(f_0/2)\sqrt{\log n})$ for all n large enough. Thus, $b_2 = O((\log n)^2 \exp(-(f_0/2)\sqrt{\log n}))$.
■

Proof of Theorem 2.5 Obviously, $k_n = [((\log n)/\alpha)^{1/d}]$ or $[((\log n)/\alpha)^{1/d}] + 1$. Define E_n as the set of all subcubes in $\{1, 2, \cdots, n\}^d$ with side length k_n, i.e., $E_n = \{\Delta \in \mathcal{S}_n; |\Delta|^{1/d} = k_n\}$. Also,

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

Apply $a(z) = b(z) = 1/4$, $z = z_n := \Lambda'(\theta)k_n^3 + x$ to Lemma 2.2.10, and then by Lemma 2.2.10,

$$P(W_n \geq z_n) - P(\overline{W}_n \geq z_n) = O(n^{-\epsilon})$$

for some $\epsilon > 0$. Hence, to prove our theorem, it is enough to show that

$$P\left(\overline{W}_n \leq \Lambda'(\theta)k_n^d + x\right) - \exp\left(-K_2 r_n e^{-\theta x}\right) \to 0. \tag{2.16}$$

Note that

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

But, by Corollary 2.1,

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

Now we prove the theorem by Proposition 2.2.3, i.e., the Chen-Stein method. It is easy to check

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

For any two different sub-cubes with side lengths k_n for which they are overlapped, the symmetric difference of them is at least $2k_n^{d-1}$. By Lemma 2.2.8,

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

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

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

Therefore, (2.16) follows. ■

2.4 Maxima on rectangles.

Like the notations in Section 2.2, $\{X, X_{i,j};\ i \geq 1, j \geq 1\}$ are i.i.d. random variables, and $\mathcal{R}_n$ is the set of all the rectangles in $\{1, 2, \cdots, n\}^2$, i.e., $\mathcal{R}_n = \{\Delta = \Delta(I, J);\ 1 \leq I \leq J \leq n\}$. Let $S_{p,q} = \sum_{i=1}^{p} \sum_{j=1}^{q} X_{i,j}$. In this section, we will study the following two statistics:

$$U = \max_{p\geq 1, q\geq 1} S_{p,q}, \quad \text{and} \quad U_n = \max_{\Delta\in\mathcal{R}_n} S_\Delta.$$

The first theorem below is about the local maximum of sums indexed by rectangles with fixed up-left corner, which is the first-hand information for studying maxima over moving rectangles.

Theorem 2.6 *Suppose condition (2.1) holds, then*

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

Compare Theorem 2.4 in section 2.3, which is on the maxima over squares, we now have limit law of U_n.

Theorem 2.7 *Suppose condition (2.1) holds, for any* $x \in R^1$,

$$P\left(U_n \le \frac{2\log n}{\theta} + \frac{\log_3 n}{\theta} + x\right) \to e^{-Ke^{-\theta x}},$$

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

From Theorem 2.1, we know that both U_n and W_n have the same scale. But obviously, $U_n \ge W_n$. Theorem 2.4 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 $U_n \sim (2\log n + \log_3 n)/\theta$. This comparison makes sense.

Now we begin to prove the two theorems. Denote $E_z = \{(p,q) \in \mathbb{N}^2;\ |pq - \Lambda'(\theta)^{-1}z| < \alpha\sqrt{z\log z}\ \}$, and

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

Lemma 2.4.1 *Suppose condition (2.1) holds. Then*

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

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 $S_k = \sum_{i=1}^n X_{1,i}$. Compare the quantity $\sum_{n\notin\Omega_z}\sqrt{n}P(S_n \ge z)$ with Lemmas 2.2.9 and 2.2.10($d = 1$), we see that the only difference between them is that $\sqrt{n}$ appears in the form term. But this term does not dominate the sum. So by checking the proofs of Lemmas 2.2.9 and 2.2.10($d = 1$), 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 $C > 0$. ■

Lemma 2.4.2 *Recall* $E_z = \{(p,q) \in \mathbb{N}^2;\ |pq - \Lambda'(\theta)^{-1}z| < \alpha\sqrt{z\log z}\ \}$. *Then for any* $\alpha > 0$

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

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

$$P(S_{p,q} > z) \sim \frac{e^{-\theta z}}{\theta\sqrt{2\pi\Lambda''(\theta)pq}} \exp\left(-\frac{\Lambda'(\theta)^2}{2pq\Lambda''(\theta)}\left(\frac{z}{\Lambda'(\theta)} - pq\right)^2\right) \tag{2.17}$$

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

$$\begin{aligned}&\theta\sqrt{2\pi\Lambda''(\theta)/\Lambda'(\theta)}e^{\theta z}(\log z)^{-1}\sum_{(p,q)\in E_z} P(S_{p,q}\geq z)\\ &\sim \frac{1}{\sqrt{z}\log z}\sum_{(p,q)\in E_z}\exp\left(-K_1\frac{(pq-y)^2}{y}\right) = \frac{1}{\sqrt{z}\log z}\sum_{k\in\Omega_z} q(k)\exp\left(-K_1\frac{(k-y)^2}{y}\right),\end{aligned}$$

where $K_1 = \Lambda'(\theta)^2/2\Lambda''(\theta)$ and $\Omega_z = \{k \in \mathbb{N};\ |k - y| \leq \alpha\sqrt{z\log z}\}$. To finish the proof, we need to show that

$$\frac{1}{\sqrt{z}\log z}\sum_{k\in\Omega_z} q(k)\exp\left(-K_1\frac{(k-y)^2}{y}\right) \to \sqrt{\pi/K_1\Lambda'(\theta)}. \tag{2.18}$$

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

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

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

$$\sum_{k\in\Omega_z} q(k)\exp\left(-K_1\frac{(k-y)^2}{y}\right) = \sum_{i=-i_z}^{i_z}\sum_{k\in A_i} q(k)\exp\left(-K_1\frac{(k-y)^2}{y}\right) + O(\sqrt{z}), \tag{2.19}$$

Now we estimate $\sum_{k\in A_i}$ of (2.19). Note that for any $k \in A_i$,

$$e^{-K_1(i\Delta)^2/y} - e^{-K_1(k-y)^2/y} = e^{-K_1(i\Delta)^2/y}(1 - e^{\phi_k(y)}),$$

where $\phi_k(y) := K_1(k-y+i\Delta)(k-y-i\Delta)/y$. It is easy to check that $\rho_i := \max_{k\in A_i}|\phi_k(y)| \leq K_1(2|i|+1)\Delta^2/y \leq C\gamma$, where C here and through all this proof is a constant which depends on X and α only, and may vary from line to line. Therefore, since $|e^x - 1| \leq |x|e^{|x|}$ for any $x \in R$, we have that

$$\sum_{k\in A_i} q(k)\exp\left(-K_1\frac{(k-y)^2}{y}\right) = (1+\rho_i)e^{-K_1(i\Delta)^2/y}\sum_{k\in A_i} q(k). \tag{2.20}$$

By Theorem 320 on p. 264 of [16]

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

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

$$\sum_{i=m+1}^{n} q(k) = (n-m)\log n + (c+1)(n-m) + O(\sqrt{n}). \tag{2.21}$$

As a consequence,

$$\sum_{k\in A_i} q(k) = \sum_{[y+i\Delta]+1}^{[y+(i+1)\Delta]} q(k) = \gamma\sqrt{y\log y} + O(\sqrt{y}). \tag{2.22}$$

By (2.19), (2.20), and (2.22), we obtain

$$\frac{1}{\sqrt{z}\log z}\sum_{k\in\Omega_z} q(k)\exp\left(-K_1\frac{(k-y)^2}{y}\right) \sim \frac{\gamma}{\sqrt{\Lambda'(\theta)\log y}}\sum_{-i_z}^{i_z}(1+\rho_i')e^{-K_1(i\gamma)^2/\log y}, \tag{2.23}$$

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

$$\frac{\gamma}{\sqrt{\log y}}\sum_{-i_z}^{i_z} e^{-K_1(i\gamma)^2/\log y} \to \int_{-\infty}^{\infty} e^{-K_1t^2}\,dt = \sqrt{\frac{\pi}{K_1}}, \quad \forall\gamma \in (0,\infty). \tag{2.24}$$

It follows that

$$\limsup_{y\to+\infty}\frac{\gamma}{\sqrt{\log y}}\sum_{-i_z}^{i_z-1}\rho_i' e^{-K_1(i\gamma)^2/\log y}\le C\gamma. \tag{2.25}$$

Thus, combining (2.23), (2.24) and (2.25), we obtain that

$$\limsup_{y\to+\infty}\left|\frac{1}{\sqrt{z}\log z}\sum_{k\in\Omega_z}q(k)e^{-K_1(i\gamma)^2/\log y}-\sqrt{\frac{\pi}{K_1\Lambda'(\theta)}}\right|\le C\gamma$$

for arbitrary given $\gamma>0$. Let $\gamma\downarrow 0$, then (2.18) follows. ■

Lemma 2.4.3 *Given $\alpha>0$ and any function p_z, let*

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

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 relying on θ. Set $B_z=\{q\in\mathbb{N};\ \beta_z^-\le pq\le\beta_z^+\}$. Then

$$\sum_{(p,q)\in E_z\setminus E_z^5}P(S_{p,q}\ge z)\le 2\sum_{\substack{1\le p\le p_z,\\(p,q)\in E_z}}P(S_{p,q}\ge z)\le 2C_\theta z^{-1/2}e^{-\theta z}\sum_{p=1}^{p_z}\sum_{q\in B_z}1$$
$$\le C_\theta z^{-1/2}e^{-\theta z}(\beta_z^+-\beta_z^-)\sum_{p=1}^{p_z}\frac{1}{p}=O(\sqrt{\log z}e^{-\theta z}\log p_z).\quad ■$$

Proof of Theorem 2.6. Combining Lemmas 2.4.1, 2.4.2 and 2.4.3 we have that for any $\epsilon>0$ there is $\alpha>0$ so that

$$\limsup_{z\to\infty}e^{\theta z}(\log z)^{-1}\left(P(U\ge z)-P(U_z^5\ge z)\right)\le\epsilon, \tag{2.26}$$

where $U_z^5 = \max_{(p,q)\in E_z^5} S_{p,q}$ and E_z^5 is as in Lemma 2.4.3 with $p_z = \exp(\epsilon\sqrt{\log z})$. We claim that

$$\lim_{z\to\infty} e^{\theta z}(\log z)^{-1}\left(P(U_z^5 \geq z) - \sum_{(p,q)\in E_z^5} P(S_{p,q} \geq z)\right) \to 0, \qquad \forall \alpha > 0. \tag{2.27}$$

If the claim is true, then by Lemmas 2.4.2, 2.4.3 and (2.26), then

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

for any $\epsilon > 0$. Then the Theorem follows by letting $\epsilon \downarrow 0$. Now we prove the claim. Observe that

$$P(\max_{\Gamma\in E_z^5} S_\Gamma \geq z) \geq \sum_{\Gamma\in E_z^5} P(S_\Gamma \geq z) - \sum_{\Gamma_1\neq\Gamma_2\in E_z^5} P(S_{\Gamma_1} \geq z, S_{\Gamma_2} \geq z).$$

To prove claim (2.27), it suffices to show that

$$e^{\theta z}(\log z)^{-1} \sum_{\Gamma_1\neq\Gamma_2\in E_z^5} P(S_{\Gamma_1} \geq z, S_{\Gamma_2} \geq z) \to 0 \tag{2.28}$$

as $z \to +\infty$. By definition of E_z^5, $|\Gamma_1\Delta\Gamma_2| \geq \exp(\epsilon\sqrt{\log z})$ for any $\Gamma_1,\ \Gamma_2 \in E_z^5$. Thus, by Lemma 2.2.8, there is a $\zeta > 0$, so that

$$P(S_{\Gamma_1} \geq z, S_{\Gamma_2} \geq z) \leq 2\exp(-\theta z - \zeta e^{\epsilon\sqrt{\log z}}), \tag{2.29}$$

as z is large enough. Therefore

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

where $E_z = \{(p,q) \in \mathbb{N}^2;\ |pq - \Lambda'(\theta)^{-1}z| < \alpha\sqrt{z\log z}\ \}$ is as before. By (2.21), we have that

$$|E_z| \le \sum_{i=\beta_z^-}^{\beta_z^+} q(i) = O(\sqrt{z\log z}),$$

where β_z^- and β_z^+ are those as in the proof of lemma 2.4.3. It follows that

$$\sum_{\Gamma_1 \neq \Gamma_2 \in E_z^5} P(S_{\Gamma_1} \ge z, S_{\Gamma_2} \ge z) = O\left(z(\log z)\exp(-\theta z - \delta e^{\epsilon\sqrt{\log z}})\right),$$

which implies (2.28). ■

Proof of Theorem 2.7. Denote $z_n = (2\log n + \log_3 n)/\theta + x$. Take $p_{z_n} = e^{(\log_2 z_n)^{1/4}}$ in the the definition of E_z^5 in Lemma 2.4.3. Of course $E_{z_n}^5$ is the corresponding subsequence of E_z^5. Let $\mathcal{R}_n^1$ be the set of all the rectangles in $\{1,2,\cdots,n\}^2$ whose length and width, say, p,q, satisfying $(p,q) \in E_{z_n}^5$. Accordingly, $W_n^1 := \max_{\Delta\in\mathcal{R}_n^1} S_\Delta$. By Lemmas 2.4.1, 2.4.2 and 2.4.3, there is $\alpha > 0$ such that

$$e^{\theta z}(\log z)^{-1} \sum_{(p,q)\in\mathbb{N}^2\setminus E_z^5} P(S_{p,q} \ge z) = O((\log z)^{-1/4}) \tag{2.30}$$

for large z, where E_z^5 is that in Lemma 2.4.3 corresponding to $p_z = \exp((\log_2 z)^{1/4})$. Still, like we did before, we view $\mathcal{R}_n$ as the union of rectangles with fixed up-left corners for all such possible corners. It follows by (2.30) that

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

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

$$P(W_n^1 \ge z_n) \to 1 - e^{-Ke^{-\theta x}}. \tag{2.31}$$

First, it is easy to see that

$$(n-L_n)^2 \sum_{(p,q)\in E_{z_n}} P(S_{p,q} \geq z_n) \leq \sum_{\Delta\in\mathcal{R}_n^1} P(S_\Delta \geq z_n) \leq n^2 \sum_{(p,q)\in E_{z_n}} P(S_{p,q} \geq z_n), \tag{2.32}$$

where L_n is the longest length and width of any rectangles in $E^5_{z_n}$. Obviously, $L_n \leq \Lambda'(\theta)z_n - \alpha\sqrt{z_n \log z_n} = O(\log n)$. By Lemmas (2.4.2) and (2.4.3),

$$n^2 \sum_{(p,q)\in E_{z_n}} P(S_{p,q} \geq z_n) \sim \frac{n^2 e^{-\theta z_n} \log z_n}{\theta\sqrt{\Lambda'(\theta)}} \sim \frac{e^{-\theta x}}{\theta\sqrt{\Lambda'(\theta)}}. \tag{2.33}$$

Thus, (2.32) and (2.33) imply that $\lambda_n := \sum_{\Delta\in\mathcal{R}_n^1} P(S_\Delta \geq z_n) \to e^{-\theta x}/\theta\sqrt{\Lambda'(\theta)}$ as $n \to \infty$. Now we use Chen-Stein method to end the proof.

For any $\Delta \in \mathcal{R}_n^1$, define $\mathcal{A}_\Delta = \{\Delta' \in \mathcal{R}_n^1;\ \Delta \cap \Delta' \neq \emptyset\}$. It is easy to count that $|\mathcal{A}_\Delta| = O((\log n)\log_2 n)$. By lemma 2.2.3, to prove (2.31), we need to verify that b_1 and b_2 go to zero. Recall $P(S_\Delta \geq z_n) \sim e^{-\theta z_n}\log z_n \sim n^{-2}$ uniformly for all $\Delta \in \mathcal{R}_n^1$. Then

$$b_1 \leq |\mathcal{A}_\Delta| \max_{\Delta'\in\mathcal{A}_\Delta} P\left(S_{\Delta'} \geq z_n\right)\lambda_n = O\left(\frac{(\log n)\log_2 n}{n^2}\right).$$

By (2.29)

$$P\left(S_\Delta \geq z_n,\ S_{\Delta'} \geq z_n\right) = O\left(n^{-2}(\log_2 n)^{-1}\exp(-\zeta e^{(\log_2 n)^{1/4}})\right).$$

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

$$b_2 \leq |\mathcal{R}_n^1||\mathcal{A}_\Delta| \max_{\Delta,\Delta'\in\mathcal{R}_n^1} P\left(S_\Delta \geq z_n,\ S_{\Delta'} \geq z_n\right) = O\left((\log n)^2(\log_2 n)^2\exp(-\zeta e^{(\log_2 n)^{1/4}})\right). \blacksquare$$

2.5 Maxima on binary trees.

We first introduce the definition of binary trees, then analyze maxima of partial sums indexed by subtrees. We call $(1,1)$ the origin, this origin has immediate two children: $(2,1)$ and $(2,2)$. Then each of the above two children has immediate two children: $(3,1),(3,2)$, and $(3,3),(3,4)$, respectively. And each of the four children has exact two children, $\cdots$. Precisely, we define Δ_n as the tree starting from (1,1) and having n generations:

$$\Delta_n = \{(k,l); l = 1, \cdots, 2^{k-1}, k = 1, 2, \cdots, n\}.$$

A subtree $\Delta = \Delta_{i,j}^m$ is a tree starting from the origin at (i,j) and has m generations in the tree Δ_n, i.e.,

$$\Delta_{i,j}^m = \{(k,l); l = (j-1)2^{k-i} + 1, \cdots, j2^{k-i}, k = i, i+1, \cdots, i+m\}$$

Let $\{X, X_{i,j};\ (i,j) \in \Delta_n\}$ be a family of i.i.d. real-valued random variables. In this section, we concern on the following statistics

$$W_n = \max_{(i,j,m):\Delta_{i,j}^m \subset \Delta_n} S_{\Delta_{i,j}^m} ,$$

As before, $S_\Delta = \sum_{\alpha \in \Delta} X_\alpha$. Our main theorem in this section is as follows.

Theorem 2.8 *Suppose condition (2.1) holds for X. Let $k_n = \inf\{i \in \mathbb{N};\ 2^i + \alpha i \geq \beta n\}$, where $\alpha = 3(\log 2)/2\theta\Lambda'(\theta)$ and $\beta = (\log 2)/\theta\Lambda'(\theta)$. Let $r_n = 4\exp\{(\log 2)n - 1.5(\log 2)k_n - 2^{k_n}\theta\Lambda'(\theta)\}$. Then*

$$P\left(W_n \leq \Lambda'(\theta)2^{k_n} + x\right) - e^{-pr_n e^{-\theta x}} \to 0,$$

where $p = 4(\theta\sqrt{2\pi\Lambda''(\theta)})^{-1}$.

Proof. Let $z_n = \Lambda'(\theta)2^{k_n} + x$. First note that k_n is either $[\log_2(\beta n)]$ or $[\log_2(\beta n)] + 1$, therefore

$$k_n \sim \log_2 n \quad \text{and} \quad \theta\Lambda'(\theta)2^{k_n} + 1.5(\log 2)k_n \geq (\log 2)n. \tag{2.34}$$

Let $\mathcal{A}_i$ be the set of all subtrees in Δ_n which has exactly i generations, $i = 1, 2, \cdots, n$. Then $|\mathcal{A}_i| = 2^{n+2-i}$. Set

$$W_n^1 = \max_{\Delta \in \mathcal{A}_{k_n-1}} S_\Delta.$$

We now show that W_n is approximated by W_n^1. Actually,

$$|P(W_n > z_n) - P(W_n^1 > z_n)| \le \sum_{\substack{1 \le i \le n \\ i \ne k_n}} 2^{n+2-i} P(S_{2^i-1} \ge z_n). \tag{2.35}$$

By Chernoff bound,

$$P(S_{2^i-1} \ge z_n) \le \exp\{-z_n I(z_n/(2^i-1))\} \tag{2.36}$$

where $I(x) = \Lambda^*(x)/x$. It is easy to check that

$$\inf_{1 \le i \le k_n-1} \left\{ \frac{z_n}{2^i-1} \right\} \ge 2\Lambda'(\theta)(1 + 2^{-k_n}x); \qquad \sup_{k_n+1 \le i \le n} \left\{ \frac{z_n}{2^i-1} \right\} \le \frac{4}{7}\Lambda'(\theta)(1 + 2^{-k_n}x).$$

for all $k_n \ge 3$. Therefore, by Lemma 2.2.6, there exists $\delta > 0$ such that

$$\inf_{1 \le i \le n, i \ne k_n-1} I(z_n/(2^i-1)) \ge \theta + \delta.$$

It follows from (2.34), (2.35) and (2.36) that

$$|P(W_n \ge z_n) - P(W_n^1 \ge z_n)| \le n 2^n e^{-(\theta+\delta)z_n} = O(ne^{-(\log 2)\delta n/4}).$$

So it is enough to show that

$$P\left(W_n^1 \le \Lambda'(\theta)2^{k_n} + x\right) - e^{-pr_n e^{-\theta x}} \to 0.$$

We now prove this by appealing the Chen-Stein method. Let $\{\Delta_i;\ i = 1, 2, \cdots, 2^{n+2-k_n}\}$ be a list of $\{\Delta_{i,j}^{k_n-1} \subset \Delta_n$ for all possible i and $j\}$. For any $i = 1, 2, \cdots, 2^{n+2-k_n}$, let $\mathcal{B}_i =$

$\{j;\ \Delta_j \cap \Delta_i \neq \phi,\ i = 1, 2, \cdots, 2^{n+2-k_n}\}$. By Corollary 2.1

$$\lambda_n := \sum_{i=1}^{2^{n+2-k_n}} P(S_{\Delta_i} > z_n) = 2^{n+2-k_n} P(S_{\Delta_1} \geq z_n)$$
$$\sim \frac{4 \cdot 2^{n-1.5k_n}}{\theta\sqrt{2\pi\Lambda''(\theta)}} e^{-2^{k_n}\theta\Lambda'(\theta)-\theta x} = \frac{4r_n e^{-\theta x}}{\theta\sqrt{2\pi\Lambda''(\theta)}}.$$

By Lemma 2.2.3, we only need to prove the following two terms go to zero:

$$b_{1,n} = \sum_{i=1}^{2^{n+2-k_n}} \sum_{j\in\mathcal{B}_i} P(S_{\Delta_j} > z_n) P(S_{\Delta_i} > z_n); \quad b_{2,n} = \sum_{i=1}^{2^{n+2-k_n}} \sum_{i\neq j\in\mathcal{B}_i} P(S_{\Delta_i} > z_n, S_{\Delta_j} > z_n).$$

First, notice that $|\mathcal{B}_i| \leq 2^{k_n}$, then $b_{1,n} = \lambda_n^2 \cdot 2^{2k_n-n-2} \to 0$. Second,

$$b_{2,n} \leq 2^{n+2} \max_{i\neq j\in\mathcal{A}_i} P(S_{\Delta_j} > z_n, S_{\Delta_i} > z_n).$$

Observe that the difference between Δ_i and Δ_j is at least 2^{k_n-1} for any i, j. Then by Lemma 2.2.8, there is $\zeta > 0$ so that

$$P(S_{\Delta_j} > z_n, S_{\Delta_i} > z_n) \leq \exp(-\theta z_n - 2^{k_n}\zeta).$$

Therefore, by (2.34)

$$b_{2,n} \leq 2^{n+2}\exp(-\theta z_n - 2^{k_n}\zeta) = O(a^{-n}),$$

where $a = 2^{\zeta/3}$. ■

Chapter 3

Maxima on Random Cubes

3.1 An inequality on empirical processes.

We will prove an inequality which will be used in the next section. First, we review some basic definitions and facts of empirical processes. Given a class of $\mathcal{F}$ of subsets of S, define $\Delta^{\mathcal{F}}(s_1, \cdots, s_n) = \#\{F \cap \{s_1, \cdots, s_n\};\ F \in \mathcal{F}\}$ for any $\{s_1, \cdots, s_n\} \subset S$, and

$$m^{\mathcal{F}}(n) = \max\{\Delta^{\mathcal{F}}(s_1, \cdots, s_n),\ s_i \in S\}, \quad V(\mathcal{F}) = \inf\{n : m^{\mathcal{F}}(n) < 2^n\},$$
$$V(\mathcal{F}) = +\infty \text{ if } m^{\mathcal{F}}(n) = 2^n \text{ for all } n.$$

Dudley([14]) calls $\mathcal{F}$ a Vapnik-Červonenkis(VC for short) class of sets if $V(\mathcal{F}) < \infty$. $V(\mathcal{F})$ is called the exponent of the VC class $\mathcal{F}$. One remarkable result of VC class given by Vapnik-Červonenkis is the following VC lemma from [28]:

$$m^{\mathcal{F}}(n) \leq n^{V(\mathcal{F})}, \quad \forall n \geq 2. \tag{3.1}$$

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 parallel to those of $[0,1]^3$. Specifically,

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

Lemma 3.1.1 *$\mathcal{F}$ is a VC class with exponent no greater than* 6*. Therefore, for any* $(y_1, \cdots, y_n) \in \mathbb{R}^n$ *and* $n \geq 2$, $\#\{(y_1, y_2, \cdots, y_n) \cap F;\ F \in \mathcal{F}\} \leq n^6$.

Proof. Take any six points $x_i = (x_i^1, x_i^2, x_i^3) \in (0,1)^3, i = 1, \cdots, 6$. Define $x_-^j = \min\{x_i^j;\ 1 \leq i \leq 6\}$ and $x_+^j = \max\{x_i^j;\ 1 \leq i \leq 6\},\ j = 1,2,3$. Denote by B the rectangular solid whose six faces parallel to x-y, y-z and z-x axes respectively and pass through $x_+^j, x_-^j,\ j = 1,2,3$. Then B contains all the six points $x_1, \cdots, x_6$. To show $\mathcal{F}$ is a VC class, we need to show that $m^{\mathcal{F}}(6) < 2^6$. We will prove this by considering two mutually exclusive cases.

Case (i). If a point, say, x_1, is in the interior of B, then any cube which covers $\{x_2, \cdots, x_6\}$ must cover x_1 too. Therefore, no cube separates $\{x_1\}$ from the other five points $\{x_i, i = 2, \cdots, 6\}$.

Case (ii). If case (i) is not true, then all the six points lie on the six faces. So, suppose there are two points in one face, say, x_1, x_2, then any cube covers $\{x_2, \cdots, x_6\}$ also covers x_1. Consequently, no cube separate $\{x_1\}$ from $\{x_2, \cdots, x_6\}$. Then the only left possibility is that each face contains exactly one point. Let l be the shortest of B's eight edge lengths. Without loss of generality, assume x_1, x_2, x_3, x_4 are on the faces of B in which two edges of the face have length l. Then any cube which covers $\{x_1, x_2, x_3, x_4\}$ also covers at least one of the two points x_5 and x_6(otherwise, it contradicts the definition of l). Thus, no cube separates $\{x_1, x_2, x_3, x_4\}$ from $\{x_5, x_6\}$.

In summary, $m^{\mathcal{F}}(6) < 2^6$. Therefore, $V(\mathcal{F}) \leq 6$. By (3.1), $m^{\mathcal{F}(n)} \leq n^6$ for $n \geq 2$. ∎

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

$$\mathcal{F}_{1,n} = \{B(x, \epsilon) \in \mathcal{F},\ x \in [0,1]^3,\ \epsilon \leq C(\log n/n)^{1/3}\}, \quad C > 0.$$

Lemma 3.1.2 *For any $C > 0$, there is a constant $D > 0$ such that*

$$P(\#\Gamma_{\mathcal{F}_{n,1}} \geq Dn(\log n)^6) = O(n^{-3}).$$

Proof. For any class of subsets $\mathcal{C}$, $\Gamma_{\mathcal{C}} := \#\{\{Y_1, \cdots, Y_n\} \cap F;\ F \in \mathcal{C}\}$. Let $r_n = (\log n/n)^{1/3}$ and $\mathcal{G}_i = \{B(Y_i, r_n) \cap F; F \in \mathcal{F}\}$ for $i = 1, 2, \cdots$. Since $\mathcal{F}$ is a VC class with exponent no greater than 6, so is $\mathcal{G}_i$. By (3.1),

$$\#\Gamma_{\mathcal{G}_1} \le \left\{ \sum_{i=1}^{n} I_{B(Y_1, Cr_n)}(Y_i) \right\}^6. \tag{3.2}$$

Therefore, for $t > 0$, by (3.2)

$$P(\#\Gamma_{\mathcal{F}_{n,1}} \ge t) \le nP(\#\Gamma_{\mathcal{G}_1} \ge t/n) \le nP\left(\sum_{i=1}^{n} I_{B(Y_1, Cr_n)}(Y_i) \ge (t/n)^{1/6} \right).$$

Apply $t = D^6 n(\log n)^6$ to the above inequality, we have that

$$P(\#\Gamma_{\mathcal{F}_{n,1}} \ge D^6 n(\log n)^6) \le nP\left(\sum_{i=2}^{n} I_{B(Y_1, Cr_n)}(Y_i) \ge D\log n - 1 \right). \tag{3.3}$$

By Lemma 2.2.1, for any $D > 20C^3$,

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

for large n, where K is a constant depends only on C. The above inequality and (3.3) yields the desired inequality by choosing D sufficiently large. ■

Let $\{X, X_i;\ i \ge 1\}$ be a sequence of i.i.d. $\mathbb{R}^d$-valued random variables with law P. P_n is the empirical law of $\{X_n\}$, i.e.,

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

We assume that $\mathcal{H}$, a class of subsets of $\mathbb{R}^d$, is a VC class with exponent v. Let $\{\mathcal{H}_n \subset \mathcal{H};\ n \ge 1\}$ be a sequence of subclass of sets. Denote $\mu_n := \sup_{V \in \mathcal{H}_n} P(V)$ and

$$\|P_n - P\|_n = \sup_{V \in \mathcal{H}_n} |P_n(V) - P(V)|.$$

We assume that there is a countable $\mathcal{H}'_n \subset \mathcal{H}_n$ such that

$$\sup_{V \in \mathcal{H}_n} |P_n(V) - P(V)| = \sup_{V \in \mathcal{H}'_n} |P_n(V) - P(V)| \quad a.s.$$

for all $n \geq 1$. Therefore, $\|P_n - P\|_n$ is measurable.

Lemma 3.1.3 *Suppose a sequence of constants* $\{t_n; n \geq 1\}$ *satisfy that* $\limsup_n (\mu_n/t_n) = C < \infty$. *Then, for* n *very large,*

$$P(\|P_n - P\|_n \geq t_n) \leq n^v \psi_n \exp(-\lambda n t_n),$$

where $\psi_n = 44 \vee 11(t_n)^{-\log_{2C'} 4}$, $\lambda = 1/2^7 C'$, *and* $C' = \max\{1, C\}$.

Proof. Let $\{\epsilon_i;\ i \geq 1\}$ be a sequence of i.i.d. Bernoulli sequence. By (11) in p.15 of [23],

$$P(\|P_n - P\|_n \geq t_n) \leq 4P\left(\sup_{F \in \mathcal{H}_n} |\frac{1}{n}\sum_{i=1}^{n} \epsilon_i I_F(X_i)| > \frac{t_n}{4}\right) \tag{3.4}$$

for all n so that $nt_n^2 \geq 8$, where $I_F(\cdot)$ is the indicator function of F. Set

$$A_n = \left\{\sup_{F \in \mathcal{H}_n} P_n(F) \leq 4C' t_n\right\}.$$

By Hoeffding's inequality(see, e.g. Appendix B in p.191 of [23]),

$$P^\epsilon\left(|\sum_{i=1}^{n} \epsilon_i I_F(X_i)| > \frac{t_n}{4}\right) \leq 2\exp\left\{-nt_n^2/(32P_n(F))\right\}. \tag{3.5}$$

Since $\limsup_n (\mu_n/t_n) = C$, there exists N_1 such that $\mu_n \leq 2C' t_n$ for all $n \geq N_1$. Note that by (3.1), $\#\{\{X_1, \cdots, X_n\} \cap F;\ F \in \mathcal{H}_n\} \leq n^v$ for $n \geq 2$. It follows from (3.4) and (3.5) that

$$\begin{aligned} P(\|P_n - P\|_n \geq t_n) &\leq 4n^v E^X \sup_{\mathcal{H}_n} P^\epsilon\left(|\sum_{i=1}^{n} \epsilon_i I_F(X_i)| > \frac{t_n}{4}\right) I_{A_n} + 4P(A_n^c) \\ &\leq 8n^v \exp(-\lambda_2 n t_n) + 4P(\|P_n - P\|_n \geq 2C' t_n) \end{aligned} \tag{3.6}$$

for $n \geq \max\{N_1, 8t_n^{-2}\}$. Repeat (3.6) to obtain

$$P(\|P_n - P\|_n \geq t_n) \leq \sum_{l=0}^{k} 8n^v \cdot 4^l \exp(-(2C')^l \lambda_2 n t_n) + 4^{k+1} P(\|P_n - P\|_n \geq (2C)^{k+1} t_n). \quad (3.7)$$

Let $k_0 = [-\log_{2C} t_n] \vee 1$, then $(2C)^{k_0+1} t_n > 1$, consequently, the probability on the RHS of (3.7) is zero because $\|P_n - P\|_n \leq 1$. Then by (3.7),

$$P(\|P_n - P\|_n \geq t_n) \leq 8n^v \exp(-\lambda_2 n t_n) \sum_{l=0}^{k_0} 4^l \leq 11 n^v \psi_n \exp(-\lambda_2 n t_n),$$

where $\psi_n = 44 \vee 11(t_n)^{-\log_{2C'} 4}$. ■

3.2 Maxima on random cubes

Throughout this section, we assume that $\{Y_i; i \geq 1\}$ is a sequence of i.i.d random variables with uniform distribution on $[0,1]^3$. Let $\{X_n; n \geq 1\}$ 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 in this section:

$$W_n = \max_{B \in \mathcal{B}} S_n(B), \text{ and } U_n = \max_{B \in \mathcal{F}} S_n(B),$$

where $\mathcal{B} = \{B = B(Y_i, r) \subset [0,1]^3;\ 1 \leq i \leq n,\ 0 < r < 1/2\}$ and $\mathcal{F}$ is the same as that in Lemma 3.1.1. Apparently, $W_{n,1} \leq W_{n,2}$.

Theorem 3.1 *Suppose condition 2.1 holds, then*

$$\lim_{n\to\infty} \frac{W_n}{\log n} = \frac{1}{\theta} \text{ a.s. and } \lim_{n\to\infty} \frac{U_n}{\log n} = \frac{1}{\theta} \text{ a.s.}$$

The following gives the asymptotic law of W_n.

Theorem 3.2 *Suppose condition 2.1 holds, then*

$$\lim_{n\to\infty} P\left(W_n \geq \frac{\log n}{\theta} + x\right) = 1 - \exp(-Ke^{-\theta x})$$

for any $x \in \mathbb{R}^1$*, where* K *is as in Lemma 2.2.4*

Before we prove Theorem 3.1, we need the following two lemmas.

Lemma 3.2.1 *Let* $\mathcal{F}$ *and* $\mathcal{F}_{n,1}$ *be the same as those in Lemma 3.1.1 and Lemma 3.1.2, respectively. Then there exists* $C > 0$ *so that*

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

Proof. Set $A_n = \{\sum_{i=1}^n I_{B(x,r)}(Y_i) \geq 4r^3 n\}$ for any $B(x,r) \in \mathcal{F}$. Since $\mathcal{F}$ is a VC class with exponent 6,

$$P(\max_{B\in\mathcal{F}\backslash\mathcal{F}_{n,1}} S_n(B) \geq 0) \leq n^6 E^Y \left\{\max_{B\in\mathcal{F}\backslash\mathcal{F}_{n,1}} P^X(S_n(B) \geq 0)(I_{A_n} + I_{A_n^c})\right\}.$$

For any $B = B(x,r) \in \mathcal{F}\backslash\mathcal{F}_{n,1}$, we have $r \geq r_n = C(\log n/n)^{1/3}$, it follows that

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

by Chernoff bound. Therefore

$$n^6 E^Y \left\{\max_{B\in\mathcal{F}\backslash\mathcal{F}_{n,1}} P^X(S_n(B) \geq 0)I_{A_n}\right\} \leq n^{6-4C^3\Lambda^*(0)} \leq 1/n^2 \tag{3.8}$$

if $C \geq (2\Lambda^*(0)^{-1})^{1/3}$. On the other hand, note $P(B(x,r_n)) = 8C^3\log n/n$,

$$E^Y \left\{\max_{B\in\mathcal{F}\backslash\mathcal{F}_{n,1}} P^X(S_n(B) \geq 0)I_{A_n}\right\} \leq P\left(\inf_{B(x,r)\subset[0,1]^3, r\geq r_n} \sum_{i=1}^n I_{B(x,r)}(Y_i) \leq 4C^3\log n\right)$$
$$\leq P\left(\inf_{B(x,r)\subset[0,1]^3} \sum_{i=1}^n I_{B(x,r_n)}(Y_i) \leq 4C^3\log n\right)$$
$$\leq P\left(\sup_{B(x,r)\subset[0,1]^3} \frac{1}{n}|\sum_{i=1}^n I_{B(x,r_n)}(Y_i) - P(B(x,r_n))| \geq 4C^3\frac{\log n}{n}\right) \leq 44n^{6-64C^3} \tag{3.9}$$

for large n, where we use Lemma 3.1.3 in the last step($C' = 2$). The last term in (3.9) is no greater than n^{-2} for $C \geq 1/2$. This fact and (3.8) prove the lemma. ■

Proof of Theorem 3.1 We only need to prove that

$$\limsup_{n\to\infty} \frac{U_n}{\log n} \leq \frac{1}{\theta} \quad a.s. \quad \text{and} \quad \liminf_{n\to\infty} \frac{W_n}{\log n} \geq \frac{1}{\theta} \quad a.s. \tag{3.10}$$

By Lemma 3.2.1, choose C so that

$$P(\max_{1\leq k\leq n} \max_{B\in\mathcal{F}\setminus\mathcal{F}_{n,1}} S_k(B) \geq 0) \leq nP(\max_{B\in\mathcal{F}\setminus\mathcal{F}_{n,1}} S_k(B) \geq 0) = O(n^{-2}). \tag{3.11}$$

Now define

$$V_n = \max_{1\leq k\leq n} \max_{B\in\mathcal{F}_{n,1}} S_k(B).$$

By Lemma 3.1.2, take D for which $B_n := \{\#\Gamma_n \geq Dn(\log n)^6\}$ has probability less than $O(1/n^2)$. Therefore, for any $\epsilon \in (0,1)$,

$$P(V_n/\log n > (1+\epsilon)/\theta) = \underbrace{E^Y P^X(V_n/\log n > (1+\epsilon)/\theta) I_{B_n}}_{G_n} + O(1/n^2).$$

But,

$$G_n \leq Dn(\log n)^6 E^Y \max_{B\in\mathcal{F}_{n,1}} P^X\left(\max_{1\leq k\leq n} \sum_{i=1}^{k} X_i \geq (1+\epsilon)/\theta\right) \leq D(\log n)^6/n^{\epsilon},$$

where we use the fact $E\exp(\theta X) = 1$ and submartingale inequality in the last inequality. Consequently, by (3.11)

$$P(\max_{1\leq k\leq n} U_k/\log n > (1+\epsilon)/\theta) = O(n^{-\epsilon}).$$

Observe that $\max_{1\leq k\leq n} U_k$ is non-decreasing, by the same arguments as (2.8), we obtain the $\limsup$ of (3.10).

Now we turn to prove lim inf of (3.10). Let $s_n = (\log n/8n\theta\Lambda'(\theta))^{1/3}$ and $B_i = B(Y_i, s_n)$. For any integer p(will be chosen specifically after given ϵ), define

$$L_n = \bigcup_{i=n^p+1}^{(n+1)^p} B(Y_i, 2s_n) \text{ and } J_n = \{1 \le j \le n^p;\ Y_j \in [s_n, 1-s_n]^3 \backslash L_n\}. \tag{3.12}$$

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

$$\inf_{n^p < k \le (n+1)^p} \left\{ \max_{B \in \mathcal{F}_{n,1}} S_n(B) \right\} \ge \max_{j \in J_n} S_n(B_j).$$

By Borel-Cantelli lemma, (3.11) implies that $\max_{B\in\mathcal{F}\backslash\mathcal{F}_{n,1}} S_n(B) \le 0$ eventually. So to prove the liminf in (3.10), it is enough to show that

$$\liminf_{n\to\infty} \frac{\max_{j\in J_n} S_n(B_j)}{\log n^p} \ge \frac{1}{\theta} \quad a.s. \tag{3.13}$$

Define

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

We claim that

$$P(N_n \le n^p/\log n) = O(n^{-2}) \tag{3.14}$$

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

$$P\left(\sum_{i=1}^{l_n} I\left(\text{at least one of } \{Y_j, 1\le j\le n^p\} \in B_i\right) \le n^p/\log n\right)$$
$$\le \frac{l_n}{l_n - n^p(\log n)^{-1}}(1 - 8s_n^3)^n = O(n^{-2}) \tag{3.15}$$

for $p > 2\theta\Lambda'(\theta)$. Thus claim (3.14) follows.

For any given $\epsilon \in (0, 1/4)$, set $z_n = (1-\epsilon)\log n^p/\theta$. Then, by (3.14)

$$P(\max_{j\in J_n} S_n(B_j)/\log n^p \le z_n) \le E^Y \prod_{i=1}^{N_n} P^X(S_n(\tilde{A}_i) \le z_n)$$

$$\le \underbrace{E^Y \exp\{-\sum_{i=1}^{N_n} P^X(S_n(\tilde{A}_i) > z_n)\} I(N_n \ge n^p/\log n)}_{F_n} + O(n^{-2}).$$

Let $\Gamma_k = \{\{Y_1, \cdots, Y_{n^p}\} \cap [s_n, 1-s_n]^3 = \{Y_1, \cdots, Y_k\}\}$. For any random variable ξ and any set B such that $P(B) > 0$, define $E(\xi|B) = E\xi I_B/P(B)$. Then, by symmetry,

$$F_n \le \max_{n^p/\log n \le k \le n^p} \zeta_k P\left(\sum_{i=1}^{k} I_{[s_n,1-s_n]^3}(Y_i) \ge n^p/\log n\right) \le \max_{n^p/\log n \le k \le n^p} \zeta_k,$$

where

$$\zeta_k = E^Y\left\{\exp\left(-\sum_{i=1}^{k} P^X(S_n(B_i) \ge z_n)\right)\Big|\Gamma_k\right\}.$$

Therefore,

$$P(\max_{j\in J_n} S_n(B_j)/\log n^p \le z_n) \le \max_{n^p/\log n \le k \le n^p} \zeta_k + O(n^{-2}). \tag{3.16}$$

Lemma 3.2.2 below will prove that $\max_{n^p/\log n \le k \le n^p} \zeta_k = O(n^{-2})$ for some p. So

$$\sum_{n\ge 1} P(\max_{j\in J_n} S_n(B_j)/\log n^p \le z_n) < \infty, \quad \forall\epsilon \in (0, 1/4).$$

By Borel-Cantelli Lemma, we have (3.13). ■

Lemma 3.2.2 *Denote $\Omega_n = \{k \in \mathbb{N};\ n^p/\log n \le k \le n\}$. For any given $\epsilon \in (0, 1/4)$, there exists $p > 3$ such that*

$$\max_{k\in\Omega_n} \zeta_k = O(n^{-2})$$

Proof. Let $N_i = \#\{1 \le j \le n^p;\ Y_j \in B_i\},\ i = 1, 2, \cdots, k$. By Bernstein's inequality(Lemma 2.2.1), for any given $x \in (0,1)$,

$$P(N_i \notin \underbrace{\left(\frac{(1-x)\log n^p}{\theta\Lambda'(\theta)}, \frac{(1+x)\log n^p}{\theta\Lambda'(\theta)}\right)}_{O_x}) \le \exp\{-(px^2/6\theta\Lambda'(\theta))\log n\} \tag{3.17}$$

for large n. It follows from Chbyshev's inequality(similar to (3.15)) that,

$$\sup_{k\in\Omega_n} P\left(\sum_{i=1}^{k} I_{O_x}(N_i) \le n^p/\log n\right) = O(n^{-2}) \tag{3.18}$$

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

$$P^X(S_n(B_i) > z_n) \ge \min_{l\in O_x} \frac{Ce^{-l\Lambda_X^*(z_n/l)}}{\sqrt{l}}. \tag{3.19}$$

By Taylor's expansion at $(1-\epsilon)\Lambda'(\theta)$, we see that $\lim_{x\to 0^+} \frac{l\Lambda_X^*(z_n/l)}{\log n^p} = 1-\epsilon$ uniformly on $l \in O_x$. So, for the given $\epsilon \in (0, 1/4)$, choose suitable $x_0 \in (0,1)$ in the definition of O_x, we then have that

$$\inf_{i;N_i\in O_x} P^X(S_n(B_i) > z_n) \ge Cn^{(0.5\epsilon-1)p}/\sqrt{\log n}$$

for some constant C depends only on p and X. Consequently, by (3.18), we have that

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

The desired equality follows from taking any $p > \max\{3,\ 2\epsilon^{-1},\ 12\theta\Lambda'(\theta)/x_0^2\}$. ■

Denote

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

and $\Omega_n = \{(r,i) \in [l_n^-, l_n^+] \times \{1, 2, \cdots, n\};\ B(Y_i, l_n^+) \in [0,1]^3\}$, and

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

Lemma 3.2.3 *Let* $z_n = \log n/\theta + x$. *Then,* $\forall \alpha > 0$, *there exists* $\beta > 0$ *such that*

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

Proof. By Lemma 3.2.1, it is enough to show that

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

for some $\beta > 0$, where $\Omega'_n = \{(0, Cr_n)\backslash(l_n^-, l_n^+)\}$ for some $C > (2\theta\Lambda'(\theta))^{-1/3}$. Recall $r_n = (\log n/n)^{1/3}$ in Lemma 3.1.2. Set $D_n = \{B \cap \{Y_1, Y_2, \cdots, Y_n\}; B = B(Y_i, r), 1 \leq i \leq n, 0 < r < Cr_n\}$. By Lemma 3.1.2, $\#D_n \leq n(\log n)^6$ with probability $1 - O(n^{-2})$. Then there exist $k_n \leq n(\log n)^6$ and $\tilde{B}_1, \cdots, \tilde{B}_{k_n}$ as the representatives of those B's in D_n which may depend on Y_i's such that $\max_{(r,i)\in\Omega'_n} S_n(B(Y_i, r)) = \max_{1\leq i\leq k_n} S_n(\tilde{B}_i)$ with probability $1 - O(n^{-2})$. Therefore,

$$\begin{aligned} & E^Y \left\{ P^X \left(\max_{(r,i)\in\Omega'_n} S_n(B(Y_i, r)) > z_n \right) \right\} \\ & \leq n(\log n)^6 E^Y \left\{ \max_{(r,i)\in\Omega'_n} P^X(S_n(B(Y_i, r)) > z_n) \right\} + O(n^{-2}). \end{aligned} \tag{3.20}$$

Set $c_0 = (\theta\Lambda'(\theta))^{-1/3}$. Note that $|n \cdot \mathrm{vol}(B(Y_i, r)) - c_0^3 \log n| \geq 3\beta c_0^3((\log n) \log_2 n)^{1/2}$ for any $(r, i) \in \Omega'_n$. Thus

$$\begin{aligned} & \left\{ |\sum_{j=1}^n I_{B(Y_i,r)}(Y_j) - n \cdot \mathrm{vol}(B(Y_i, r))| \leq (3\beta c_0^3/2)\sqrt{(\log n) \log_2 n} \right\} \\ \subset & \left\{ |\sum_{j=1}^n I_{B(Y_i,r)}(Y_j) - c_0^3 \log n| \geq (3\beta c_0^3/2)\sqrt{(\log n) \log_2 n} \right\} := G_n. \end{aligned}$$

So, by Bernstein's inequality, $P(G_n^c) \leq 2e^{-c_0^3\beta^2 \log_2 n}$ for n large. It follows that for any $(r, i) \in \Omega'_n$.

$$P^X(S_n(B(Y_i, r)) > z_n) \leq P^X(S_n(B(Y_i, r)) > z_n)I_{G_n} + \frac{2e^{-x}}{n} I_{G_n^c}, \tag{3.21}$$

where we obtain the second term of above by using the Chernoff bound for $P^X(S_n(B(Y_i,r)) > z_n)$. By Corollary 2.1, there exists a constant C_1 depending only on X so that

$$P^X(S_n(B(Y_i,r)) > z_n)I_{G_n} \leq \frac{e^{-\theta x}}{n(\log n)^{C_1\beta^2+1/2}}$$

uniformly on $(r,i) \in \Omega_n'$. Combining this with (3.20) and (3.20), we finally get

$$P(\max_{(r,i)\in\Omega_n'} S_n(B(Y_i,r)) > z_n) \leq \frac{e^{-\theta x}}{n(\log n)^{C_1\beta^2-5.5}} + \frac{4e^{-x}}{(\log n)^{c_0^3\beta^2-6}}$$

for n large. The proof is completed by choosing β sufficiently large. ■

Lemma 3.2.4 *Let $T_n = \{1 \leq i \leq n;\ Y_i \in [l_n^+, 1-l_n^+]^3\}$. Then,*

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

where K is as that in Lemma 2.2.4.

Proof. First, by Lemma 2.2.4,

$$\lambda_n \leq nP(\max_{i\geq 1} S_n > z_n) \to Ke^{-\theta x} \quad a.s. \tag{3.22}$$

By Bernstein's inequality(Lemma 2.2.1),

$$P(|T_n - n(1-2l_n^+)^3| \geq \log n) \leq n^{-\xi} \tag{3.23}$$

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

$$A_i = \left\{ |\sum_{j=1}^n I(\|Y_j - Y_i\| \leq l_n^+) - 8n(l_n^+)^3| \leq h_n \right\}, \quad i = 1,2,\cdots,n.$$

It is easy to check that

$$P(A_1^c) \leq 2\exp\{-h_n^{1/3}/3\}. \tag{3.24}$$

for n large. Consequently, by Chebyshev's inequality(like the argument used in obtaining (3.15)),

$$P\left(\sum_{i=1}^{n} I(A_i) \le n - nP(A_1^c)^{1/2}\right) = O(P(A_1^c)^{1/2}) = O\left(\exp(-(\log n)^{1/6})\right). \tag{3.25}$$

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

$$L_i = \left\{|\sum_{j=1}^{n} I(l_n^- \le \|Y_j - Y_i\| \le l_n^+) - 8n((l_n^+)^3 - (l_n^-)^3)| \le h_n'\right\}, \quad i = 1, 2, \cdots, n.$$

By Bernstein's inequality again,

$$P(L_i^c) \le 2\exp\{-(\theta\Lambda'(\theta)/4\beta){h_n'}^{1/3}\} \tag{3.26}$$

for n large. Thus

$$P\left(\sum_{i=1}^{n} I(L_i) \le n - nP(L_1^c)^{1/2}\right) = O(P(L_1^c)^{1/2}) = O(\exp(-(\log n)^{1/20})). \tag{3.27}$$

From (3.23), (3.25) and (3.27), it follows that with probability approaching to 1, at least $n - n^{3/4} - nP(A_1^c)^{1/2} - nP(L_1^c)^{1/2} = n - o(n)$ of $\{Y_i; 1 \le i \le n\}$

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 $8n(l_n^+)^3 - h_n \sim (\theta\Lambda'(\theta))^{-1}(\log n + \beta\sqrt{(\log n)\log_2 n})$ elements of $\{Y_1, Y_2, \cdots, Y_n\}$;

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

By Lemmas 2.2.4, 2.2.9 and 2.2.10, there exists $\gamma > 0$ for which

$$nP(\max_{k \in Q_n} S_k \ge z_n) \to Ke^{-\theta x}, \tag{3.28}$$

where $Q_n = \mathbb{N} \cap \{k;\ |k - \log n/\theta\Lambda'(\theta)| \leq \gamma\sqrt{(\log n)\log_2 n}\}$. Therefore, by (3.28), and b) and c) above,

$$P^Y\left(\lambda_n \geq (n - o(n))(1/n)\{Ke^{-\theta x} + o(1)\}\right) \to 0,$$

which together with (3.22) proves the Lemma. ■

Proof of Theorem 3.2. Define $V_{n,i} = \max_{l_n^- \leq r \leq l_n^+} S_n(B(Y_i, r)), \quad i = 1, 2, \cdots, n$. Then $W_{n,1} = \max_{i \in T_n} V_{n,i}$, where $T_n = \{1 \leq i \leq n; Y_i \in [l_n^+, 1 - l_n^+]^3\}$ as before. By Lemma 3.2.3, to prove the theorem, it is enough to show that

$$P\left(\max_{i \in T_n} V_{n,i} > z_n\right) \to 1 - \exp(-K^* e^{-\theta x}). \tag{3.29}$$

By Lemma 2.2.3, we have

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

where $E^Y e^{-\lambda_n} \to \exp(-Ke^{-\theta x})$ by Lemma 3.2.4 and Dominated Convergence Theorem. Also,

$$\begin{aligned} b_1 &:= E^Y\left\{\sum_{i \in T_n}\sum_{j \in T_n} P^X(V_{n,j} > z_n)P^X(V_{n,i} > z_n)I(d(Y_j, Y_i) \leq 2l_n^+)\right\} \\ &\leq \frac{e^{-2\theta x}}{n^2} \cdot n^2 \cdot P(d(Y_1, Y_2) \leq 2l_n^+) = e^{-2\theta x}O(\log n/n) \end{aligned}$$

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

$$\begin{aligned} b_2 &:= E^Y\left\{\sum_{i \in T_n}\sum_{j \in T_n} P^X(V_{n,j} > z_n, V_{n,i} > z_n)I(d(Y_j, Y_i) \leq 2l_n^+)\right\} \\ &= n^2 E^Y\left\{P^X(V_{n,1} > z_n, V_{n,2} > z_n)(I_{\Psi_n} + I_{\Psi'_n})\right\}, \end{aligned}$$

where

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

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

$$n^2 E^Y \left\{P^X(V_{n,1} > z_n, V_{n,2} > z_n) I_{\Psi_n'}\right\} \le e^{-\theta x}(\log n)^{-3\delta}. \tag{3.30}$$

It is easy to check that $B(Y_2, l_n^-) \backslash B(Y_1, l_n^+)$ has volume

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

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

$$P(\underbrace{\sum_{i=1}^n I_{B(Y_2, l_n^-)\backslash B(Y_1, l_n^+)}(Y_i) \le nv_n - \sqrt{\log n}}_{E_n}) \le \exp(-C(\log n)^{1/3+\delta}). \tag{3.31}$$

By (3.26),

$$\begin{aligned} &n^2 E^Y \left\{P^X(V_{n,1} > z_n, V_{n,2} > z_n) I_{\Psi_n}\right\} \\ &\le Cn^2(\log n)\log_2 n E^Y \left\{\max_{l_n^- \le r_1, r_2 \le l_n^+} P^X(S_n(B(Y_i, r_i)) > z_n,\ i = 1, 2) I_{\Psi_n \cap L_1 \cap L_2}\right\} \\ &\quad + \underbrace{n^2 P^{Y_1, Y_2}(\Psi_n)(n^{-1}e^{-\theta x}) \cdot 4e^{-(\log n)^{0.1}}}_{O((\log n)e^{-(\log n)^{0.1}})}. \end{aligned} \tag{3.32}$$

For any $r_1, r_2 \in (l_n^-, l_n^+)$, $B(Y_2, r_1)\backslash B(Y_1, r_2) \supset B(Y_2, l_n^-)\backslash B(Y_1, l_n^+)$ on Ψ_n. Thus, by Lemma 2.2.8 and (3.31), there exists $\zeta > 0$ such that

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

for any fixed $Y_1, Y_2 \in [l_n^+, 1-l_n^+]^3$. Consequently, the first term of the RHS of (3.32) is less than

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

Combining the above equality, (3.32) and (3.30), we conclude that $b_2 \to 0$. ■

Chapter 4

The Comparison of Two Pictures

4.1 Some auxiliary lemmas.

Let $\{X,\ X_i,\ i=1,2,\cdots\}$ and $\{Y,\ Y_i, i=1,2,\cdots\}$ be two sequences of i.i.d. random vectors taking values in some space Σ(not necessary $\mathbb{R}^d$). X and Y may not necessarily have the same distribution. $F(\cdot,\cdot):\mathcal{X}\times\mathcal{X}\to\mathbb{R}^1$, is usually called score function in terms of protein matching problems. Throughout all this chapter, we assume that $F(\cdot,\cdot)$ is non-lattice. Denote $\Lambda(t)=\Lambda_F(t)=\log Ee^{tF(X,Y)}$, $t\in\mathbb{R}^1$. If condition (2.1) holds for $Z=F(X,Y)$, then there exists unique $\theta=\theta_F>0$ so that $Ee^{\theta F(X,Y)}=1$. Set $\mu_F=EF(X,Y)$, $\phi_X(x)=\log\{E^Y e^{\theta F(x,Y)}\}$ and $h_X=E\left\{e^{\theta F(X,Y)}\left(\log E^Y e^{\theta F(X,Y)}\right)\right\}$. Define $\phi_Y(y)$ and h_Y by symmetry.

Lemma 4.1.1 *Suppose condition (2.1) holds for $Z=F(X,Y)$. Then, for any $a\in(0,1),\epsilon>0$ and sequence $\{\gamma_n;\ n\geq 1\}$ so that $\gamma_n\to\gamma\in\mathbb{R}^1$, there exists $\delta>0$ such that*

$$\max_{an\leq k\leq n} P\left(\frac{1}{n}\sum_{i=1}^{n}F(X_i,Y_i)\geq\gamma_n,\ \left|\frac{1}{k}\sum_{i=1}^{k}\phi_Y(Y_i)-h_Y\right|\geq\epsilon\right)=o(e^{-(\theta\gamma+\delta)n})$$

Proof. For any sequence $\{k_n,n\geq 1\}$, define

$$Z_n=\left(\frac{1}{n}\sum_{i=1}^{n}F(X_i,Y_i),\frac{1}{k_n}\sum_{i=1}^{k_n}\phi_Y(Y_i)\right)\in\mathbb{R}^2.$$

To prove this lemma, it is enough to show that for the given $c > 0$ and $\epsilon > 0$, there exists $\delta > 0$ such that

$$p_n := P\left(Z_n \in [\gamma_n, \infty) \times (h_Y - \epsilon, h_Y + \epsilon)^c\right) = o(e^{-(\theta\gamma+\delta)n}) \tag{4.1}$$

for all subsequences k_n so that $k_n/n \to a' \in [a, 1]$. Let $F_\eta = [\gamma - \eta, \infty) \times \{y;\ |y - h_Y| \geq \epsilon\}$. Then F_η is a closed set in $\mathbb{R}^2$. It is easy to see that $p_n \leq P\left(Z_n \in F_\eta\right)$ for n large enough. For any $(u, v) \in \mathbb{R}^2$

$$\begin{aligned}&\frac{1}{n}\log Ee^{n(u,v)\cdot Z_n} = \frac{n-k_n}{n}\log Ee^{uF(X,Y)} + \frac{k_n}{n}\log Ee^{uF(X,Y)+(nv/k_n)\phi_Y(Y)} \\ &\to (1-a')\log Ee^{uF(X,Y)} + a'\log\left\{Ee^{uF(X,Y)}(E^X e^{\theta F(X,Y)})^{v/a'}\right\} := g(u,v).\end{aligned}$$

Clearly, $g(u, v)$ is finite and differentiable for any $(u, v) \in \mathbb{R}^2$. For any $\eta > 0$, by the Gatner-Ellis Theorem(cf. Theorem 2.3.6 of [12]),

$$\limsup_{n\to\infty}(\log p_n)/n \leq -\inf_{(x,y)\in F_\eta} I((x,y)),$$

where $I((x, y)) = \sup_{(u,v)\in\mathbb{R}^2}\{ux + vy - g(u, v)\}$. Note that $F_\eta \downarrow F_0$ as $\eta \downarrow 0$. Therefore, by taking $\eta \downarrow 0$, we obtain

$$\limsup_{n\to\infty}(\log p_n)/n \leq -\inf_{(x,y)\in F_0} I((x,y)). \tag{4.2}$$

It is easy to see that $g(\theta, v) \leq a\log\left\{Ee^{\theta F(X,Y)}(E^X e^{\theta F(X,Y)})^{v/a}\right\}$ since $E^{\theta F(X,Y)} = 1$ and $a' \geq a$. Therefore, for any $x \geq \gamma$ and $y \leq h_Y - \epsilon$,

$$I((x,y)) - \theta\gamma \geq a \cdot \sup_{v\leq 0}\{\underbrace{vh_Y - v\epsilon - \log(Ee^{\theta F(X,Y)}(\overbrace{E^X e^{\theta F(X,Y)})^v}^{\xi(v)})}_{\psi(v)}\}. \tag{4.3}$$

Obviously, $\psi(0) = 0$, and observe that $E(e^{\theta F(X,Y)}\xi'(v)) \to h_Y$ as $v \uparrow 0$. It then follows that

$$\psi'(v) = h_Y - \epsilon - \left(Ee^{\theta F(X,Y)}\xi(v)\right)^{-1} E\left(e^{\theta F(X,Y)}\xi'(v)\right) < 0$$

for $v < 0$ and $|v|$ small enough. Therefore, by (4.3), there exists $\delta > 0$ such that $I((x, y)) > \theta\gamma + 2\delta$ for all $x \geq \gamma$ and $y \leq h_Y - \epsilon$. By the same arguments, it is also true for $x \geq \gamma$ and $y \geq h_Y + \epsilon$. Thus, $\inf_{(x,y)\in F_0} I((x, y)) > \theta\gamma + 2\delta$ for some $\delta > 0$, which together with (4.2) yields (4.1). ■

Lemma 4.1.2 *Suppose condition (2.1) holds for $Z = F(X, Y)$. Then, $\forall\epsilon > 0$,*

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\phi_Y(Y_i) - h_Y\right| \leq \epsilon\right) \leq 2e^{-n(h_Y-\epsilon)}.$$

Proof. By the Chernoff bound(see e.g. p.p. 31 in [12])

$$P\left(\frac{1}{n}\left|\sum_{i=1}^{n}\phi_Y(Y_i) - nh_Y\right| \leq \epsilon\right) \leq 2e^{-n\inf_{x\in A} J(x)},$$

where $J(x) = \sup_{t\in\mathbb{R}^1}\{tx - \log E\left(E^X e^{\theta F(X,Y)}\right)^t\}$ and $A = \{x \in \mathbb{R}^1;\ |x - h_Y| \leq \epsilon\}$. For any $x \in A$, we have $x \geq h_Y - \epsilon$. Note that $Ee^{\theta F(X,Y)} = 1$. We then have

$$J(x) \geq h_Y - \epsilon - \log E(E^X e^{\theta F(X,Y)}) = h_Y - \epsilon$$

for any $x \in A$. ■

Lemma 4.1.3 *Suppose $E\{F(X_1, Y)e^{\theta F(X_2,Y)}\} < 0$, where X_1 and X_2 are i.i.d. Let $M(t) = E\exp(tF(X_1, Y) + \theta F(X_2, Y))$ and $t_0 = \sup\{t > 0; M(t) < 1\}$. Then*
(i) $t_0 \in (0, \theta)$.
(ii) There exists $\delta \in (\mu_F, 0)$ such that $\gamma_1 := \sup_{0<t<t_0}\{\delta t - \log\Lambda_F(t)\} > 0$ and $\gamma_2 := \sup_{0<t<t_0}\{\delta t - \log M(t)\} > 0$.

Proof. (i) Note that $M(0) = 1$ and $M'(0) = EF(X_1, Y)e^{\theta F(X_2,Y)} < 0$, so there exists some $t > 0$ such that $M(t) < 1$. Since $Ee^{\theta F(X,Y)} = 1$, by Hölder's inequality, $M(t) \geq M(\theta)^{t/\theta} = (E^Y(E^X e^{\theta F(X,Y)})^2)^{t/\theta} > (Ee^{\theta F(X,Y)})^{2t/\theta} = 1$ for all $t > \theta$. Thus, $t_0 < \theta$.
(ii) By (i), $\log\Lambda(t_0/2) < 0$ and $\log M(t_0/2) < 0$. Define

$$\delta = \frac{1}{2}\max\left\{EF(X, Y),\ 2t_0^{-1}\log\Lambda(t_0/2).\ 2t_0^{-1}\log M(t_0/2)\right\}, \quad ■$$

4.2 Main results and their proofs

In this section, capital letters such as I, J, K etc. are indices taking values in $\mathbb{N}^2$. We follow the convention that $(i_1, i_2) = I \leq J = (j_1, j_2)$ if and only if $i_1 \leq j_1$ and $i_2 \leq j_2$. We say $I < J$ if at least one of the previous two inequalities is strict. For convenience, sometimes we write $I \leq m \in \mathbb{N}$ to denote that $I \leq (m, m)$.

Let $\{X, X_I;\ I \in \mathbb{N}^2\}$ and $\{Y, Y_I;\ I \in \mathbb{N}^2\}$ be two sets of i.i.d. random elements taking values in the space $\mathcal{X}$ with σ-algebra $\mathcal{B}$. A map $F(\cdot,\cdot) : (\mathcal{X}, \mathcal{B}) \to \mathbb{R}$ is a measurable function. Define $\Gamma_\Delta = \{1, 2, \cdots, \Delta\}^2$. In this section we consider

$$W_n = \max_{\substack{0 \leq I, J \leq n-\Delta \\ \Delta > 0}} \left\{ \sum_{K \in \Gamma_\Delta} F(X_{I+K}, Y_{J+K}) \right\}.$$

In one of the following theorems, the following condition is needed:

$$EF(X_1, Y)e^{\theta F(X_2, Y)} < 0, \quad EF(X, Y_1)e^{\theta F(X, Y_2)} < 0. \tag{4.4}$$

We also assume that condition (2.1) holds for the random variable $F(X, Y)$. Therefore, there exists $\theta = \theta_F > 0$ such that $Ee^{\theta F(X,Y)} = 1$. Define a measure α^* on $(\mathcal{X}^2, \mathcal{B}^2)$ by

$$\frac{d\alpha^*}{d(\mu_X \times \mu_Y)} = e^{\theta F},$$

where μ_X and μ_Y are the distributions of X and Y, respectively. For any two probability measures μ, ν on $(\mathcal{X}^2, \mathcal{B}^2)$, recall the definition of relative entropy $H(\nu|\mu)$:

$$H(\nu|\mu) = \begin{cases} \int_{\mathcal{X}} \left(\log \frac{d\nu}{d\mu}\right) d\nu, & \text{if } \nu \ll \mu, \\ +\infty, & \text{otherwise.} \end{cases}$$

The following assumption will be used later.

$$H(\alpha^*|\mu_X \times \mu_Y) > 2\max(H(\alpha^*_X|\mu_X), H(\alpha^*_Y|\mu_Y)). \tag{4.5}$$

A detailed discussion about (4.5) is given in [10] and [11]. To understand the above condition, we give the following easy lemma without proof.

Lemma 4.2.1 *Denote* $\xi_X = E^Y e^{\theta F(X,Y)}$ *and* $\xi_Y = E^X e^{\theta F(X,Y)}$. *Then,* $h_X = E^X(\xi_X \log \xi_X)$, *and* $h_Y = E^Y(\xi_Y \log \xi_Y)$, *and (4.5) is equivalent to*

$$\frac{1}{2}\theta\Lambda'(\theta) > \max\{E^X(\xi_X \log \xi_X), E^Y(\xi_Y \log \xi_Y)\}.$$

We will use the following function:

$$\delta(x) = \sum_{i=-\infty}^{+\infty} \exp\{-\beta(i+x)^2\},\ x \in R^1,$$

where $\beta = 2\Lambda'(\theta)^2/\Lambda''(\theta)$, where all $\Lambda(\cdot)$ in this section is $\Lambda_{F(X,Y)}$. The main results are the following.

Theorem 4.1 *Suppose conditions (2.1) and (4.5) hold. Then,*

$$\lim_{n\to\infty} \frac{W_n}{\log n} = \frac{4}{\theta} \quad a.s.$$

Theorem 4.2 *Suppose condition (2.1) holds for* $Z = F(X,Y)$. *Also, assume conditions (4.4) and (4.5) hold. Denote* $t_n = \log \delta\left(2\sqrt{\frac{\log n}{\theta\Lambda'(\theta)}}\right)$. *Then*

$$P\left(W_n \le \frac{1}{\theta}\left\{4\log n - \frac{1}{2}\log_2 n + t_n\right\} + x\right) \to e^{-K^* e^{-\theta x}},$$

where $K^* = \sqrt{\Lambda'(\theta)/(8\pi\theta\Lambda''(\theta))}$.

As shown in [11], condition (4.4) is not required in the one-dimensional counterpart of Theorem 4.2. We need this condition here because we don't have one-dimensional structure as in [11].

The following example is motivated by the analysis of protein sequences. See [10]

Example. On some space Σ with $1 < |\Sigma| < \infty$, define $F(x,x) = 1$ and $F(x,y) = -m\sqrt{2}$ for $x \neq y$, $x, y \in \Sigma$ with $m \in \mathbb{N}$. Then, $F(\cdot,\cdot)$ is non-lattice. It is easy to check that conditions (2.1) and (4.4) hold for sufficiently large m. Furthermore, condition (4.5) holds for sufficiently large m provided the following condition((1.7) in [10]) is true:

$$\max\left\{\sum_{i\in\Sigma}\mu_X(i)\mu_Y(i)\log\mu_Y(i), \sum_{i\in\Sigma}\mu_X(i)\mu_Y(i)\log\mu_X(i)\right\} < -\frac{1}{2}\theta e^{-\theta}.$$

Proof of Theorem 4.1. First, we prove the upper bound, i.e.

$$\limsup_{n\to\infty}\frac{W_n}{\log n} \leq \frac{4}{\theta} \quad a.s.$$

Note that W_n is non-decreasing in n. By the same arguments as (2.7), it is enough to show that for any given $\epsilon > 0$, there exists $\delta > 0$ such that

$$P\left(W_n \geq \frac{4(1+\epsilon)\log n}{\theta}\right) \leq \frac{1}{n^\delta} \tag{4.6}$$

for n is sufficiently large. By Lemma 2.2.9, there exist $\gamma > 1$ and $\delta_1 > 0$ such that

$$\sum_{(I,J,\Delta)\in\mathcal{B}_n} P\left(\sum_{K\in\Gamma_\Delta} F(X_{I+K}, Y_{J+K}) > \frac{4(1+\epsilon)\log n}{\theta}\right) \leq \frac{1}{n^{\delta_1}},$$

for n sufficiently large, where $\mathcal{B}_n = \{(I,J,\Delta) \in \mathbb{N}^5;\ (I \vee J) + \Delta \leq n, \frac{\Delta}{\sqrt{\log n}} \notin (\gamma^{-1}, \gamma)\}$. Obviously, the cardinality of $\mathcal{B}'_n =: \{(I,J,\Delta) \in \mathbb{N}^5;\ I+\Delta, J+\Delta \leq n, \frac{\Delta}{\sqrt{\log n}} \in (\gamma^{-1},\gamma)\}$ is at most $(\gamma-\gamma^{-1})^2 n^4(\log n)$. We then have

$$\begin{aligned} &P\left(\max_{(I,J,\Delta)\in\mathcal{B}'_n}\sum_{K\in\Gamma_\Delta} F(X_{I+K}, Y_{J+K}) > \frac{4(1+\epsilon)\log n}{\theta}\right) \\ &\leq |\mathcal{B}_n| \max_{\gamma^{-1}\sqrt{\log n}\leq\Delta\leq\gamma\sqrt{\log n}} P\left(\sum_{K\in\Gamma_\Delta} F(X_K, Y_K) \geq \frac{4(1+\epsilon)\log n}{\theta}\right) \leq \frac{(\gamma-\gamma^{-1})^2(\log n)}{n^{4\epsilon}}, \end{aligned}$$

where we use the Chebyshev inequality in the last inequality and the fact that $E\exp(\theta F(X,Y))$ $= 1$. Therefore, (4.6) follows.

Now we prove the lower bound. By subsequence arguments as in the proof of liminf in (2.7), we need to show that for any given $\gamma \in (0,1)$, there exists $\delta > 0$ such that

$$P(W_n \leq z_n) \leq n^{-\delta}, \tag{4.7}$$

where $z_n = 4\gamma \log n/\theta$. Let $l_n = [2\sqrt{\log n/\theta\Lambda'(\theta)}]$, $\Gamma_n = \{1, 2, \cdots, l_n\}^2$,

$$\mathcal{C}_n = \left\{(sl_n, tl_n);\ (s,t) \in \mathbb{N}^2;\ 0 \le s, t \le \left[\frac{n}{l_n}\right] - 1\right\},$$
$$W'_n = \max_{(I,J)\in\mathcal{C}_n^2} \sum_{K\in\Gamma_n} F(X_{I+K}, Y_{J+K}),$$

where $\mathcal{C}_n^2 = \mathcal{C}_n \times \mathcal{C}_n$. Obviously, $W_n \ge W'_n$ for all $n \ge 1$. Define $E_{\Gamma,\Gamma'}(\epsilon) = E_{X,\Gamma}(\epsilon) \times E_{Y,\Gamma'}(\epsilon)$, where

$$E_{X,\Gamma}(\epsilon) = \left\{|\frac{1}{l_n^2}\sum_{I\in\Gamma} \phi_X(X_I) - h_X| < \epsilon\right\},\quad \text{and}$$
$$E_{Y,\Gamma'}(\eta) = \left\{|\frac{1}{l_n^2}\sum_{I\in\Gamma'} \phi_Y(Y_I) - h_Y| < \epsilon\right\},\quad \Gamma,\ \Gamma' \in \mathcal{C}_n,\ \epsilon > 0,$$

and $\phi_X(x)$ and h_X etc. are those as in section 4.1. Note that $z_n/l_n^2 \to \gamma\Lambda'(\theta)$, by Lemma 4.1.1, there exists positive δ depends only on ϵ such that

$$|P(W_n > z_n) - \underbrace{P(\bigcup_{(I,J)\in\mathcal{C}_n^2}\{\sum_{K\in\Gamma_n} F(X_{I+K}, Y_{J+K}) > z_n,\ E_{I+\Gamma_n, J+\Gamma_n}(\epsilon)\})}_{\mathcal{H}_n}|$$
$$\le n^4 e^{-(\gamma\theta\Lambda'(\theta)+\delta)l_n^2} \le n^{-\delta'}, \tag{4.8}$$

where $\delta' = \delta/\theta\Lambda'(\theta)$ by taking $\gamma > 1 - \delta'$.

Now we use Chen-Stein method to estimate $\mathcal{H}_n$. Observe that for any $(I_0, J_0) \in \mathcal{C}_2$, there are at most $2[n/l_n]^2$ pairs of $(I,J) + \Gamma_n \in \mathcal{C}_n^2 + \Gamma_n$ which may intersect $(I_0, J_0) + \Gamma_n$. So by Theorem 1 from [1], we have

$$P(\mathcal{H}_n^c) \le \exp(-[n/l_n]^4 \xi_n) + \left[\frac{n}{l_n}\right]^6 \xi_n^2 + \left[\frac{n}{l_n}\right]^6 (\eta_n + \eta'_n), \tag{4.9}$$

where $\xi_n = P\left(\sum_{K\in\Gamma_n} F(X_I, Y_I) > z_n, E_{\Gamma_n,\Gamma_n}(\epsilon)\right)$

$$\eta_n = P\left(\sum_{I\in\Gamma_n} F(X_I, Y_I) \geq z_n, \sum_{I\in\Gamma_n} F(X_I, Y_{I+l_n}) \geq z_n, E_{\Gamma_n,\Gamma_n}(\epsilon) \cap E_{\Gamma_n,\Gamma_n+l_n}(\epsilon)\right),$$
$$\eta_n' = P\left(\sum_{I\in\Gamma_n} F(X_I, Y_I) \geq z_n, \sum_{I\in\Gamma_n} F(X_{I+l_n}, Y_I) \geq z_n, E_{\Gamma_n,\Gamma_n}(\epsilon) \cap E_{\Gamma_n+l_n,\Gamma_n}(\epsilon)\right).$$

Since $Ee^{F(X,Y)} = 1$, by Chebyshev's inequality, we have that

$$\left[\frac{n}{l_n}\right]^6 \xi_n^2 \leq n^6 e^{-2\theta z_n} = \frac{1}{n^{8\gamma-6}} < \frac{1}{n^{1.1}} \tag{4.10}$$

for $\gamma > 9/10$. Now we estimate ξ_n more precisely. By Lemma 4.1.1, $\xi_n \sim P(S_{l_n^2} > z_n)$ for $\gamma < 1$ but sufficiently close to 1, where S_n is the sum of n i.i.d. random variables having the same distribution as $F(X,Y)$. By Proposition 2.1, $P(S_{l_n^2} > z_n) \sim Cl_n^{-1}\exp(-l_n^2\Lambda^*(z_n/l_n^2))$. By Taylor expansion, $l_n^2\Lambda^*(z_n/l_n^2) = \theta z_n + O((\gamma-1)^2)l_n^2$ as $\gamma \uparrow 1$. Therefore,

$$\exp(-[n/l_n]^4\xi_n) \sim \exp\left(-Cn^4(\log n)^{-1.5}e^{-z_n+O((\gamma-1)^2)l_n^2}\right) \leq e^{-Cn^{2(1-\gamma)}} \tag{4.11}$$

for $\gamma < 1$ but sufficiently close to 1, where C is a constant that may vary even in the same line. Now we turn to estimate η_n. Note that $E_{\Gamma_n,\Gamma_n}(\epsilon) \cap E_{\Gamma_n,\Gamma_n+l_n}(\epsilon) \subset E_{X,\Gamma_n}(\epsilon)$, by independence and Chebyshev's inequality

$$\begin{aligned}
\eta_n &= E^X P^Y\left(\sum_{I\in\Gamma_n} F(X_I, Y_I) \geq z_n\right)^2 1_{E_{X,\Gamma_n}} \\
&\leq E^X \exp\left(-2l_n^2 \sup_{t\in\mathbb{R}^1}\left\{t\frac{z_n}{l_n^2} - \frac{1}{l_n^2}\sum_{I\in\Gamma_n}\phi(X_I)\right\}\right) 1_{E_{X,\Gamma_n}} \\
&\leq \exp\left\{-2l_n^2\left(\{\gamma t\Lambda'(\theta) - h_X\} - \epsilon + o_n(1)\right)\right\} P(E_{X,\Gamma_n}) \\
&\leq 2\exp\left(-l_n^2\left\{2\theta\Lambda'(\theta) - h_X - 2\epsilon - 2(1-\gamma)\theta\Lambda'(\theta) + o_n(1)\right\}\right),
\end{aligned} \tag{4.12}$$

where we use Lemma 4.1.2 in the last inequality for $P(E_{X,\Gamma_n})$. By Lemma 4.2.1, there is $\epsilon_0 > 0$ such that $\epsilon_0 = \theta\Lambda'(\theta)/2 - h_X(\theta)$. Choose ϵ and γ sufficiently small so that $2\epsilon + 2(1-\gamma)\theta\Lambda'(\theta) < \epsilon_0/2$. Therefore $\eta_n = O(n^{-6-(\epsilon_0/4)})$. Similarly, we also get $\eta_n' = O(n^{-6-(\epsilon_0/4)})$. Consequently, by (4.8), (4.9), (4.10), and (4.11), we have (4.7). ■

Proof Theorem 4.2 Let $z_n = (4\log n - \log_2 n/2 + t_n)/\theta + x$,

$$E_n = \left\{\Delta \in \mathbb{N}; |\Delta - \sqrt{z_n/\Lambda'(\theta)}| \le \eta\sqrt{\log z_n}\right\}, \quad \text{and}$$
$$\Theta_n = \left\{(I,J,\Delta) \in \mathbb{N}^2 \times \mathbb{N}^2 \times E_n;\ I+\Gamma_\Delta, J+\Gamma_\Delta \subset \{1,2,\cdots,n\}^2\right\}, \quad \text{and}$$
$$W_n^1 = \max_{(I,J,\Delta)\in\Theta_n} \sum_{K\in\Gamma_\Delta} F(X_{I+K}, Y_{J+K}).$$

By arguments similar to (2.14), we have

$$P(W_n > z_n) - P(W_n^1 > z_n) \to 0 \tag{4.13}$$

for some $\eta > 0$ as $n \to \infty$. Like the proof of the lower bound in Theorem 4.1, define $G_{\Gamma,\Gamma'}(\epsilon) = G_{X,\Gamma}(\epsilon) \cap G_{Y,\Gamma'}(\epsilon)$, where

$$G_{X,\Gamma}(\epsilon) = \left\{|\frac{1}{|\Gamma|}\sum_{I\in\Gamma}\phi_X(X_I) - h_X| < \epsilon\right\},$$
$$G_{Y,\Gamma'}(\eta) = \left\{|\frac{1}{|\Gamma'|}\sum_{I\in\Gamma'}\phi_Y(Y_I) - h_Y| < \epsilon\right\}, \quad \epsilon > 0,$$

and $\phi_X(x)$ and h_X etc. are the same as in Lemma 4.2.1. In particular, denote $G_{(I,J,\Delta)}(\epsilon) = G_{(I+\Delta,J+\Delta)}(\epsilon)$. By Lemma 4.1.1, there exists positive δ that depends only on ϵ such that

$$|P(W_n^1 > z_n) - \underbrace{P(\bigcup_{(I,J,\Delta)\in\Theta_n}\{\sum_{K\in\Gamma} F(X_{I+K}, Y_{J+K}) > z_n,\ G_{(I,J,\Delta)}(\epsilon)\})}_{\mathcal{H}_n}|$$
$$\le O(\sqrt{\log n})\cdot n^4 \max_{\Delta\in E_n} e^{-(\theta\Lambda'(\theta)+\delta)\Delta^2} \le n^{-\delta'} \tag{4.14}$$

for some $\delta' > 0$. By (2.9), Lemma 2.3.3 and the same arguments as in (2.15), we obtain

$$\sum_{(I,J,\Delta)\in\Theta_n} \underbrace{P(\sum_{K\in\Gamma_\Delta} F(X_{I+K}, Y_{J+K}) > z_n)}_{S_{(I,J,\Delta)}} \sim \frac{1}{\theta}\sqrt{\frac{\Lambda'(\theta)}{2\pi\Lambda''(\theta)}}\, e^{-\theta x}. \tag{4.15}$$

Since $P(S_{(I,J,\Delta)} > z_n) > C z_n^{-1/2} e^{-\theta z_n}$ for some constant $C > 0$ by Corollary 2.1, we have from Lemma 4.1.1 that $P(S_{(I,J,\Delta)} > z_n) \sim P(S_{(I,J,\Delta)} > z_n,\ G_{(I,J,\Delta)}(\epsilon)\})$. Thus, by (4.15),

$$\lambda_n := \sum_{(I,J,\Delta)\in\Theta_n} P(S_{(I,J,\Delta)} > z_n,\ G_{(I,J,\Delta)}(\epsilon)) \sim \frac{1}{\theta}\sqrt{\frac{\Lambda'(\theta)}{2\pi\Lambda''(\theta)}}\ e^{-\theta x}. \tag{4.16}$$

Therefore, by (4.13), (4.14) and (4.16), it sufficies to show that

$$P(\mathcal{H}_n) - e^{-\lambda_n} \to 0. \tag{4.17}$$

Actually, by Lemma 2.2.3 and the Chen-Stein method, we have

$$|P(\mathcal{H}_n) - e^{-\lambda_n}| \le b_1 + b_2,$$

where

$$\begin{aligned} b_1 &:= \sum_{(I,J,\Delta)\in\Theta_n} \sum_{(I',J'\Delta')\in\mathcal{A}_n} P(S_{(I,J,\Delta)} > z_n) P(S_{I',J',\Delta'} > z_n), \\ b_2 &= \sum_{(I,J,\Delta)\in\Theta_n} \sum_{(I',J'\Delta')\in\mathcal{A}_n} P(S_{(I,J,\Delta)} > z_n,\ S_{(I',J',\Delta')} > z_n, G_{(I,J,\Delta)}(\epsilon) \cap G_{(I',J',\Delta')}(\epsilon)), \end{aligned}$$

and $\mathcal{A}_n = \{(I', J', \Delta') \in \Theta_n; \text{either } I + \Gamma_\Delta \cap I' + \Gamma_{\Delta'} \ne \emptyset \text{ or } J + \Gamma_\Delta \cap J' + \Gamma_{\Delta'} \ne \emptyset\}$. Note that $|\mathcal{A}_n| = O((\log n)^2)$ and $E\exp(\theta F(X,Y)) = 1$. Therefore

$$b_1 \le \lambda_n |\mathcal{A}_n| e^{-\theta z_n} = O\left(\frac{(\log n)^3}{n^2}\right).$$

Next, we need to show that $b_2 \to 0$. If $(I', J', \Delta') \in \mathcal{A}_{(I,J,\Delta)}$, there are two possibilities: Case (i): both two X-squares and both two Y-squares have intersections; Case (ii): only

X-squares or Y-squares have intersections. Therefore

$$\begin{aligned} b_2 \le\ & O(n^4(\log n)^3) \max_{I_1} P(S_{(I,J,\Delta)} > z_n, S_{(I',J',\Delta')} > z_n, G_{(I,J,\Delta)}(\epsilon) \cap G_{(I',J',\Delta')}(\epsilon)) \\ & + O(n^6(\log n)^3) \max_{I_2} P(S_{(I,J,\Delta)} > z_n, S_{(I',J',\Delta')} > z_n, G_{(I,J,\Delta)}(\epsilon) \cap G_{(I',J',\Delta')}(\epsilon)) \\ & + O(n^6(\log n)^3) \max_{I_3} P(S_{(I,J,\Delta)} > z_n, S_{(I',J',\Delta')} > z_n, G_{(I,J,\Delta)}(\epsilon) \cap G_{(I',J',\Delta')}(\epsilon)) \\ =:\ & A_n + B_n + C_n, \end{aligned}$$

where

$$\begin{aligned} I_1 &= \{(I', J', \Delta') \in \Theta_n; (I + \Gamma_\Delta) \cap (I' + \Gamma_{\Delta'}) \neq \emptyset,\ (J + \Gamma_\Delta) \cap (J' + \Gamma_{\Delta'}) \neq \emptyset\}, \\ I_2 &= \{(I', J', \Delta') \in \Theta_n;\ (I + \Gamma_\Delta) \cap (I' + \Gamma_{\Delta'}) \neq \emptyset,\ (J + \Gamma_\Delta) \cap (J' + \Gamma_{\Delta'}) = \emptyset\}, \end{aligned}$$

and I_3 is defined obviously by the similar way to I_2.

Now we first bound B_n. On I_2, suppose $\Delta \ge \Delta'$,

$$\begin{aligned} & P(S_{(I,J,\Delta)} > z_n, S_{(I',J',\Delta')} > z_n, G_{(I,J,\Delta)}(\epsilon) \cap G_{(I',J',\Delta')}(\epsilon)) \\ &= E^X P^Y \left(\sum_{K \in \Gamma_\Delta} F(X_{I+K}, Y_{J+K}) \ge z_n \right) P^Y \left(\sum_{K \in \Gamma_{\Delta'}} F(X_{I'+K}, Y_{J'+K}) \ge z_n \right) \\ & \quad \cdot 1(G_{X,I+\Gamma_\Delta}(\epsilon) \cap G_{X,I'+\Gamma_{\Delta'}}(\epsilon)) \\ &\le e^{-2\theta z_n} E^X \exp\left(\sum_{K \in (I+\Gamma_\Delta)} \phi_X(X_K) + \sum_{K \in (I'+\Gamma_{\Delta'})} \phi_X(X_K) \right) \\ & \quad \cdot 1(G_{X,I+\Gamma_\Delta}(\epsilon) \cap G_{X,I'+\Gamma_{\Delta'}}(\epsilon)) \\ &\le e^{-2\theta z_n + \Delta^2(h_X(\theta) + 3\epsilon)} \end{aligned} \tag{4.18}$$

for large n, where we use Chebyshev's inequality in the first inequality above, and the definition of $G_{X,I+\Gamma_\Delta}(\epsilon)$ and Lemma 4.1.2 in the second one. By Lemma 4.2.1, there exists $\epsilon_0 > 0$ such that

$$\frac{\theta \Lambda'(\theta)}{2} > \max\{h_X, h_Y\} + \epsilon_0.$$

Therefore, take ϵ small enough in (4.18), we have $2\theta z_n - \Delta^2(h_X(\theta) + \epsilon) > (3 + \epsilon_0) z_n/2$, and therefore

$$I_2 = O\left(\frac{(\log n)^3}{n^{2\epsilon_0}}\right). \tag{4.19}$$

Similarly, $I_2 = O((\log n)^3/n^{2\epsilon_0})$ too. The desired conclusion then follows from Lemma 4.2.2 below. ■

Lemma 4.2.2 $A_n \to 0$.

Proof. For any $(I, J, \Delta), (I', J', \Delta') \in \Theta_n$, there exist two squares R, R' with areas Δ^2 and Δ'^2, and two vectors K_0, L_0(possibly negative) such that

$$\begin{aligned}
&P(S_{(I,J,\Delta)} > z_n, S_{(I',J'\Delta')} > z_n) \\
&= P(\sum_{k\in R} F(X_K, Y_{K+K_0}) > z_n, \sum_{L\in R'} F(X_L, Y_{L+L_0}) > z_n).
\end{aligned}$$

By the definition of E_n, the cardinality of $R_1 := R\backslash R'$ is larger than $a_\theta\sqrt{\log n}$ for some constant $a_\theta > 0$ for large n. Consequently, the above term is equal to

$$\begin{aligned}
&E^Y P^X\left((\sum_{K\in R_1} + \sum_{K\in R\cap R'})F(X_K, Y_{K+K_0}) > z_n, \sum_{L\in R'} F(X_L, Y_{L+L_0}) > z_n\right) \\
&\le E^Y P^X\left(\sum_{K\in R\cap R'} F(X_K, Y_{K+K_0}) > z_n - \delta'|R_1|\right) \\
&\quad + E^Y\left\{P^X(\sum_{K\in R_1} F(X_K, Y_{K+K_0}) > \delta'|R_1|)\cdot P^X(\sum_{L\in R'} F(X_L, Y_{L+L_0}) > z_n)\right\} \\
&= D_n + F_n, \text{ say,}
\end{aligned}$$

where $\delta' = 2\delta$ and $\delta \in (\mu_F, 0)$ chosen in (ii) of Lemma 4.1.3. By Chebyshev's inequality, $D_n \le e^{-\theta_0 z_n + \delta|R_1|} = O(n^{-4}\exp(-b_\theta\sqrt{\log n}))$ for some positive constant b_θ which depends on θ only. So we just need to show that F_n vanishes as $n \to \infty$. Let a_n and b_n be the cardinalities of $(R_1 + K_0)\backslash(R' + L_0)$ and $(R_1 + K_0) \cap (R' + L_0)$. By Chebyshev's inequality and Fubini's theorem, we have

$$
\begin{aligned}
F_n &\le e^{-\theta z_n} e^{-\delta' |R_1| t} E^Y \prod_{K \in R_1 + K_0} E^X e^{tF(X, Y_K)} \prod_{L \in R' + L_0} E^X e^{\theta F(X, Y_L)} \\
&= e^{-\theta z_n} e^{-\delta' |R_1| t} (E e^{tF(X,Y)})^{a_n} \left(E^Y (E^X e^{tF(X,Y)}) E^X e^{\theta F(X,Y)} \right)^{b_n} \\
&= e^{-\theta z_n} e^{-\delta' |R_1| t} \Lambda_F(t)^{a_n} M(t)^{b_n} \qquad (4.20)
\end{aligned}
$$

for any $t \in (0, t_0)$, where we use the fact $E \exp(\theta F(X, Y)) = 1$ in the last equality. Note that $a_n + b_n = |R_1|$, $\Lambda_F(t) < 1$ and $M(t) < 1$ for all $t \in (0, t_0)$ by Lemma 4.1.3, so if $a_n \ge |R_1|/2$, then minimizing the term in (4.20) over all $t \in (0, t_0)$, we obtain by Lemma 4.1.3 again that

$$
F_n \le e^{-\theta z_n} e^{-|R_1| \gamma_1 / 2} = O(n^{-4} e^{-c_\theta \sqrt{\log n}}) \qquad (4.21)
$$

for some constant $c_\theta > 0$, since $|R_1| \ge a_\theta \sqrt{\log n}$. By the same arguments, we prove that inequality (4.21) is also true when $b_n \ge |R_1|/2$. Therefore $A_n \to 0$. ∎

Chapter 5

The Comparison of Two Protein Chains

5.1 Main results and their proofs

Let $\{X, X_i,\ i = 1, 2, \cdots\}$ be a sequence of i.i.d. random variables on Σ(not necessarily on $\mathbb{R}^d$) and let the same be true of $\{Y, Y_i,\ i = 1, 2, \cdots\}$, $\{U, U_i,\ i = 1, 2, \cdots\}$ and $\{V, V_i,\ i = 1, 2, \cdots\}$. We assume that both U and V have the uniform distribution on $[0, 1]^3$. For $\{U_k,\ k = 1, 2, \cdots\ n\}$ and any fixed $i \in \{1, 2, \cdots, n\}$, let $\{u_{i,p},\ p = 1, 2, \cdots, n\}$ be a permutation of $\{1, 2, \cdots, n\}$ such that

$$0 = \|U_{u_{i,1}} - U_i\| < \|U_{u_{i,2}} - U_i\| \le \|U_{u_{i,3}} - U_i\| \le \cdots \le \|U_{u_{i,n}} - U_i\|.$$

So $\{u_{i,p}\}$ is well defined with probability one. By the same way, we obtain $\{v_{i,p}\}$ from $\{V_k,\ k = 1, 2, \cdots, n\}$. For simplicity, denote $U_{i,p} = U_{u_{i,p}}$, $V_{i,p} = V_{v_{i,p}}$, $X_{i,p} = X_{u_{i,p}}$ and $Y_{i,p} = Y_{v_{i,p}}$ for all $1 \le i \le n$ and $1 \le p \le n$. Based upon $\{u_{i,p}\}$ and $\{v_{i,\ p}\}$, we construct the statistic

$$W_n = \max_{1 \le i,j,m \le n} \sum_{p=1}^{m} F(X_{i,p}, Y_{j,p}),$$

where $F(\cdot,\cdot) : \Sigma^2 \to \mathbb{R}$ is a given real-valued function and is assumed to be non-lattice throughout all this chapter. Denote $\Lambda_F(t) = \log E\exp(tF(X,Y))$ and $\Lambda_F^*(\cdot)$ the corresponding rate function of $F(X,Y)$. If condition 2.1 holds for $Z = F(X,Y)$, then by Lemma 2.2.4, there is a positive constant K depends on $F(X,Y)$ so that

$$\lim_{x\to+\infty} e^{\theta x} P\left(\max_{n\geq 1}\sum_{i=1}^{n} F(X_i,Y_i) > x\right) \to K. \tag{5.1}$$

Our main theorem in this section is as follows.

Theorem 5.1 *Suppose condition (2.1) holds for $Z = F(X,Y)$. Also, assume conditions (4.4) and (4.5) hold. Then*

$$P\left(W_n > 2\log n/\theta + x\right) \to 1 - e^{-Ke^{-\theta x}}$$

for any $x \in R^1$, where K is as in (5.1)

As shown in [11], condition (4.4) is not required in the one dimensional counterpart of Theorem 5.1 here and Theorem 4.2 in the previous chapter. We need this condition here because we don't have one dimensional structure as in [11].

Before we prove this theorem, we need some lemmas. In what follows, we will use $B(x,r)$ to stand for the box with center x and side length r. Thus the volume of such a box is $8r^3$. Let $B_{U_i,C} = B(U_i, C(\log n/n)^{1/3})$. $B_{V_i,C}$ is defined by the same way. Also,

$$\tau_{i,j} = \sum_{p=1}^{n} 1_{B_{U_i,C}}(U_p) \wedge \sum_{p=1}^{n} 1_{B_{V_i,C}}(V_p)$$

Based on these $\tau_{i,j}$, we construct the statistic

$$W_{n,1} = \max_{\substack{1\leq m\leq \tau_{i,j} \\ 1\leq i,j\leq n}} \sum_{p=1}^{m} F(X_{i,p}, Y_{j,p}).$$

For convenience, through all this section, we set $z_n = 2\log n/\theta + x$.

Lemma 5.1.1 *Suppose condition (2.1) holds for $Z = F(X,Y)$. Then, there exists $C >$*

$1 + 2/\theta\Lambda'(\theta)$ *such that*

$$P\left(W_n > z_n\right) - P\left(W_{n,1} > z_n\right) \to 0$$

for any $x \in R^1$.

Proof. Take C such that $\alpha := 4C^3 > 3/\Lambda_F^*(0)$. The difference between the above two probabilities is less than

$$\begin{aligned} & n^2 P^{X,Y}\left(\max_{1\le m\le n}\sum_{p=1}^{m} F(X_{1,p}, Y_{1,p}) > z_n\right) P^{U,V}(\tau_{1.1} \le \alpha \log n) \\ & + n^2 \max_{\alpha \log n \le m \le n} P^{X,Y}\left(\sum_{p=1}^{m} F(X_{1,p}, Y_{1,p}) > z_n\right). \end{aligned} \tag{5.2}$$

Note that $E\exp(\theta F(X,Y)) = 1$, by Doob's submartingale inequality, the first probability in (5.2) is less than $e^{-\theta x}n^{-2}$. By Crámer's large deviation theorem, the second probability of (5.2) is less than $\exp(-\alpha\Lambda_F^*(0)\log n)$. By Bernstein's inequality, we have that

$$\begin{aligned} & P(\tau_{1,1} \notin (4C^3\log n, 12C^3\log n) \\ & \le 2P\left(\left|\sum_{p=1}^{n} 1_{B_{U_1,C}}(U_p) - 8C^3\log n\right| \ge 4C^3 \log n\right) \\ & \le 4e^{-2C^3\log n/3}. \end{aligned} \tag{5.3}$$

Consequently, the sum of two terms in (5.2) is less than

$$\frac{4e^{-\theta x}}{n^{2C^3/3}} + \frac{1}{n^{\alpha\Lambda^*(0)-2}}$$

for all $n \ge 1$. ■

Define

$$\Omega_n = \{n \geq 1;\ |m - 2\log n/\theta\Lambda'(\theta)| \leq C\sqrt{(\log n)\log_2 n}\},\tag{5.4}$$

$$\Omega_{i,j} = \{m;\ 1 \leq m \leq \tau_{i,j}\} \cap \Omega_n,$$

$$W_{n,2} = \max_{\substack{m\in\Omega_{i,j}\\ 1\leq i,j\leq n}} \sum_{p=1}^{m} F(X_{i,p}, Y_{j,p}).$$

Then we have another approximation of W_n.

Lemma 5.1.2 *Suppose condition (2.1) holds for $Z = F(X, Y)$. Then there exists a constant $C > 0$ such that*

$$P\left(W_n > z_n\right) - P\left(W_{n,2} > z_n\right) \to 0$$

for any $x \in R^1$.

Proof. By Lemma 5.1.1, it is enough to prove that

$$P\left(W_{n,1} > z_n\right) - P\left(W_{n,2} > z_n\right) \to 0.$$

Actually, by (5.3), the LHS of above is less than

$$n^2 P\left(\max_{\substack{m\notin\Omega_{1,1}\\ 1\leq m\leq\tau_{1,1}}} \sum_{p=1}^{m} F(X_{1,p}, Y_{1,p}) \geq z_n\right)$$

$$\leq 12C^3(n^2\log n)E^{U,V} \max_{m\in\Omega_n^c} P^{X,Y}\left(\sum_{p=1}^{m} F(X_p, Y_p) \geq z_n\right) + n^{2-2C^3/3}.$$

By Corollary 2.1 we know that the maximum in the right hand side of the above inequality is less than $e^{-\theta x}n^{-2}\exp(-C'C^2\log_2 n)$, where $C' > 0$ is also a constant depending on $F(X, Y)$. Taking C big enough gives the desired conclusion. ■

Define

$$l_n^{\pm} = \frac{1}{2}\left(\frac{2\log n}{\theta\Lambda'(\theta)n}\right)^{1/3}\left(1 \pm \bar{C}\sqrt{\frac{\log_2 n}{\log n}}\right),$$

$$T_{i,j}^{\pm} = \sum_{p=1}^{n} 1_{B(U_i,l_n^{\pm})}(U_p) \wedge \sum_{p=1}^{n} 1_{B(V_j,l_n^{\pm})}(V_p),$$

$$\Psi_{i,j} = \left\{m;\ T_{i,j}^{-} \le m \le T_{i,j}^{+}\right\} \cap \Omega_n,$$

$$W_{n,3} = \max_{\substack{m\in\Psi_{i,j}\\ 1\le i,j\le n}} \sum_{p=1}^{m} F(X_{i,p}, Y_{j,p}).$$

In the definition of $l_n^{\pm}$ above, we use the notation $\bar{C}$ rather than C. The reason is that we want to distinguish from this $\bar{C}$ from C appearing in Ω_n.

Lemma 5.1.3 *Suppose condition (2.1) holds for $Z = F(X, Y)$. Then for any fixed C appeared in (5.4) and $\bar{C}$ in the definition of $l_n^{\pm}$,*

$$P(W_n > z_n) - P(W_{n,3} > z_n) \to 0.$$

Proof. The above difference is less than

$$n^2 P\left(\max_{\substack{T_{1,1}^{+}<m\\ m\in\Omega_n}} \sum_{p=1}^{m} F(X_{i,p}, Y_{j,p}) \ge t_n\right) + n^2 P\left(\max_{\substack{T_{1,1}^{-}>m\\ m\in\Omega_n}} \sum_{p=1}^{m} F(X_{i,p}, Y_{j,p}) \ge t_n\right).$$

By symmetry, we only need to estimate the first term of above. Actually

$$n^2 P^{X,Y}\left(\max_{\substack{T_{1,1}^{+}<m\\ m\in\Omega_n}} \sum_{p=1}^{m} F(X_{i,p}, Y_{j,p}) \ge t_n\right)$$
$$\le n^2 \sum_{m\in\Omega_n} P\left(\sum_{p=1}^{m} F(X_p, Y_p) \ge z_n\right) P^{U,V}(T_{1,1}^{+} \le m).$$

By Corollary 2.1, we have

$$P\left(\sum_{p=1}^{m} F(X_p, Y_p) \ge z_n\right) \sim \frac{C_1 e^{-\theta x}}{n^2\sqrt{\log n}} \exp(-\overbrace{C_2(m\Lambda'(\theta) - z_n)^2/\log n}^{I_n}),$$

where $C_1, C_2, \cdots$ appeared here and hereinafter are constants depend on the function $F(X,Y)$. By Bernstein's inequality,

$$P(T^+_{1,1} < m) \le 2P\left(\sum_{i=1}^{n} 1_{B(U_1, l_n^+)} < m\right)$$
$$\le 2\exp(-\underbrace{C_3(m\Lambda'(\theta) - z_n - C_4 C\sqrt{(\log n)\log_2 n}\ )^2/\log n}_{I'_n}).$$

By considering m such that $|m\Lambda'(\theta) - z_n| \ge (\log n)^{1/2}(\log_2 n)^{1/4}$, we have that $I_n + I'_n \ge C_5(\log_2 n)^{1/2}$. Therefore,

$$n^2 \sum_{m\in\Omega_n} P\left(\sum_{p=1}^{m} F(X_p, Y_p) \ge z_n\right) P(T^+_{1,1} \le m)$$
$$\le \frac{C_5 e^{-\theta x}}{\sqrt{\log n}} \sum_{m\in\Omega_n} e^{I_n + I'_n} \le 2C_5\bar{C}e^{-\theta x}\sqrt{\log_2 n}\cdot e^{-C_5\sqrt{\log_2 n}}. \quad \blacksquare$$

Let's recall some definitions in section 4.1. We defined $\phi_X(x) = \log\{E^Y \exp(\theta F(x,Y))\}$ and $h_X = E\left\{e^{\theta F(X,Y)}\left(\log E^Y e^{\theta F(X,Y)}\right)\right\}$. And $\phi_Y(y)$ and h_Y are defined similarly. Now we define

$$G_{(i,j,m)}(\epsilon) = \left\{|\frac{1}{m}\sum_{p=1}^{m}\phi_X(X_{i,p}) - h_X| < \epsilon\right\} \bigcap \left\{|\frac{1}{m}\sum_{p=1}^{m}\phi_Y(Y_{j,p}) - h_Y| < \epsilon\right\}$$

Lemma 5.1.4 *Suppose condition (2.1) holds for $Z = F(X,Y)$. Then for any $\epsilon > 0$ small enough, there exist $C > 0$ and $\bar{C} > 0$ such that*

$$b_{1,n} := \sum_{1\le i,j\le n} P^{X,Y}\left(\bigcup_{m\in\Psi_{i,j}}\left\{\sum_{p=1}^{m} F(X_{i,p}, Y_{j,p}) \ge z_n,\ G_{(i,j,m)}(\epsilon)\right\}\right) \to Ke^{-\theta x}$$

in probability for any $x \in R$, where K is as in (5.1).

Proof. It is easy to see that

$$b_{1,n} \le n^2 P\left(\max_{m\ge 1}\sum_{p=1}^{m} F(X_p, Y_p) \ge z_n\right) \to Ke^{-\theta x}$$

by (5.1). We also know that from Lemmas 2.2.9 and 2.2.10 and (5.1)

$$n^2 P^{X,Y}\left(\max_{m\in\Omega_n}\sum_{p=1}^{m}F(X_{1,p},Y_{1,p})\ge z_n\right)\to Ke^{-\theta x}\quad a.s. \tag{5.5}$$

if C is bigger than a positive constant C_0, where Ω_n is as in (5.4). By Lemma 4.1.1

$$E^{U,V}\sum_{1\le i,j\le n}P^{X,Y}\left(\bigcup_{m\in\Psi_{i,j}}\left\{\sum_{p=1}^{m}F(X_{i,p},Y_{j,p})\ge z_n,\, G_{(i,j,m)}(\epsilon)^c\right\}\right)$$
$$\le n^2|\Omega_n|\max_{m\in\Omega_n}P\left(\sum_{p=1}^{m}F(X_{1,p},Y_{1,p})\ge z_n,\, G_{(1,1,m)}(\epsilon)^c\right)=O(n^{-\delta_0})$$

for some positive δ_0 depending only on $F(X,Y)$ and ϵ appearing in the definition of $G_{(1,1,m)}(\epsilon)$. Therefore, to prove the lemma, it is enough to prove that

$$\sum_{1\le i,j\le n}P^{X,Y}\left(\max_{m\in\Psi_{i,j}}\sum_{p=1}^{m}F(X_{i,p},Y_{j,p})\ge z_n\right)\to Ke^{-\theta x}. \tag{5.6}$$

Recall the definitions of Ω_n and $T^{\pm}_{i,j}$. Take $\bar{C} > C_0 + \theta\Lambda'(\theta)C/6$. Then by Bernstein's inequality,

$$P(T^{+}_{1,1}\ge J^{+}_n \text{ or } T^{-}_{1,1}\le J^{-}_n)\le 4\exp\left\{-C_5\left(\frac{6\overline{C}}{\theta\Lambda'(\theta)}-C\right)^2\log_2 n\right\}. \tag{5.7}$$

Thus

$$-n^2P^{X,Y}\left(\max_{m\notin\Omega_n}\sum_{p=1}^{m}F(X_{1,p},Y_{1,p})\ge z_n\right) \tag{5.8}$$
$$\le n^2P^{X,Y}\left(\max_{m\in\Omega_n}\sum_{p=1}^{m}F(X_{1,p},Y_{1,p})\ge z_n\right)-b_{1,n}$$
$$\le\sum_{1\le i,j\le n}P^{X,Y}\left(\max_{m\in\Omega_n}\sum_{p=1}^{m}F(X_{i,p},Y_{j,p})\ge z_n\right)1(T_{i,j}\le J^{+}_n \text{ or } T_{i,j}\ge J^{-}_n)). \tag{5.9}$$

The term in (5.8) goes to zero because of (5.1) and (5.5). The expectation on U, V of the term in equality (5.9) vanishes because of (5.5) and (5.7). Thus, (5.6) follows. ■

Proof of Theorem 5.1. By Lemma 5.1.3, it is enough to show that

$$P\left(W_{n,3} > z_n\right) \to 1 - e^{-Ke^{-\theta x}}.$$

Define

$$A_{i,j} = \bigcup_{m \in \Psi_{i,j}} \left\{ \sum_{p=1}^{m} F(X_{i,p}, Y_{j,p}) \geq z_n, G_{(i,j,m)}(\epsilon) \right\}, \quad 1 \leq i, j \leq n.$$

Evidently, $|\Omega_n| = O(\sqrt{(\log n) \log_2 n})$. By Lemma 4.1.1, we have

$$\begin{aligned} & P\left(W_{n,3} > z_n\right) - P\left(\cup_{m \in \Psi_{i,j}} A_{i,j}\right) \\ & \leq n^2 |\Omega_n| \max_{m \in \Psi_{i,j}} P\left(\sum_{p=1}^{m} F(X_{1,p}, Y_{1,p}) \geq z_n, G_{(1,1,m)}(\epsilon)^c \right) = O(n^{-\delta_0}) \end{aligned}$$

for some constant $\delta_0 > 0$. Hereafter, when we say that some number is constant, it always mean that it depends on $F(X,Y)$ and ϵ only. Recalling $b_{1,n}$ defined in Lemma 5.1.4, we have

$$\begin{aligned} & |P\left(\cup_{m \in \Psi_{i,j}} A_{i,j}\right) - e^{-Ke^{-\theta x}}| \\ & \leq E^{U,V} \left| P^{X,Y}\left(\cup_{m \in \Psi_{i,j}} A_{i,j}\right) - e^{-b_{1,n}} \right| + E^{U,V} \left| e^{-b_{1,n}} - e^{-Ke^{-\theta x}} \right|. \end{aligned}$$

By Lemma 5.1.4 and the Dominated Convergence Theorem, the second term of the RHS in the above inequality goes to 0. It is enough to show the first term goes to 0. Actually, by Lemma 2.2.3,

$$E^{U,V} \left| P^{X,Y}\left(W_{n,3} \leq z_n\right) - e^{-b_{1,n}} \right| \leq E^{U,V} b_{2,n} + E^{U,V} b_{3,n},$$

where

$$b_{2,n} = \sum_{i,j=1}^{n} \sum_{(k,l)\in\Gamma_{i,j}} P(A_{i,j})P(A_{k,l}),\ b_{3,n} = \sum_{i,j=1}^{n} \sum_{(k,l)\in\Gamma_{i,j}} P(A_{i,j}\cap A_{k,l}), \quad \text{and}$$
$$\Gamma_{i,j} = \{(k,l)\in\{1,2,\cdots,n\}^2; B(U_k,l_n^+)\cap B(U_i,l_n^+)\neq\emptyset \text{ or } B(V_l,l_n^+)\cap B(V_j,l_n^+)\neq\emptyset\}.$$

It is easy to see by Doob's submartingale inequality that $P(A_{1,1}) \leq e^{-\theta x}n^{-2}$ *a.s.* for each $n \geq 1$. It follows that $E^{U,V}b_{2,n} \leq e^{-2\theta x}n^{-2}E^{U,V}(\#\Gamma_{1,1}) \leq e^{-2\theta x}n^{-1}E^{U,V}\#\Xi$, where $\Xi = \{1 \leq k \leq n;\ B(U_k,l_n^+)\cap B(U_1,l_n^-)\neq\emptyset\}$. But note that $\Xi = \sum_{i=2}^{n} 1\{d(U_i,U_1) < l_n^+\}$, so $E\Xi = O(\log n)$, and thus $E^{U,V}b_{2,n} = O(e^{-2\theta x}n^{-1}\log n)$.

On the other hand, by using symmetry, we see that

$$\begin{aligned}
&E^{U,V}b_{3,n} \\
&= n^2E^{U,V}\sum_{(k,l)\in\Gamma_{1,1}} P(A_{k,l}\cap A_{1,1}) \\
&\leq n^3E^{U,V}P(A_{2,1}\cap A_{1,1}) + n^3E^{U,V}P(A_{1,2}\cap A_{1,1}) + n^4P(A_{2,2}\cap A_{1,1})1\{d(U_1,U_2)\leq 2l_n^+\}.
\end{aligned}$$

We will provide two lemmas in the following to estimate the three terms of the RHS in the above inequality. With them, our proof is complete. ■

Lemma 5.1.5 $E^{U,V}P(A_{2,2}\cap A_{1,1})1\{d(U_1,U_2)\leq 2l_n^+\} = o(n^{-4})$.

Proof. Let $c_n = (\log n)^{-\delta}n^{-1/3}, \delta\in(1/3,2/3)$. Recall in the previous proof that $P(A_{2,2}\cap A_{1,1}) \leq P(A_{1,1}) \leq n^{-2}$. Also,

$$P(d(U_1,U_2)\leq 2l_n^+) \leq 8C^3(\log n)/n,\ P(d(U_1,U_2)\leq c_n) \leq 8/n(\log n)^{3\delta}.$$

Then

$$\begin{aligned}
&E^{U,V}P(A_{2,2}\cap A_{1,1})1\{d(U_1,U_2)\leq 2l_n^+\} \\
&\leq E^{U,V}P(A_{2,2}\cap A_{1,1})(1_{E_{U,1}\cap E_{V,1}} + 1_{E_{U,2}\cap E_{V,2}}) + O(n^{-4}(\log n)^{1-3\delta}), \qquad (5.10)
\end{aligned}$$

where $E_{U,1} = \{d(U_1, U_2) \in (c_n, 2l_n^+)\}$, $E_{V,1} = \{d(V_1, V_2) \in (c_n, 2l_n^+)\}$, $E_{U,2} = \{d(U_1, U_2) \leq 2l_n^+\}$ and $E_{V,2} = \{d(V_1, V_2) > 2l_n^+\}$. By independence, on $E_{U,2} \cap E_{V,2}$,

$$\begin{aligned} &P(A_{2,2} \cap A_{1,1}) = E^X P^Y(A_{2,2}) P^Y(A_{1,1}) \\ &\leq \left(4C^2(\log n)\log_2 n\right) E^{U,V} \max_{m_1, m_2 \in \Omega_n} E^X \left\{ \prod_{k=1}^{2} P^Y \left(\sum_{p=1}^{m_k} F(X_{k,p}, Y_{1,p}) \geq z_n, G_{(k,1,m_k)} \right) \right\}, \end{aligned}$$

where Ω_n is as in (5.4). Fix $m_1,\ m_2 \in \Omega_n$. By the same argument as deriving (4.19) via using condition (4.5), we have that

$$e_1 := E^X \prod_{k=1}^{2} P^Y \left(\sum_{p=1}^{m_k} F(X_{k,p}, Y_{1,p}) \geq z_n, G_{(k,1,m_k)} \right) = O(n^{-3-\gamma})$$

for some $\gamma > 0$, which depend on ϵ but not m_k, $k = 1, 2$. Consequently,

$$n^4 \cdot E^{U,V} \left\{ P(A_{2,2} \cap A_{1,1}) 1_{E_{U,2} \cap E_{V,2}} \right\} = o\left((\log n)^3 n^{-\gamma}\right). \tag{5.11}$$

Now let's estimate the rest term $E^{U,V} P(A_{2,2} \cap A_{1,1}) 1_{E_{U,1} \cap E_{V,1}}$in (5.10).

Define

$$\tau_{i,j}(r) = \sum_{p=1}^{n} 1_{B(U_i,r)}(U_p) \wedge \sum_{p=1}^{n} 1_{B(V_j,r)}(V_p).$$

Then, $\Psi_{i,j} = \{\tau_{i,j}(r);\ l_n^- \leq r \leq l_n^+,\ \tau_{i,j}(r) \in \Omega_n\}$. Conditional on U and V,

$$\begin{aligned} &P(A_{1,1} \cap A_{2,2}) \\ &\leq \bar{C}^2 (\log n)(\log_2 n) \max P(\underbrace{\sum_{p=1}^{\tau_{1,1}(r)} F(X_{1,p}, Y_{1,p}) \geq z_n,}_{A_1(r)} \underbrace{\sum_{p=1}^{\tau_{2,2}(s)} F(X_{2,p}, Y_{2,p}) \geq z_n}_{A_2(s)}), \end{aligned}$$

where the maximum is taken over all r, s so that $r, s \in (l_n^-, l_n^+)$ and $\tau_{1,1}(r), \tau_{2,2}(s) \in \Omega_n$. For any such pair r, s, w.l.o.g., we assume $r \geq s$, it is easy to check(moving parallel or vertically without changing the biggest distance $d(V_1, V_2)$ between two faces) that the volume of

$B(V_1,r)\backslash B(V_2,s)$ is greater than or equal to $4s^2 d(V_1,V_2) \geq \kappa(\log n)^{2/3-\delta}/n$ for some constant $\kappa > 0$, and the same is true for $B(U_1,r)\backslash B(U_2,s)$. Recalling the definition of $\tau_{1,1}(r)$, by symmetry, we assume w.l.o.g. that

$$\sum_{i=1}^{n} 1_{B(V_1,r)}(V_i) \leq \sum_{i=1}^{n} 1_{B(U_1,r)}(U_i). \tag{5.12}$$

By Bernstein inequality

$$P^{V_3,\cdots,V_n}\underbrace{(\sum_{i=1}^{n} 1_{B(V_1,r)\backslash B(V_2,s)}(V_i) \leq \frac{\kappa}{2}(\log n)^{2/3-\delta})}_{H_n} \leq \exp(-C(\log n)^{2/3-\delta})$$

for some $C > 0$. Define

$$\Gamma_1 = \{1 \leq p \leq n; U_{1,p} \in B(U_1,r)\backslash B(U_2,s) \text{ and } V_{1,p} \in B(V_1,r)\backslash B(V_2,s)\},$$
$$\Gamma_2 = \{1 \leq p \leq n; U_{1,p} \in B(U_1,r) \cap B(U_2,s) \text{ and } V_{1,p} \in B(V_1,r)\backslash B(V_2,s)\}.$$

Remember (5.12), on H_n^c, there are only two possibilities: either $\#\Gamma_1 \geq (\kappa/4)(\log n)^{2/3-\delta}$ or $\#\Gamma_2 \geq (\kappa/4)(\log n)^{2/3-\delta}$. Now we deal with these two cases separately.

Case 1. $\#\Gamma_1 \geq (\kappa/4)(\log n)^{2/3-\delta}$ on H_n^c.

On $E_U \cap E_{V,1} \cap H_n^c$,

$$\begin{aligned} &P^{X,Y}(A_1(r) \cap A_2(s)) \\ &\leq P^{X,Y}\left(\sum_{p\in\Omega_n\backslash\Gamma_1} + \sum_{p\in\Gamma_1} F(X_{1,p},Y_{1,p}) \geq z_n,\ \sum_{p\in\Omega_n} F(X_{2,p},Y_{2,p}) \geq z_n\right). \end{aligned}$$

Observe that $\sum_{p\in\Gamma_1} F(X_{1,p}, Y_{1,p})$ is independent of random variables $\sum_{p\in\Omega_n\setminus\Gamma_1} F(X_{1,p}, Y_{1,p})$ and $\sum_{p\in\Omega_n} F(X_{2,p}, Y_{2,p})$ conditional on $E_U \cap E_{V,1} \cap H_n^c$. Then, on $E_U \cap E_{V,1} \cap H_n^c$,

$$
\begin{aligned}
& P^{X,Y}(A_1(r) \cap A_2(s)) \\
\leq\ & P^{X,Y}\left(\sum_{p\in\Gamma_1} F(X_{1,p}, Y_{1,p}) \geq |\Gamma_1|\mu/2\right) \cdot P^{X,Y}\left(\sum_{p\in\Omega_n} F(X_{2,p}, Y_{2,p}) \geq z_n\right) \\
& + P^{X,Y}\left(\sum_{p\in\Omega_n\setminus\Gamma_1} F(X_{1,p}, Y_{1,p}) \geq z_n - |\Gamma_1|\mu_F/2\right),
\end{aligned}
$$

where $\mu_F = EF(X,Y)$ as before. On $E_U \cap E_{V,1} \cap \Gamma_1$, by Crámer's large deviation theorem and the submartingale inequality, we have that

$$
\begin{aligned}
P^{X,Y}\left(\sum_{p\in\Gamma_1} F(X_{1,p}, Y_{1,p}) \geq |\Gamma_1|\mu/2\right) &\leq e^{-C(\log n)^{2/3-\delta}\Lambda_F^*(\mu/2)}, \\
P^{X,Y}\left(\sum_{p\in\Omega_n} F(X_{2,p}, Y_{2,p}) \geq z_n\right) &\leq \frac{1}{n^2}.
\end{aligned}
$$

By Markov's inequality and using the fact that $E\exp(\theta F(X,Y)) = 1$, we obtain

$$
P^{X,Y}\left(\sum_{p\in\Omega_n\setminus\Gamma_1} F(X_{1,p}, Y_{1,p}) \geq z_n - |\Gamma_1|\mu/2\right) \leq \frac{1}{n^2} e^{-C(\log n)^{2/3-\delta}}.
$$

But

$$
\begin{aligned}
& E^{U,V} P^{X,Y}(A_{2,2} \cap A_{1,1}) 1_{E_{U,1}\cap E_{V,1}\cap H_n} \\
\leq\ & P^{X,Y}(A_{1,1}) P^{U_1,U_2,V_1,V_2}(H_n) P^{U_2,\cdots,U_n,V_2,\cdots,V_n}(E_{U,1} \cap E_{V,1}) \\
\leq\ & O\left(n^{-4}(\log n)^2 e^{-(\log n)^{2/3-\delta}}\right). \qquad (5.13)
\end{aligned}
$$

Combining all the above arguments, we finally get

$$
\begin{aligned}
& E^{U,V} P^{X,Y}(A_{2,2} \cap A_{1,1}) 1(E_U \cap E_{V,1} \cap \{\#\Gamma_1 \geq (\kappa/4)(\log n)^{2/3-\delta}\}) \\
=\ & O(n^{-4}(\log n)^2 e^{-C(\log n)^{2/3-\delta}}) \qquad (5.14)
\end{aligned}
$$

for some constant $C > 0$.

Case 2. $\#\Gamma_2 \geq (\kappa/4)(\log n)^{2/3-\delta}$ on H_n^c.

On $E_U \cap E_{V,1} \cap H_n^c$, the same arguments as those in case 1 can be applied to show that

$$\begin{aligned}&P^{X,Y}(A_1(r) \cap A_2(s))\\ &\leq P\left(\sum_{p\in\Gamma_2} F(X_{1,p}, Y_{1,p}) \geq \lambda|\Gamma_2|, \sum_{p=1}^{\tau_{2,2}(s)} F(X_{2,p}, Y_{2,p}) \geq z_n\right) 1(\tau_{2,2}(s) \in \Omega_n)\\ &+O\left(n^{-2}e^{-(\log n)^{2/3-\delta}}\right)\end{aligned}$$

for any fixed $\lambda \in (\mu_F, 0)$. Recalling the definitions of Γ_2 and $\tau_{2,2}(s)$, we know that not all the essential U_i's corresponding to $\{U_{1,p},\ p \in \Gamma_2\}$ necessarily belong to $\{V_{2,q};\ 1 \leq q \leq \tau_{2,2}(s)\}$. Define

$$\begin{aligned}A_U &= \{1 \leq i \leq n;\ U_i = U_{1,p} = U_{2,q} \text{ for some } p \in \Gamma_2 \text{ and some } 1 \leq q \leq \tau_{2,2}(s)\},\\ B_U &= \{1 \leq i \leq n;\ U_i = U_{1,p} \neq U_{2,q} \text{ for some } p \in \Gamma_2 \text{ and for all } 1 \leq q \leq \tau_{2,2}(s)\}.\end{aligned}$$

Consequently, $A_U \cap B_U = \emptyset$ and $A_U \cup B_U = \Gamma_2$. Then, by Markov's inequality,

$$\begin{aligned}T_n &:= P\left(\sum_{p\in\Gamma_2} F(X_{1,p}, Y_{1,p}) \geq \lambda|\Gamma_2|, \sum_{p=1}^{\tau_{2,2}(s)} F(X_{2,p}, Y_{2,p}) \geq z_n\right) 1(\tau_{2,2}(s) \in \Omega_n)\\ &\leq n^{-2}e^{-\lambda t|\Gamma_2|}E^X\{\prod_{p\in A_U\cup B_U} E^Y e^{tF(X_{1,p},Y)} \cdot \prod_{\substack{1\leq p\leq\tau_{2,2}(s)\\ \tau_{2,2}(s)\in\Omega_n}} E^Y e^{\theta F(X_{2,p},Y)}\}.\\ &= n^{-2}e^{-\lambda t|\Gamma_2|}\Lambda(t)^{|B_U|}M(t)^{|A_U|}\end{aligned}$$

for all $t > 0$, where $\Lambda(t) = Ee^{tF(X,Y)}$ and $M(t) = Ee^{tF(X,Y_1)+\theta F(X,Y_2)}$. If $|B_U| \geq |\Gamma_2|/2$, then by Lemma 4.1.3, $T_n \leq n^{-2}(e^{-2\lambda t}\Lambda(t))^{|B_U|}$ for all $t \in (0, t_0)$. By Lemma 4.1.3 again and choosing $\lambda = t_0/2$, $T_n \leq n^{-2}\exp(-|\Gamma_2|\gamma_1)$ for some $\gamma_1 > 0$. If $|A_U| \geq |\Gamma_2|/2$, note that $\Lambda(t) < 1$ for all $t \in (0, t_0)$, and repeat the same arguments as above to get the same bound for T_n up to another number γ_2. Consequently, $T_n = O(n^{-2}\exp(-C(\log n)^{2/3-\alpha}))$ for some

$C > 0$. Therefore, by computations similar to those in (5.13), we obtain

$$\begin{aligned} & E^{U,V} P^{X,Y}(A_{2,2} \cap A_{1,1}) 1(E_U \cap E_{V,1} \cap \{\#\Gamma_2 \geq (\kappa/4)(\log n)^{2/3-\delta}\}) \\ = \; & O(n^{-4}(\log n)^2 e^{-C(\log n)^{2/3-\delta}}), \end{aligned} \tag{5.15}$$

which together with (5.14) implies that

$$E^{U,V} P^{X,Y}(A_{2,2} \cap A_{1,1}) 1(E_U \cap E_{V,1}) = O(n^{-4}(\log n)^2 e^{-C(\log n)^{2/3-\delta}}).$$

Combining this with (5.10) and (5.11) yields the desired inequality. ■

Lemma 5.1.6 $E^{U,V} P(A_{1,1} \cap A_{1,2}) = o(n^{-3})$.

Proof. Set $c_n = (\log n)^{-h} n^{-1/3}$, $h \in (0, 1/6)$. Recall in the previous proof that $P(A_{1,1} \cap A_{1,2}) \leq P(A_{1,1}) \leq n^{-2}$. Also, $P(d(V_1, V_2) \leq c_n) \leq 8/n(\log n)^{3h}$. Therefore

$$\begin{aligned} & E^{U,V} P(A_{1,1} \cap A_{1,2}) \\ \leq \; & E^{U,V} P(A_{1,1} \cap A_{1,2}) \{I(d(V_1, V_2) \geq 2l_n) + I(c_n < d(V_1, V_2) < 2l_n)\} + 8(\log n)^{-3h}. \end{aligned}$$

Similar to the estimate of $E^{U,V} P(A_{2,2} \cap A_{1,1}) I_{E_{U,2} \cap E_{V,2}}$ in (5.11), we get

$$E^{U,V} P(A_{1,1} \cap A_{1,2}) I(d(V_1, V_2) \geq 2l_n) = o((\log n)^3 n^{-\gamma})$$

for some $\gamma > 0$. So our reasoning task is to estimate $E^{U,V} P(A_{1,1} \cap A_{1,2}) I(c_n < d(V_1, V_2) < 2l_n)$. Recall that

$$\begin{aligned} \tau_{i,j}(r) &= \sum_{p=1}^{n} 1_{B(U_i, r)}(U_p) \wedge \sum_{p=1}^{n} 1_{B(V_j, r)}(V_p), \\ \Psi_{i,j} &= \{\tau_{i,j}(r);\ l_n^- \leq r \leq l_n^+,\ \tau_{i,j}(r) \in \Omega_n\}. \end{aligned}$$

Then, if $d(V_1, V_2) \in (c_n, 2l_n)$,

$$\begin{aligned}&E^{U,V} P(A_{1,1} \cap A_{1,2})\\ &\le \bar{C}^2 (\log n)(\log_2 n) E^{U,V} \max P(\sum_{p=1}^{\tau_{1,1}(r)} F(X_{1,p}, Y_{1,p}) \ge z_n, \sum_{p=1}^{\tau_{1,2}(s)} F(X_{1,p}, Y_{2,p}) \ge z_n),\end{aligned}$$

where the maximum is taken over all r, s so that $r, s \in (l_n^-, l_n^+)$ and $\tau_{1,1}(r), \tau_{1,2}(s) \in \Omega_n$. For any such pair r, s, w.l.o.g., we assume $r \ge s$, it is easy to check(moving parallel or vertically without changing the biggest distance $d(V_1, V_2)$ between two faces) that the volume of $B(V_1, r) \backslash B(V_2, s) \ge 4s^2 d(V_1, V_2) \ge C(\log n)^{2/3-h}/n$ for some constant $C > 0$. It is also easy to check by Bernstein inequality that

(i) with probability at least $1 - \exp(-\sqrt{\log n})$,

$$\#\{3 \le p \le n;\ V_i \in B(V_1, r) \backslash B(V_2, s)\} \ge C(\log n)^{2/3-h},$$

(ii) with probability at least $1 - \exp(-\sqrt{\log n})$,

$$\#\{3 \le p \le n;\ V_i \in B(U_1, r) \backslash B(U_1, s)\} \le C\sqrt{(\log n) \log_2 n}\ .$$

The above two facts imply that there exists a set $\Gamma_3 \subset \{1, 2, \cdots, n\}$ such that $|\Gamma_3| \ge C(\log n)^{2/3-h}$, $U_{1,p} \in B(U_1, s)$ and $V_{1,p} \in B(V_1, r) \backslash B(V_2, s)$ for all $p \in \Gamma_3$. Then by using the same argument as that in getting (5.14) and (5.15) , we obtain that

$$P\left(\sum_{p=1}^{\tau_{1,1}(r)} F(X_{1,p}, Y_{1,p}) \ge z_n, \sum_{p=1}^{\tau_{1,2}(s)} F(X_{1,p}, Y_{2,p}) \ge z_n\right) = o(n^{-2} \exp\{-(\log n)^{2/3-h}\}).$$

Combining all above arguments, we have

$$E^{U,V} P(A_{1,1} \cap A_{1,2}) = O\left(n^{-3} e^{-\sqrt{\log n}} (\log n) \log_2 n\right),$$

Completing the proof. ■

Chapter 6

Future Projects and Open Problems

In this short chapter, we will comment on some results obtained in the previous chapters, and list some conjectures and future projects.

Problem 6.1 The one-dimensional setting of Theorem 2.4 originally arose from studying GI/G/I queue in [17] and was latter applied to the CUSUM method and the BLAST program. It is natural to investigate possible applications of Theorem 2.5 to queuing theory.

Problem 6.2 Let $\{X_i; i \geq 1\}$ 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 2.2 and 2.3 the asymptotic behavior of $\max_{k \geq 1} S_{f(k)}$ when $f(x) = x^p, p = 2, 3, \cdots$, and the case $F(x) = x$ is treated in Lemma 2.2.4. It is interesting to see what may happen for general $f(x)$, particularly, the case when $f(x)/x \to \infty$ and whether the fluctuation theory still works.

Problem 6.3 We impose the condition $Ee^{tX} < \infty$ for any $t \in \mathbb{R}$ in Theorems 2.2 and 2.3. It is interesting to see what happens in the case that the moment generating function does not exist but $E|X_1|^p < \infty$ for some $p \geq 1$.

Problem 6.4 Theorem 2.7 is a result on $d = 2$. It is interesting to ask what happens when $d \geq 3$. The key is to address the following number theoretic question: Denote $q_d(k) = \#\{(p_1, \cdots, p_d) \in \mathbb{N}^d;\ p_1 \cdots p_d = k\}$. What is the asymptotic behavior of $q_d(k)$?

Comment 1 We dealt with the maximum indexed by squares and rectangles in Theorems 2.4 and 2.7. One should not have much difficulty in handling general convex sets by understanding the local behavior as in Theorems 2.2 and 2.6 and then using the Chen-Stein method globally.

Problem 6.5 We obtain results for binary trees in Section 2.5. It is more interesting and potentially more useful to extend the results to Galton-Watson tree. For example, each offspring has a random number of offsprings rather than two decendents in the binary tree case.

Open problem 1 Let U_n be as in Section 3.2. It is interesting to know the limit law of U_n. The author is not able to do this at this moment. So he poses the following open problem: Suppose condition 2.1 holds. Does there exist a constant $K' > 0$ and a sequence of numbers $\{a_n; n \geq 1\}$ so that $0 \leq a_n = o(\log n)$ and

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

for all $x \in \mathbb{R}^1$?

Problem 6.6 Recall that in Theorem 4.2 we did not consider rotated squares in the definition of W_n, i.e., we just considered one way of comparison for the two given squares. There are actually four ways to compare. So what happens if rotations are allowed in the definition of W_n?

Problem 6.7 An astronomer suggested to the author that a generalized form of Theorem 4.2 can be applied to the comparison of two galaxies in the universe. The generalized setting in 4.2 has the following form: Random variables X_I are not independent, but instead are related in a Gaussian manner. However, X_I and X_J become asymptotically independent as the distance between I and J goes to infinity.

Problem 6.8 As shown in [11], condition (4.4) is not required in the one-dimensional counterpart of Theorem 4.2, so it would be nice if one can remove condition (4.4) from Theorems 4.2 and 5.1.

Bibliography

[1] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximation: The Chen-Stein method. *Annals of Probability*, 17(1):9–25, 1989.

[2] R. Arratia, L. Gordon, and M. S. Waterman. An extreme value theory for sequence matching. *Ann. Statist.*, 14:971–993, 1986.

[3] R. Arratia, L. Gordon, and M. S. Waterman. The Erdös-Rényi strong law in distribution for coin tossing and sequence matching. *Ann. Statist.*, 18:539–570, 1990.

[4] R. Arratia, P. Moris, and M. S. Waterman. Stochastic scrabble:large deviations for sequences with scores. *J. Appl. Probab.*, 25:106–119, 1988.

[5] R. Arratia and M. S. Waterman. Critical phenomena in sequence matching. *Ann. Probab.*, 13:1236–1249, 1985.

[6] R. R. Bahadur and R. Ranga Rao. On deviations of sample mean. *Ann. Math. Statist*, 31:1015–1027, 1960.

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

[8] Y. S. Chow and H. Teicher. *Probability Theory, Independence, Interchangeability, Martingales*. Springer-Verlag New York, second edition, 1988.

[9] P. Deheuvels and L. Devroye. Limit law of Erdös-Rényi-Shepp type. *Ann. Probab.*, 15:1363–1386, 1987.

[10] A. Dembo, S. Karlin, and O. Zeitouni. Critical phenomena for sequence matching with scoring. *Ann. Probab.*, 22(4):1993–2021, 1994.

[11] A. Dembo, S. Karlin, and O. Zeitouni. Limit distribution of maximal non-aligned two-sequences segmental score. *Ann. Probab.*, 22(4):2022–2039, 1994.

[12] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Springer, second edition, 1998.

[13] J. D. Deuschel and D. W. Stroock. *Large Deviations.* Academic Press Inc., 1989.

[14] R. M. Dudley. Central limit theorems for empirical measures. *Ann. Probab.*, 6:899–929, 1978.

[15] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. Wiley, New York, second edition, 1971.

[16] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers.* Oxford:Claredon, New York: Oxford University Press, 5th edition, 1979.

[17] D. Iglehart. Extreme values in the GI/G/1 queue. *Ann. Math. Statist.*, 43:627–635, 1972.

[18] S. Karlin. Statistical studies of biomolecular sequences: Score-based methods. *Phi. Trans. R. Soc. Lond. B*, 344:391–402, 1994.

[19] S. Karlin and S. F. Altschul. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. natn. Acad. Sci. U.S.A.*, 87:2264–2268, 1990.

[20] S. Karlin and F. Ost. Maximal length of common words among random letter sequences. *Ann. Probab.*, 16(535-563), 1988.

[21] S. Karlin and Z.Y. Zhu. Clusters of charged residues in protein structures. *Proc. Natl. Acad. Sci.*, 93(16):8350–8355, August 6 1996.

[22] V. V. Petrov. *Sums of Independent Random Variables.* Springer-Verlag, Berlin, 1975.

[23] D. Pollard. *Convergence of Stochastic Processes.* Springer-Verlag, New York, 1984.

[24] H. Rootzen. Maxima and exceedances of stationary markov chains. *Adv. Appl. Prob.*, 20:371–390, 1988.

[25] D. Siegmund. Approximate tail probabilities for the maxima of some random fields. *Ann. Probab.*, 16:487–501, 1988.

[26] D. Siegmund and B. Yakir. Tail probabilities for the null distribution of scaning statistics. *Bernoulli, to appear*, 1999.

[27] F. Spitzer. *Principles of Random Walk.* Van Nostrand, Princeton, 1964.

[28] V. N. Vapnik and A. Ja. Červonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab. Appl.*, 16:264–280, 1971.