Weighted Approximations for Studentized U and U-type statistics

by

Masoud Mollaalizadeh Nasari, M.Sc.

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfilment of requirements of the degree of

Doctor of Philosophy

School of Mathematics and Statistics
Ottawa-Carleton Institute of Mathematics and Statistics
Carleton University
Ottawa, Ontario, Canada
May 2010

Your file Votre référence
ISBN: 978-0-494-67881-7
Our file Notre référence
ISBN: 978-0-494-67881-7

To Abdullah Nasari and Leila Pourfard

Abstract

This dissertation deals with weak approximations of processes of non-degenerate U and U-type statistics when the variance of their respective kernel does not necessarily exist. In order to establish the results in this context, the concepts of self-normalization and Studentization are suitably introduced and used. This in turn leads to having data-based computable Studentized U and U-type statistics for estimating various parameters of interest via specific kernel functions. One immediate example of an application of such statistics is establishing asymptotic confidence intervals for the parameter of interest in a non-parametric manner. The work in this thesis was inspired by the 2008 contribution of Csörgő, Szyszkowicz and Wang [17] on studying U-statistics type processes for changepoint problems .

In Chapter 1 some general results on U-statistics are quoted or derived. Chapter 2 is dedicated to reviewing some results on weak convergence and weighted approximations in probability of processes of partial sums of i.i.d. random variables in a historical context, with an emphasis put on the case when the observations lie in the domain of attraction of the normal law. In Chapter 3 pseudo-self-normalized processes of U and U-type statistics are introduced. Employing the method of truncations, a weighted weak convergence result is derived under more relaxed conditions, in comparison to the classical integrability conditions. In chapter 4, in the spirit of the jackknife method for estimating the variance, properly Studentized sequences of U and U-type statistics processes are introduced. It is shown that these newly introduced processes coincide asymptotically in probability with the pseudo-self-normalized sequences of processes introduced in Chapter 3. Thus, they become computable, asymptotically

equivalent versions of the latter ones.

Acknowledgements

I am heartily thankful to my outstanding supervisors, Professors Csörgő and Mojirsheibani, whose encouragement, guidance and unlimited support from the initial to the final stage made writing this dissertation possible. I am grateful to them for unselfishly sharing their outstanding mathematical ideas with me, for their patience, for believing in me and for their constant help to overcome the barriers and issues during my PhD program. With no doubt Miklós Csörgő and Majid Mojirsheibani are my best teachers. I wish also to thank them very much for the numerous inspiring and joyful hours of discussion that have contributed a lot to my mathematical knowledge. Not only are Csörgő and Mojirsheibani excellent mathematicians but also they have wonderful personalities.

I am grateful to Miklós Csörgő, Barbara Szyszkowicz and Qiying Wang for sharing the preliminary version of their paper [17] with me, which inspired the work in this thesis.

I wish to sincerely thank Valerie Daley, the former Graduate Administrator of the School of Mathematics and Statistics, for her very kind help in finding me a source of financial support that made continuing my program possible.

I wish to thank Gillian Murray, the coordinator of Laboratory of Research in Statistics and Probability (LRSP) for her kindly help in preparing letters of recommendation and publishing my research results in LRSP's Technical Reports series.

I would like to thank Graduate studies of Carleton University and my supervisors for their financial support during the last five years.

My heartfelt gratitude goes to my wonderful family for their constant encourage-

ment and unlimited support without which accomplishing this would not have been possible.

Contents

Chapter 1

U-statistics; Art of Symmetry

1.1 Introduction

Ever since their introduction by Hoeffding [27], U-statistics have been of great interest and subject of intensive study. These elegant mathematical objects explore the notion of symmetry to serve of the manifold goal of Statistics that is to make inference regarding some usually unknown parameters of the underlying distribution of a random sample. Their role especially in the practical area of nonparametric statistics is undeniable.

Both exact and asymptotic behavior of U-statistics have been widely studied, and numerous contributions in both areas have been accomplished. The vast field of asymptotics of this class of statistics will certainly be witnessing more progresses from researchers to come. The present thesis in hand is devoted to investigating the asymptotic behavior of U-statistics in terms of weak convergence, on the space $D[0,1]$, with respect to weighted sup-norm metric.

In this thesis we shall specify the notion of "Lemma" to results which are quoted from books, papers and other references. The notions of "Theorem", "Proposition" and "Corollary" are used for those results derived by the author.

1.2 Definition and examples

Let $X_1, X_2, \ldots,$ be a sequence of non-degenerate real-valued independent and identically distributed (i.i.d.) random variables with distribution F. Consider the parameter (parametric function) $\theta = \theta(F)$ for which there is an unbiased estimator in terms of a Borel-measurable real-valued function $h = h(x_1, \ldots, x_m)$, which is symmetric in its arguments. That is to say, given a random sample $X_1, \ldots, X_n$ of size $n \geq m$ on

F, we have

$$\theta = \theta(F) := E(h(X_1, \ldots, X_m)) = \int_{\mathbb{R}^m} h(x_1, \ldots, x_m)\, dF(x_1) \ldots dF(x_m) < \infty.$$

The corresponding U-statistic (cf. Serfling [35] or Hoeffding [27]) is

$$\begin{aligned} U_n = U(X_1, \ldots, X_n) &= \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \ldots < i_m \leq n} h(X_{i_1}, \ldots, X_{i_m}) \\ &= [n]^{-m} \sum_{1 \leq i_1 \neq \ldots \neq i_m \leq n} h(X_{i_1}, \ldots, X_{i_m}), \end{aligned} \tag{1.2.1}$$

where $[n]^{-m} := \dfrac{(n-m)!}{n!}$.

Definition 1.2.1. *The integer $m \geq 1$, $m \leq n$, is called the* order *of the U-statistic U_n, and the symmetric estimator h of the parameter θ is called the kernel.*

Remark 1.2.1. *It is crucial to observe that the assumption of symmetry of the kernel h is not a real restriction. This is certainly true since one always can replace the nonsymmetric kernel h by*

$$\frac{1}{m!} \sum_{C(m)} h(x_1, \ldots, x_m),$$

where $C(m)$ denotes the set of all permutation of the set $\{1, \ldots, m\}$. Although the assumption of symmetry, for h, imposes no real restriction in theory, in some circumstances it is better to keep an initial nonsymmetric kernel in order to avoid technical difficulties caused by symmetrization. Such circumstances will for example be seen to emerge in Chapter 4 of this thesis.

Remark 1.2.2. *Taking $h(x) = x$ reduces a U-statistic to a partial sum. Therefore, U-statistics can be viewed as a generalization of partial sums. Although our results for U-statistics in this thesis agree with those for the corresponding partial sums, the proofs are stated, and are valid only for U-statistics of order $m \geq 2$. In fact, as it will be seen later on, we shall make use of previously obtained results for partial sums, due to Csörgő, Szyszkowicz and Wang CsSzW in [14], [15], [16], to establish our own results in Sections 3 and 4.*

1.3 Examples of U-statistics

A large number of well-known statistics are actually U-statistics. The following examples of some well known U-statistics can be found in Serfling [35] and Wilks [39].

Example 1: Let $\theta = \int x dF(x)$, the mean of the distribution function F, and consider the kernel $h(x) = x$. Then the corresponding U-statistic is of order 1 is

$$U_n = \binom{n}{1}^{-1} \sum_{i=1}^{n} X_i = \bar{X}_n,$$

the sample mean.

Example 2: Let $\theta = \int x^k dF(x)$, the k-th moment of F. Consider the kernel $h(x) = x^k$. Then, the corresponding U-statistic of order 1 is

$$U_n = \frac{1}{n} \sum_{i=1}^{n} X_i^k,$$

the sample k-th moment.

Example 3: Let $\theta = \int_{-\infty}^{a} dF(x) = P(X_1 \leq a) = F(a)$, where a is a real number. Consider the kernel $h(a) = \mathbf{1}_{(x \leq a)}$. Then $E\big(\mathbf{1}_{(x \leq a)}\big) = F(a)$, and the corresponding

U-statistic of order 1 is

$$U_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{(X_i \le a)} = F_n(a),$$

the sample random distribution function.

Example 4: Let $\theta = \int (x-\mu)^2 dF(x)$, the variance of F, where $\mu = \int x dF(x)$, and consider the kernel $h(x_1, x_2) = \frac{1}{2}(x_1 - x_2)^2$. Then the corresponding U-statistic of order 2 is

$$\begin{aligned} U_n &= \frac{2}{n(n-1)} \sum_{1 \le i_1 < i_2 \le n} \frac{1}{2}(X_{i_1} - X_{i_2})^2 \\ &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2, \end{aligned}$$

the sample variance.

Example 5: Let $\theta = (\int x dF(x))^m = \mu^m$, where m is a positive integer. Consider the kernel $h(x_1, \ldots, x_m) = \prod_{j=1}^{m} x_j$. Then, the corresponding U-statistic of order m is

$$U_n = \binom{n}{m}^{-1} \sum_{1 \le i_1 < \ldots < i_m \le n} \prod_{j=1}^{m} X_{i_j}.$$

Example 6: Let $T_n(X_1, \ldots, X_n)$ be a given statistic and m and d be two positive integers such that $m + d = n$. For any $S = \{i_1, \ldots, i_m\} \subset \{1, \ldots, m\}$, define $T_{m,S} = T_m(X_{i_1}, \ldots, X_{i_m})$ which is the statistic T_n computed after X_i, $i \in \{1, \ldots, n\} - S$ are deleted from the original random sample $X_1, \ldots, X_n$. Considering now the kernel

$$h(X_1, \ldots, X_m) = \frac{m}{d}[T_m(X_1, \ldots, X_m) - T_n(X_1, \ldots, X_n)]^2,$$

the corresponding U-statistic is

$$U_n = \binom{n}{m}^{-1} \sum_{1 \le i_1 < \ldots < i_m \le n} \frac{m}{d} [T_m(X_{i_1}, \ldots, X_{i_m}) - T_n(X_{i_1}, \ldots, X_{i_n})]^2.$$

The latter U-statistic is known as the deleted jackknife variance estimator.

Fisher's k-statistic for estimation of cumulants, Kendal τ, Gini's mean difference and Wilcoxon one-sample statistic are other examples of U-statistics that can be found, for example, in [35].

1.4 U-statistic versus kernel

Both the kernel h and its corresponding U-statistic U_n are unbiased estimators of the parameter θ. Considering that both U_n and h are unbiased estimators of θ, what is the reason of choosing U_n (instead of h) to work with? To answer this question we first note that, for a kernel $h(x_1, \ldots, x_m)$ and a random sample $X_1, \ldots, X_n$, $n \ge m$, the corresponding U-statistic U_n can be represented in terms of conditional expectation of the kernel h on the order statistics (cf., Serfling [35], Section 5.1.4), i.e.,

$$U_n = E(h(X_1, \ldots, X_m) \mid X_{(1)}, \ldots, X_{(n)}),$$

where $X_{(1)}, \ldots, X_{(n)}$ are the order statistics of the sample $X_1, \ldots, X_n$. The following lemma (cf., Serfling [35], Section 5.1.4) provides an answer to our initial question of this section.

Lemma 1.4.1. . *Let $S = S(x_1, \ldots, x_n)$ be an unbiased estimator of $\theta = \theta(F)$ based* *on a random sample* $X_1, \ldots, X_n$ *from the distribution* F. *Also let*

$$S^*(x_1, \ldots, x_n) = \frac{1}{n!} \sum_{1 \le \sigma(1) < \ldots < \sigma(n) \le n} S(x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

Then the corresponding U-statistic, i.e. $U_n = E(S^*|X_{(1)}, \ldots, X_{(n)})$ *(cf. Remark 1.2.1), is also unbiased and*

$$Var_F(U_n) \le Var_F(S),$$

with equality if and only if $P_F(U_n = S) = 1$.

1.5 Variance of a U-statistics

Consider a U-statistic U_n of order $m \le n$, with mean θ, for which the kernel h is square integrable, i.e., $Eh^2(X_1, \ldots, X_m) < \infty$. Since U_n is a sum of a number of terms, the calculation of its variance yields cross-product terms of the following form:

$$\xi_c = E[(h(X_1, \ldots, X_c, X_{c+1}, \ldots, X_m) - \theta)(h(X_1, \ldots, X_c, X_{c+1}, \ldots, X_m) - \theta)].$$

It is not hard to see that the number of times that each ξ_c, $c = 1, \ldots, m$, appears is $\binom{n}{m}\binom{m}{c}\binom{n-m}{m-c}$.

In order to simplify ξ_c, (for future use), we now define the projections $\tilde{h}_c$, $c = 1, \ldots, m$, as follows.

$$\tilde{h}_c(x_1, \ldots, x_c) := E(h(X_1, \ldots, X_c, X_{c+1}, \ldots, X_m) - \theta | X_1 = x_1, \ldots, X_c = x_c).$$

Remark 1.5.1. *One should note that due to symmetry of the kernel* h, *stating the* *latter definition with only the first c elements does not violate the generality of it. It can be readily seen that* $E(\tilde{h}_c(X_1,\ldots,X_c)) = 0$ *and* $\tilde{h}_m = h-\theta$. *We note that the most important projection to us, in this thesis, is the one of the form that conditions on one observation only, i.e.,* $\tilde{h}_1(.)$. *These projections form the sequence* $\tilde{h}_1(X_1),\ldots,\tilde{h}_1(X_n)$ *of i.i.d. random variables. As we will see, this sequence will play an important role in deriving weak convergence of the underlying non-degenerate (cf., Section 1.6) U-statistic* U_n, *for which we have* $E(\tilde{h}_1(X_1))^2 > 0$.

Now in order to further simplify ξ_c, an application of tower property leads to

$$\xi_c = E\{E[(h(X_1,\ldots,X_c,X_{c+1},\ldots,X_m)-\theta)$$

$$(h(X_1,\ldots,X_c,X_{c+1},\ldots,X_m)-\theta)]\big|X_1,\ldots,X_c\}$$

$$= E(\tilde{h}_c(X_1,\ldots,X_c))^2$$

$$= Var(\tilde{h}_c(X_1,\ldots,X_c)).$$

The variance of the U-statistic U_n of order $m \le n$, is given in the following Lemma 1.5.1 (cf. Serfling [35], Section 5.2.1).

Lemma 1.5.1. *The variance of* U_n *is given by*

$$Var(U_n) = \binom{n}{m}^{-1} \sum_{c=1}^{m} \binom{n}{m}\binom{n-m}{m-c}\xi_c$$

and satisfies

(i) $\dfrac{m^2}{n}\xi_1 \le Var(U_n) \le \dfrac{m}{n}\xi_m;$

(ii) $(n+1)Var(U_{n+1}) \le Var(U_n);$

(iii) $Var(U_n) = \frac{m^2\xi_1}{n} + O(n^{-2}), \quad n \to \infty.$

1.6 The concept of degeneracy; An extension

In order to avoid some future difficulties in this thesis, we shall extend the usual definition of the concept of degeneracy of U-statistics U_n of order $m \le n$. This extension will include both U-statistics, with a symmetric kernel, and sums of the form

$$\sum_{i_1 \neq \ldots \neq i_m} L(X_{i_1}, \ldots, X_{i_m}),$$

where L is not necessarily symmetric. We shall, simply, call them statistics.

In order to state our definition of degeneracy, first observe that for a Borel-measurable function $L : \mathbb{R}^m \to \mathbb{R}$, $m \ge 1$, with mean $\mu = E(L(X_1, \ldots, X_m))$ and $c \le m$, we have

$$\begin{aligned}
Var & \ \big(E(L(X_1, \ldots, X_m) - \mu | X_1, \ldots, X_c)\big) \\
= & \ E\big(E^2(L(X_1, \ldots, X_m) - \mu | X_1, \ldots, X_c)\big) \\
= & \ E\big\{E^2\big(E(L(X_1, \ldots, X_m) - \mu | X_1, \ldots, X_{c+1}) \ |X_1, \ldots, X_c\big)\big\} \\
\le & \ E\big\{E\big(E^2(L(X_1, \ldots, X_m) - \mu | X_1, \ldots, X_{c+1}) \ |X_1, \ldots, X_c\big)\big\} \\
= & \ E\big(E^2(L(X_1, \ldots, X_m) - \mu | X_1, \ldots, X_{c+1})\big) \\
= & \ Var\ \big(E(L(X_1, \ldots, X_m) - \mu | X_1, \ldots, X_{c+1})\big).
\end{aligned}$$

The above argument, which holds when L is not necessarily symmetric, implies

$$0 \le Var\big(E(L(X_1, \ldots, X_m) - \mu | X_1)\big)$$

$$\leq \ldots \leq Var\big(E(L(X_1,\ldots,X_m)-\mu|X_1,\ldots,X_{m-1})\big) \leq Var\big(L(X_1,\ldots,X_m)\big).$$

Definition 1.6.1. *The function $L:\mathbb{R}^m \to \mathbb{R}$, $m \geq 1$, which is Borel-measurable with mean $\mu = E(L(X_1,\ldots,X_m))$, is said to be degenerate of rank r, $1 \leq r \leq m$, if r is the first integer for which we have*

$$Var\big(E(L(X_1,\ldots,X_m)-\mu|X_1)\big)$$

$$= \ldots = Var\big(E(L(X_1,\ldots,X_m)-\mu|X_1,\ldots,X_{r-1})\big) = 0 \ \ and$$

$$Var\big(E(L(X_1,\ldots,X_m)-\mu|X_1,\ldots,X_r)\big) > 0.$$

Remark 1.6.1. *The sum $\sum_{i_1\neq\ldots\neq i_m} L(X_{i_1},\ldots,X_{i_m})$ is called degenerate of rank r when its summand L is so.*

Remark 1.6.2. *When $r=1$, the corresponding statistic is called non-degenerate and when $r=m$, it is called a complete degenerate one.*

Remark 1.6.3. *It is crucial to observe that our definition of degeneracy extends the usual concept of degeneracy for random variables. Moreover, it can be readily seen that*

$$Var\big(E(L(X_1,\ldots,X_m)-\mu|X_1,\ldots,X_c)\big) = 0 \text{ if and only if}$$

$$E(L(X_1,\ldots,X_m)-\mu|X_1,\ldots,X_c) = 0 \ a.s.$$

In virtue of the latter relation, the case $r=m$ (complete degeneracy) is equivalent to

$$E(L(X_1,\ldots,X_m)-\mu|X_1) = \ldots = E(L(X_1,\ldots,X_m)-\mu|X_1,\ldots,X_{m-1}) = 0 \ a.s.$$

We shall use the latter definition of complete degeneracy in our proofs (cf. Proposi-

tions 1.8.1, 1.8.2 and 1.8.3).

Remark 1.6.4. *When L is symmetric, our definition of degeneracy coincides with the one defined by Borovskikh in [4].*

1.7 Decomposition(Reduction)

It is clear that except in the case of a kernel of order $m = 1$, U-statistics are not sums of independent random variables. Therefore the theory of partial sums can not be applied directly to this class of statistics. To overcome this barrier on our way to establishing weak convergence results for U-statistics in our Chapters 3 and 4, we introduce and use a decomposition of statistics of the form

$$\sum_{1\le i_1\neq\ldots\neq i_m\le n} L(X_{i_1},\ldots,X_{i_m}).$$

This decomposition, which remains valid when L is symmetric, works by adding and subtracting appropriate terms as follows.

Let $L(x_1,\ldots,x_m)$ with mean μ be degenerate of rank r, $r\le m$. Now we write

$$\sum_{1\le i_1\neq\ldots\neq i_m\le n} \left(L(X_{i_1},\ldots,X_{i_m}) - \mu\right)$$

$$= \sum_{1\le i_1\neq\ldots\neq i_m\le n} \Big\{ \sum_{d=r}^{m} (-1)^{m-d} \sum_{1\le j_1<\ldots<j_d\le m} E(L(X_{i_1},\ldots,X_{i_m}) - \mu \mid X_{i_{j_1}},\ldots X_{i_{j_d}})$$

$$+\sum_{c=r}^{m-1}\sum_{1\le k_1<\ldots<k_c\le m}\sum_{d=1}^{c}(-1)^{c-d}\sum_{1\le j_1<\ldots<j_d\le c} E(L(X_{i_1},\ldots,X_{i_m}) - \mu| \, X_{i_{k_{j_1}}},\ldots X_{i_{k_{j_d}}})\Big\} \tag{1.7.1}$$

Remark 1.7.1. *The above decomposition will coincide with the so called Hoeffding* *(cf. for example [35] or [4]) decomposition of U-statistics when L is symmetric in its arguments.*

Remark 1.7.2. *When $r = 1$, (1.7.1) leads to partial sums of i.i.d. random variables. These partial sums in turn will determine the limiting behavior of the underlying statistics on the left hand side of equality as in (1.7.1).*

Remark 1.7.3. *Although in this thesis we shall only consider non-degenerate U-statistics, i.e., the case $r = 1$ for which by the definition of degeneracy we have $E(\tilde{h}_1(X_1))^2 > 0$, in light of (1.7.1), we shall present our proofs based on complete degenerate statistics, i.e., when $r = m$.*

1.8 Some properties of complete degenerate statistics

At this stage we shall state and prove some results concerning complete degenerate statistics for future use. The following result gives an upper bound for the second moment of complete degenerate statistics as follows.

Proposition 1.8.1. *If $L : \mathbb{R}^m \to \mathbb{R}$, $m \geq 2$, is complete degenerate with mean $\mu = EL(X_1, \ldots, X_m)$ and $EL^2(X_1, \ldots, X_m) < \infty$, then,*

$$E\big([n]^{-m} \sum_{1 \leq i_1 \neq \ldots \neq i_m \leq n} (L(X_{i_1}, \ldots, X_{i_m}) - \mu)\ \big)^2 \leq\ [n]^{-m}\ E\ (L(X_1, \ldots, X_m) - \mu)^2.$$

Proof of Proposition 1.8.1

Let $\hat{L}_{1\ldots m} := \frac{1}{m!}\sum_{C_m} L_{\sigma_1\ldots\sigma_m}$, where $L_{\sigma_1\ldots\sigma_m} := L(X_{\sigma_1},\ldots,X_{\sigma_m})$ and C_m denotes the set of all permutations $\sigma_1,\ldots,\sigma_m$ of $1,\ldots,m$. It is clear that

$$\sum_{1\le i_1\ne\ldots\ne i_m\le n}(\hat{L}_{i_1,\ldots,i_m}-\mu)=\sum_{1\le i_1\ne\ldots\ne i_m\le n}(L_{i_1,\ldots,i_m}-\mu).$$

Now write

$$E\ (\ [n]^{-m}\textstyle\sum_{1\le i_1\ne\ldots\ne i_m\le n}(\hat{L}_{i_1,\ldots,i_m}-\mu)\ )^2$$

$$=([n]^{-m})^2\textstyle\sum_{1\le i_1\ne\ldots\ne i_m\le n}E(\hat{L}_{i_1,\ldots,i_m}-\mu)^2$$

$$+([n]^{-m})^2\textstyle\sum_{j=1}^{m-1}\sum_{1\le i_1\ne\ldots\ne i_{2m-j}\le n}E[\ (\hat{L}_{i_1,\ldots,i_j,i_{j+1},\ldots,i_m}-\mu)$$

$$\times\ (\hat{L}_{i_1,\ldots,i_j,i_{m+1},\ldots,i_{2m-j}}-\mu)]$$

$$+([n]^{-m})^2\textstyle\sum_{1\le i_1\ne\ldots\ne i_{2m}\le n}E\ (\ (\hat{L}_{i_1,\ldots,i_m}-\mu)\ (\hat{L}_{i_{m+1},\ldots,i_{2m}}-\mu)\ )$$

$$=\ [n]^{-m}\ E\ (\hat{L}_{1,\ldots,m}-\mu)^2$$

$$+([n]^{-m})^2\textstyle\sum_{j=1}^{m-1}\sum_{1\le i_1\ne\ldots\ne i_{2m-j}\le n}E\{E[\hat{L}_{i_1,\ldots,i_j,i_{j+1},\ldots,i_m}-\mu\mid X_{i_1},\ldots,X_{i_j}]$$

$$\times\ E[\hat{L}_{i_1,\ldots,i_j,i_{m+1},\ldots,i_{2m-j}}-\mu|\ X_{i_1},\ldots,X_{i_j}]\}$$

$$+([n]^{-m})^2\textstyle\sum_{1\le i_1\ne\ldots\ne i_{2m}\le n}E(\hat{L}_{i_1,\ldots,i_m}-\mu)\ E(\hat{L}_{i_{m+1},\ldots,i_{2m}}-\mu)$$

$$=\ [n]^{-m}\ E\ (\hat{L}_{1,\ldots,m}-\mu)^2$$

$$\le[n]^{-m}\ \tfrac{m!}{(m!)^2}\textstyle\sum_{C_m}E\ (L_{\sigma(1),\ldots,\sigma(m)}-\mu)^2.$$

$$=[n]^{-m}E\ (L_{1,\ldots,m}-\mu)^2.$$

The last equality above is true provided that $EL^2_{\sigma_1,\ldots,\sigma_m}=EL^2_{1\ldots m}$. $\square$

It is easy to observe that when L is symmetric in its arguments, the inequality in Proposition 1.8.1 becomes equality.

A companion of Proposition 1.8.1 when the summand L is symmetric and depends on the index i_m, where $1 \leq i_1 < \ldots < i_m \leq n$, reads as follows.

Proposition 1.8.2. *If $L_{i_m} : \mathbb{R}^m \to \mathbb{R}$, $m \geq 2$, is symmetric, centered and complete degenerate such that $EL^2_{i_m}(X_{i_1}, \ldots, X_{i_m}) < \infty$, then*

$$E\Big(\sum_{1 \leq i_1 < \ldots < i_m \leq n} L_{i_m}(X_{i_1}, \ldots, X_{i_m})\Big)^2 \leq \sum_{i_m=m}^{n} \binom{i_m}{m-1} E\big(L^2_{i_m}(X_1, \ldots, X_m)\big).$$

Proof of Proposition 1.8.2

First let

$$\sum_{i_m=m}^{n} \sum_{i_{m-1}=m-1}^{i_m-1} \ldots \sum_{i_1=1}^{i_2-1} L_{i_m}(X_{i_1}, \ldots, X_{i_m}) := \sum_{i_m=m}^{n} Y_{i_m},$$

and for $i_m \neq i'_m$ write

$$E\Big(\sum_{i_m=m}^{n} Y_{i_m}\Big)^2 = \sum_{i_m=m}^{n} E(Y_{i_m})^2 + \sum_{i_m=m}^{n} \sum_{i'_m=m}^{n} E\big(Y_{i_m} Y_{i'_m}\big). \tag{1.8.1}$$

We now show that $E\big(Y_{i_m} Y_{i'_m}\big) = 0$. To do so, assume that $i_m < i'_m$ and write

$$\begin{aligned} E\big(Y_{i_m} Y_{i'_m}\big) &= E\big[E\big(Y_{i_m} Y_{i'_m}\big) | X_1, \ldots, X_{i_m}\big] \\ &= E\big[Y_{i_m} E\big(Y_{i'_m} | X_1, \ldots, X_{i_m}\big)\big]. \end{aligned} \tag{1.8.2}$$

To deal with the conditional expectation $E\big(Y_{i'_m} | X_1, \ldots, X_{i_m}\big)$ in the latter relation, first recall that

$$Y_{i'_m} = \sum_{i'_{m-1}=m-1}^{i'_m} \ldots \sum_{i'_1=1}^{i'_2-1} L_{i'_m}(X_{i'_1}, \ldots, X_{i'_{m-1}}, X_{i'_m}).$$

Now consider the case when $i'_m - i_m < m-1$, then $\{i'_1, \ldots, i'_{m-1}\} \cap \{i_1, \ldots, i_m\} \neq \phi$. Hence, the complete degeneracy of $L_{i'_m}$ implies that

$$E(L_{i'_m}(X_{i'_1}, \ldots, X_{i'_m})|X_1, \ldots, X_{i_m}) = 0 \; a.s.,$$

and this implies that

$$E\big(Y_{i'_m}|X_1, \ldots, X_{i_m}\big) = 0.$$

The other case is when $i'_m - i_m \geq m-1$. In this case $Y_{i'_m}$ can be written as

$$Y_{i'_m} = \sum_{I_1} L_{i'_m}(X_{i'_1}, \ldots, X_{i'_m}) + \sum_{I_2} L_{i'_m}(X_{i'_1}, \ldots, X_{i'_m}),$$

where $I_1 = \{i'_{m-1}, \ldots, i'_1\} \subset \{i_1, \ldots, i_m\}$ and $I_2 = \{i'_{m-1}, \ldots, i'_1\} \subset \{i_{m+1}, \ldots, i'_m\}$. The same argument as that of the previous case implies that on I_1 we have

$$E(L_{i'_m}(X_{i'_1}, \ldots, X_{i'_m})|X_1, \ldots, X_{i_m}) = 0 \; a.s.$$

Observe that on I_2, $\{i'_1, \ldots, i'_{m-1}\} \cap \{i_1, \ldots, i_m\} = \phi$. Therefore,

$$E(L_{i'_m}(X_{i'_1}, \ldots, X_{i'_m})|X_1, \ldots, X_{i_m}) = E(L_{i'_m}(X_{i'_1}, \ldots, X_{i'_m})) = 0.$$

The last relation is true since $L_{i'_m}$ is centered. Thus,

$$E\big(Y_{i'_m}|X_1, \ldots, X_{i_m}\big) = 0.$$

The last relation together with (1.8.2) and (1.8.1) implies that

$$E\Big(\sum_{1\le i_1<\ldots<i_m\le n} L_{i_m}(X_{i_1},\ldots,X_{i_m})\Big)^2 = \sum_{i_m=m}^{n} E\Big(\sum_{\substack{1\le i_1<\ldots<i_{m-1}\\ i_{m-1}<i_m}} L_{i_m}(X_{i_1},\ldots,X_{i_m})\Big)^2.$$

We now proceed via a similar argument as in the proof of Proposition 1.8.1, for estimating $E\big(\sum_{\substack{1\le i_1<\ldots<i_{m-1}\\ i_{m-1}<i_m}} L_{i_m}(X_{i_1},\ldots,X_{i_{m-1}},X_{i_m})\big)^2$ as follows.

$$E\Big(\sum_{\substack{1\le i_1<\ldots<i_{m-1}\\ i_{m-1}<i_m}} L_{i_m}(X_{i_1},\ldots,X_{i_{m-1}},X_{i_m})\Big)^2$$

$$= \sum_{\substack{1\le i_1<\ldots<i_{m-1}\\ i_{m-1}<i_m}} E\big(L^2_{i_m}(X_{i_1},\ldots,X_{i_{m-1}},X_{i_m})\big)$$

$$+\sum_{j=1}^{m-2}\sum_{1\le i_1<\ldots<i_{2m-j}\le n} E\big[L_{i_m}(X_{i_1},\ldots,X_{i_j},X_{i_{j+1}},\ldots,X_{i_{m-1}},X_{i_m})|X_{i_1},\ldots,X_{i_j},X_{i_m}\big]$$
$$\times E\big[L_{i_m}(X_{i_1},\ldots,X_{i_j},X_{i_{m+1}},\ldots,X_{i_{2m-j}},X_{i_m})|X_{i_1},\ldots,X_{i_j},X_{i_m}\big]$$

$$+\sum_{1\le i_1<\ldots<i_{2m-2}\le n} E\big[L_{i_m}(X_{i_1},\ldots,X_{i_{m-1}},X_{i_m})|X_{i_m}\big]\ E\big[L_{i_m}(X_{i_{m+1}},\ldots,X_{i_{2m-2}},X_{i_m})|X_{i_m}\big]$$

$$= \sum_{\substack{1\le i_1<\ldots<i_{m-1}\\ i_{m-1}<i_m}} E\big(L^2_{i_m}(X_{i_1},\ldots,X_{i_m})\big)$$

$$\le \binom{i_m}{m-1} E\big(L^2_{i_m}(X_1,\ldots,X_m)\big).$$

Now the proof of Proposition 1.8.2 is complete. □

The following Proposition 1.8.3 deals with the martingale property of degenerate U-statistics and statistics of the form of a sum over non symmetric summands.

Proposition 1.8.3. *Let* $\mathfrak{F}_n := \sigma(X_1,\ldots,X_n)$ *be the smallest* σ*-field generated by the i.i.d. sequence* $X_1,\ldots,X_n$*. If for* $m\ge 2$*,* $L:\mathbb{R}^m\to\mathbb{R}$*,* $L_n:\mathbb{R}^m\to\mathbb{R}$ *and*

$L_{i_m} : \mathbb{R}^m \to \mathbb{R}$ *are centered complete degenerate functions, then, for* $m \leq n$*, we have that*

$$(a)\ \{ \sum_{1\leq i_1 \neq \ldots \neq i_m \leq n} L(X_{i_1}, \ldots, X_{i_m}), \mathfrak{F}_{\mathfrak{n}}\} \text{ is a martingale,}$$

$$(b)\ \{ \sum_{1\leq i_1 \neq \ldots \neq i_m \leq K} L_n(X_{i_1}, \ldots, X_{i_m}), \mathfrak{F}_K, \} \text{ is a martingale,}$$

where $m \leq K \leq n$, $\mathfrak{F}_K := \sigma(X_1, \ldots, X_K)$ *and*

$$(c)\ \{ \sum_{1\leq i_1 \neq \ldots \neq i_m \leq n} L_{i_m}(X_{i_1}, \ldots, X_{i_m}), \mathfrak{F}_{\mathfrak{n}}\} \text{ is a martingale.}$$

Proof of Proposition 1.8.3

To prove part (a), write

$$\begin{aligned} E \quad (& \sum_{1\leq i_1 \neq \ldots \neq i_m \leq n+1} L(X_{i_1}, \ldots, X_{i_m}) | X_1, \ldots, X_n) \\ = & \sum_{1\leq i_1 \neq \ldots \neq i_m \leq n} L(X_{i_1}, \ldots, X_{i_m}) \\ + & \sum_{\substack{1\leq i_2 \neq \ldots \neq i_m \leq n \\ i_1 = n+1}} E(L(X_{n+1}, X_{i_2}, \ldots, X_{i_m}) \big| X_1, \ldots, X_n) \\ + & \cdots + \sum_{\substack{1\leq i_1 \neq \ldots \neq i_{m-1} \leq n \\ i_m = n+1}} E(L(X_{i_1}, \ldots, X_{i_{m-1}}, X_{n+1}) \big| X_1, \ldots, X_n) \\ = & \sum_{1\leq i_1 \neq \ldots \neq i_m \leq n} L(X_{i_1}, \ldots, X_{i_m}). \end{aligned}$$

When L is symmetric in its arguments, then Part (a) coincides with the well known (forward) martingale property of $U-$statistics, and the proof can be found, for example, in [4] or [35].

Part (b) follows from the following simple argument.

$$E\Big(\sum_{1\le i_1 \neq \ldots \neq i_m \le K+1} L_n(X_{i_1}, \ldots, X_{i_m}) \Big| \, X_1, \ldots, X_K \Big)$$

$$= \sum_{1\le i_1 \neq \ldots \neq i_m \le K} L_n(X_{i_1}, \ldots, X_{i_m})$$

$$+ \sum_{\substack{1\le i_2 \neq \ldots \neq i_m \le K \\ i_1 = K+1}} E\big(L_n(X_{K+1}, X_{i_2}, \ldots, X_{i_m}) \big| \, X_1, \ldots, X_K\big)$$

$$+ \cdots + \sum_{\substack{1\le i_1 \neq \ldots \neq i_{m-1} \le K \\ i_m = K+1}} E\big(L_n(X_{i_1}, \ldots, X_{i_{m-1}}, X_{K+1}) \big| \, X_1, \ldots, X_K\big)$$

$$= \sum_{1\le i_1 \neq \ldots \neq i_m \le K} L_n(X_{i_1}, \ldots, X_{i_m}).$$

A similar argument yields

$$E\Big(\sum_{1\le i_1 \neq \ldots \neq i_m \le n+1} L_{i_m}(X_{i_1}, \ldots, X_{i_m}) \Big| \, X_1, \ldots, X_n \Big)$$

$$= \sum_{1\le i_1 \neq \ldots \neq i_m \le n} L_{i_m}(X_{i_1}, \ldots, X_{i_m})$$

$$+ \sum_{\substack{1\le i_2 \neq \ldots \neq i_m \le n \\ i_1 = n+1}} E\big(L_{i_m}(X_{n+1}, X_{i_2}, \ldots, X_{i_m}) \big| \, X_1, \ldots, X_n\big)$$

$$+ \cdots + \sum_{\substack{1\le i_2 \neq \ldots \neq i_m \le n \\ i_{m-1} = n+1}} E\big(L_{i_m}(X_{i_1}, \ldots, X_{n+1}, \ldots, X_{i_m}) \big| \, X_1, \ldots, X_n\big)$$

$$+ \sum_{\substack{1\le i_1 \neq \ldots \neq i_{m-1} \le n \\ i_m = n+1}} E\big(L_{n+1}(X_{i_1}, \ldots, X_{i_{m-1}}, X_{n+1}) \big| \, X_1, \ldots, X_n\big)$$

$$= \sum_{1\le i_1 \neq \ldots \neq i_m \le n} L_{i_m}(X_{i_1}, \ldots, X_{i_m}).$$

Now the proof of Proposition 1.8.3 is complete. □

Remark 1.8.1. *Proposition 1.8.3 is more general than what we will need it for in our dealing with truncation arguments in the future. To demonstrate the role of Proposition 1.8.3 in our proofs, consider, for example,*

$$\begin{aligned}
L_n(X_{i_1},X_{i_2}) &:= h(X_{i_1},X_{i_2})\mathbf{1}_{(|h(X_{i_1},X_{i_2})|\le n^{\frac{3}{2}})} - E(h(X_{i_1},X_{i_2})\mathbf{1}_{(|h(X_{i_1},X_{i_2})|\le n^{\frac{3}{2}})}) \\
&- E\big(h(X_{i_1},X_{i_2})\mathbf{1}_{(|h(X_{i_1},X_{i_2})|\le n^{\frac{3}{2}})} - E(h(X_{i_1},X_{i_2})\mathbf{1}_{(|h(X_{i_1},X_{i_2})|\le n^{\frac{3}{2}})})\big|X_{i_1}\big) \\
&- E\big(h(X_{i_1},X_{i_2})\mathbf{1}_{(|h(X_{i_1},X_{i_2})|\le n^{\frac{3}{2}})} - E(h(X_{i_1},X_{i_2})\mathbf{1}_{(|h(X_{i_1},X_{i_2})|\le n^{\frac{3}{2}})})\big|X_{i_2}\big)
\end{aligned}$$

and

$$\begin{aligned}
L_{i_2}(X_{i_1},X_{i_2}) &:= h(X_{i_1},X_{i_2})\mathbf{1}_{(|h(X_{i_1},X_{i_2})|\le i_2^{\frac{3}{2}})} - E(h(X_{i_1},X_{i_2})\mathbf{1}_{(|h(X_{i_1},X_{i_2})|\le i_2^{\frac{3}{2}})}) \\
&- E\big(h(X_{i_1},X_{i_2})\mathbf{1}_{(|h(X_{i_1},X_{i_2})|\le i_2^{\frac{3}{2}})} - E(h(X_{i_1},X_{i_2})\mathbf{1}_{(|h(X_{i_1},X_{i_2})|\le i_2^{\frac{3}{2}})})\big|X_{i_1}\big) \\
&- E\big(h(X_{i_1},X_{i_2})\mathbf{1}_{(|h(X_{i_1},X_{i_2})|\le i_2^{\frac{3}{2}})} - E(h(X_{i_1},X_{i_2})\mathbf{1}_{(|h(X_{i_1},X_{i_2})|\le i_2^{\frac{3}{2}})})\big|X_{i_2}\big),
\end{aligned}$$

where h is real valued and integrable. Now, since L_n and L_{i_2} are complete degenerate, Parts (b) and (c) of Proposition 1.8.3 imply that

$$\{\sum_{1\le i_1\neq i_2\le K} L_n(X_{i_1},X_{i_2}), \mathfrak{F}_K, 2\le K\le n\} \text{ and } \{\sum_{1\le i_1\neq i_2\le n} L_{i_2}(X_{i_1},X_{i_2}), \mathfrak{F}_n\}$$

are martingales.

1.9 Convergence in distribution

In this section we briefly summarize some results on the central limit theorems (CLT) and the weak convergence of non-degenerate $U-$statistics.

The first central limit theorem for non-degenerate U-statistics was established by Hoeffding [27] under the condition that the second moment of the kernel of the underlying U-statistic, U_n, exists. The following Lemma 1.9.1 can be found in Section 5.5.1 of Serfling [35].

Lemma 1.9.1. *If $E(h^2(X_1, \ldots, X_m)) < \infty$, $E(h(X_1, \ldots, X_m)) = \theta$ and U_n is non-degenerate, i.e., is of rank $r = 1$ with $\xi_1 = E(\tilde{h}_1(X_1))^2 > 0$. Then, as $n \to \infty$,*

$$n^{\frac{1}{2}}(U_n - \theta) \longrightarrow_d N(0, m^2\xi_1),$$

where $\longrightarrow_d$ stands for convergence in distribution.

Remark 1.9.1. *For future consideration in this thesis, it is crucial to note that there are two conditions in Lemma 1.9.1, namely (i) $Eh^2 < \infty$ and (ii) $0 < \xi_1$. Condition (i) implies that in (ii) we also have $\xi_1 < \infty$, since*

$$\begin{aligned} \xi_1 &= E(\tilde{h}_1(X_1))^2 \\ &= E\{E^2(h(X_1, \ldots, X_m) - \theta \big| X_1)\} \\ &\le E(h(X_1, \ldots, X_m) - \theta)^2 \\ &\le E(h^2(X_1, \ldots, X_m)) < \infty. \end{aligned}$$

It is obvious that having $\xi_1 < \infty$, does not imply the square integrability of the kernel h. In other words, having $\xi_1 < \infty$ does not necessarily guarantee the existence of the variance of U_n.

We shall state our weak convergence results under more relaxed conditions than those mentioned in (i) and (ii) of Remark 1.9.1. However, before going any further, we spell out a classical result on weak convergence of processes of U-statistics that is due to Miller and Sen [30]. Define

$$\begin{aligned} Y_n^*(t) &= 0 && \text{for } 0 \le t \le \frac{m-1}{n}, \\ Y_n^*(\frac{k}{n}) &= \frac{k(U_k - \theta)}{m\sqrt{nVar(\tilde{h}_1(X_1))}} && \text{for } k = m, \ldots, n \end{aligned}$$

and for $t \in \left[\frac{k-1}{n}\ ,\ \frac{k}{n}\right]$, $k = m, \ldots, n$, let

$$Y_n^*(t) = Y_n^*(\frac{k-1}{n}) + n(t - \frac{k-1}{n}) \left(Y_n^*(\frac{k}{n}) - Y_n^*(\frac{k-1}{n}) \right).$$

Then the cited result of Miller and Sen can be stated as follows.

Lemma 1.9.2. *If*

(I) $0 < E\big[(h(X_1, X_2, \ldots, X_m) - \theta)(h(X_1, X_{m+1}, \ldots, X_{2m-1}) - \theta)\big] = Var(\tilde{h}_1(X_1))$

and

(II) $Eh^2(X_1, \ldots, X_m) < \infty$,

then, as $n \to \infty$,

$$Y_n^*(t) \to_d W(t) \quad on\ (C[0,1], \rho),$$

where $\{W(t), 0 \le t \le 1\}$ *is a standard Wiener process, and* $\longrightarrow_d$ *now means weak convergence on the function space* $C[0,1]$ *equipped with the sup-norm* ρ.

1.10 Almost sure behavior of U-statistics

In this section we shall summarize two results, for future use, concerning the almost sure behavior of U-statistics.

The following lemma is the classical SLLN due to Hoeffding in [27], which we quote from Section 5.4 of Serfling [35].

Lemma 1.10.1. *If* $E|h(X_1,\ldots,X_m)| < \infty$ *and* $E(h(X_1,\ldots,X_m)) = \theta$, *then, as* $n \longrightarrow \infty$,

$$U_n \longrightarrow \theta \quad a.s.$$

The following Marcinkiewicz type law of large numbers for U-statistics is due to Giné and Zinn [23].

Lemma 1.10.2. *If* $E|h(X_1,\ldots,X_m)|^{m/s} < \infty$ *with* $s > m$ *then, as* $n \longrightarrow \infty$,

$$n^{m-s}\binom{n}{m}^{-1} \sum_{1\leq i_1<\ldots<i_m\leq n} \left|h(X_1,\ldots,X_m)\right| \longrightarrow 0, \quad a.s.,$$

$$i.e.,\ n^{-s} \sum_{1\leq i_1<\ldots<i_m\leq n} \left|h(X_1,\ldots,X_m)\right| \longrightarrow 0 \quad a.s.$$

1.11 Maximal inequalities

This section is devoted to stating some maximal inequalities for future reference. The following lemma is a well-known inequality established by Chow [6]. This maximal inequality extends Kolmogorov's maximal inequality for sums of independent random variables, to semi-martingales (cf., section 1.12).

Lemma 1.11.1. *Let* (y_k) *be a semi-martingale relative to the* σ*-fields* $(\mathfrak{F}_k)$, c_k *be a* *sequence of non increasing positive numbers and $\epsilon > 0$. Then,*

$$\epsilon P(\max_{1\le k\le m} c_k y_k \ge \epsilon) \le c_1 E(y_1^+) + \sum_{k=2}^{m} c_k E(y_k^+ - y_{k-1}^+) - c_m \int_{\{\max_{1\le k\le m} c_k y_k < \epsilon\}} y_m^+ dP$$

$$\le c_1 E(y_1^+) + \sum_{k=2}^{m} c_k E(y_k^+ - y_{k-1}^+)$$

$$= \sum_{k=1}^{m-1} (c_k - c_{k+1}) E(y_k^+) + c_m E(y_m^+),$$

where $z^+ = \max\{0, z\}$.

A direct consequence of Lemma 1.11.1 is another well-known maximal inequality when $c_k = 1$, for each $k = 1, 2, \ldots$, and it can be stated as follows.

Lemma 1.11.2. *Let (y_k) be a martingale relative to the σ-fields $(\mathfrak{F}_k)$. Then,*

$$\epsilon P(\max_{1\le k\le m} |y_k| \ge \epsilon) \le E(|y_m|).$$

An alternative way of deriving Lemma 1.11.2, can be found in [26].

1.12 Appendix

The following definition of semi-martingale is due to Doob [20].

Definition 1.12.1. *Let $\{X_t, t \in T\}$ be a stochastic process, with $E(|X_t|) < \infty$, $t \in T$, and suppose that to each $t \in T$ corresponds a Borel field (σ-field) $\mathfrak{F}_t$ of measurable ω sets such that*

(i) $\mathfrak{F}_s \subset \mathfrak{F}_t, \quad s < t$;

(ii) X_t is either measurable with respect to the σ-field $\mathfrak{F}_t$ or is equal, for almost all ω, to a function that is;

either

(iiia) $X_s = E(X_t|\mathfrak{F}_s)$, with probability 1, whenever $s < t$,

or

(iiib) the process is real and $X_s \leq E(X_t|\mathfrak{F}_s)$ with probability 1, whenever $s < t$, then, $\{X_t, t \in T\}$ is called a semi martingale.

Remark 1.12.1. *An application of Jensen's inequality implies that when Φ is a real convex function and $\{X_t, t \in T\}$ is a semi martingale, then, so is $\{\Phi(X_t), t \in T\}$.*

Remark 1.12.2. *It follows directly from the definition that every martingale is a semi martingale.*

Remark 1.12.3. *Semi martingales are called submartingales too.*

Chapter 2

Partial Sums; A rich theory

2.1 Introduction

The reduction procedure presented in Section 1.7 provides the framework for our study of weak approximations of non-degenerate U-statistics. It was already mentioned in Section 1.7 (cf. Remark 1.7.2 of Chapter 1) that applying the reduction principle to centered non-degenerate statistics results in partial sums of centered i.i.d. random variables. As a matter of fact, this reduction reduces the underlying U-statistic, asymptotically in probability, to partial sums of i.i.d. random variables which we shall assume to be in the *domain of attraction of the normal law* (DAN). In other words, deriving weak convergence results for non-degenerate U-statistics is reduced to deriving weak convergence of partial sums. Therefore, this section is devoted to a review of some important results on partial sums of i.i.d. random variables. Our main attention will be given to the partial sums of i.i.d. random variables that are in the domain of attraction of the normal law.

2.2 CLT in the domain of attraction of the normal law

The well known classical central limit theorem (CLT) for a sequence $X_1, X_2, \ldots$ of i.i.d. random variables with mean μ and variance $0 < \sigma^2 < \infty$ states that as $n \longrightarrow \infty$

$$P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \leq x\right) \longrightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-\frac{t^2}{2})dt, \ x \in \mathbb{R}.$$

Perhaps the most significant progress in proving CLT under more relaxed moment conditions than those of the classical one, is the result obtained by Giné, Götze and Mason [24]. In their work, these authors answer the following question: " When is the Student t-statistic asymptotically standard normal?". To state their result we shall first define the concept of domain of attraction of the normal law (DAN) and briefly address the closely related ideas of Studentization and self-normalization.

Definition 2.2.1. *A sequence $X_1, X_2, \ldots,$ of i.i.d. random variables is said to be in the domain of attraction of the normal law ($X \in DAN$) if there exist sequences of constants A_n and $B_n > 0$ such that, as $n \to \infty$,*

$$\frac{\sum_{i=1}^{n} X_i - A_n}{B_n} \longrightarrow_d N(0,1).$$

Remark 2.2.1. *It is known that in the preceding definition, A_n can be taken as $nE(X)$ and $B_n = n^{1/2}\ell_X(n)$, where $\ell_X(n)$ is a slowly varying function at infinity (i.e., $\lim_{n\to\infty} \frac{\ell_X(nk)}{\ell_X(n)} = 1$ for any $k > 0$), defined by the distribution of X.*

With the possibility of X_1 having an infinite variance, properly normalizing the partial sums of i.i.d. random variables with mean zero $S_n = \sum_{i=1}^{n} X_i$, $n \geq 1$, such that convergence in distribution takes place, is the first question that comes to mind. To answer this question first consider the case when $\sigma^2 = Var(X_1) < \infty$. Now the classical SLLN for i.i.d. random variables implies that

$$\frac{1}{n\sigma^2} \sum_{i=1}^{n} X_i^2 \to 1 \ a.s.$$

This shows that the sequence $V_n = (\sum_{i=1}^{n} X_i^2)^{1/2}$ tends to infinity with the same rate as that of $\sqrt{n}\sigma$, i.e. V_n has the same rate as the right normalizer in the classical CLT for S_n. Hence V_n is a good candidate as a normalizing sequence. Having V_n as a normalizing candidate, how fast does it grow to infinity as $n \to \infty$, when the underlying sequence is in DAN and its variance does not exist?

To answer to this question we quote the following lemma from [15].

Lemma 2.2.1. *Let $X_1, X_2, \ldots,$ be a sequence of i.i.d. random variables in DAN such that $E(X_1^2) = \infty$. Let $B_n \nearrow \infty$, a set of constants for which*

$$\frac{1}{B_n} \sum_{i=1}^{n} (X_i - E(X_1)) \longrightarrow_d N(0,1).$$

Then, as $n \to \infty$,

$$\frac{1}{B_n^2} \sum_{i=1}^{n} X_i^2 \longrightarrow_P 1.$$

By virtue of Definition 2.2.1 and Remark 2.2.1, a restatement of the latter lemma, which we shall call Raikov's theorem, reads as follows.

Lemma 2.2.2. *Let $X_1, X_2, \ldots,$ be a centered sequence of i.i.d. random variables in DAN. Then, as $n \to \infty$,*

$$\frac{V_n^2}{n\ell^2(n)} \longrightarrow_P 1.$$

Let $\bar{X}_n = \dfrac{S_n}{n}$, $V_n = (\sum_{i=1}^n X_i^2)^{1/2}$ and recall that the classical Student t-statistic $T_n(X)$ is of the following form

$$\begin{aligned} T_n(X) &= \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i}{(\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2)^{1/2}} \\ &= \frac{S_n/V_n}{\sqrt{(n - (S_n/V_n)^2)/(n-1)}}. \end{aligned}$$

It is noteworthy to mention that when X_1 is normally distributed with mean μ and variance σ^2, then $T_n(X - \mu)$ is the well known Student t-statistic with $n-1$ degrees of freedom.

In this thesis, we shall call forms similar to S_n/V_n *self-normalized* and those of the form similar to $T_n(X)$ *Studentized.*

In light of our earlier setup, we now state the classical result of Giné, Götze and

Mason [24] in the following lemma.

Lemma 2.2.3. *The following two statements are equivalent:*

(i) $X_1 \in DAN$*;*

(ii) There exists a finite constant μ *such that, as* $n \to \infty$,

$$T_n(X - \mu) \to_d N(0,1).$$

Moreover, if either (i) or (ii) holds, then $\mu = E(X_1)$.

Giné, Götze and Mason [24], also proved that, if the sequence S_n/V_n, $n \geq 1$, is stochastically bounded, then they are uniformly sub-Gaussian in the sense that

$$\sup_{n\geq 1} E\, e^{\frac{tS_n}{V_n}} \leq 2e^{ct^2}$$

for all $t \in \mathbb{R}$ and some finite c.

The latter property, which obviously does not hold true if we replace V_n by $\sigma\sqrt{n}$ when $\sigma < \infty$, in turn, constitutes a basic requirement in the proof of the convergence of the moments of S_n/V_n to those of $N(0,1)$.

2.3 Evolution of Donsker's theorem

The fascinating idea of invariance principle was first used in the work of Kac [28] and then conceived and fundamentally explored as a functional CLT by Erdős and Kac [21] (cf. [7] for review). Namely, they obtained the asymptotic distributions of certain functionals of partial sums of i.i.d. random variables via first proving that the limiting distribution of each functional of their interest exists and is independent of the initial distribution of the original observations. This enabled them to choose any initial distribution, based on which the derivation of the limiting distribution would be most convenient. As the limiting distribution remains invariant when passing to the general case from initial distribution, they called their approach the invariance principle. Passing to the general case, the results were also connected to appropriate functionals of Brownian motion.

In 1951 Donsker [18] extended the invariance principle to processes of partial sums, when he proved his well known functional central limit theorem for centered i.i.d. random variables $X_1, X_2, \cdots$ and via interpolation of their partial sums $S_n = \sum_{i=1}^{n} X_i$, $(S_0 = 0)$, he constructed the process $\{S_n(t), 0 \le t \le 1\}$ on $C[0,1]$ according to:

$$S_n(t) = \frac{1}{\sigma\sqrt{n}} S_{[nt]} + \frac{1}{\sigma\sqrt{n}}(nt - [nt]) X_{[nt]+1}.$$

He showed that, as $n \to \infty$, the distribution of $\{S_n(t), 0 \le t \le 1\}$ coincides with that of a standard Wiener process $\{W(t), 0 \le t \le 1\}$, i.e., as $n \to \infty$,

$$S_n(t) \longrightarrow_d W(t) \text{ on } (C[0,1], \rho), \tag{2.3.1}$$

where $\longrightarrow_d$ stands for weak convergence on the function pace $C[0,1]$ equipped with the sup-norm ρ.

Remark 2.3.1. *By Theorem 5.2 of Billingsley [3] (cf. also footnote of page 76 of [3]), the weak convergence in (2.3.1) is equivalent to:*

$$h(S_n(.)) \longrightarrow_d h(W(.)) \text{ on } (C[0,1], \rho),$$

for all $h : C[0,1] \to \mathbb{R}$ that are continuous with respect to the sup-norm metric, almost surely with respect to Wiener measure.

The weak convergence result of Donsker [18] requires the existence of a finite positive variance. A significant next step was made by Csörgő, Szyszkowicz and Wang [CsSzW] in [14] and [15]. They proved that Donsker's theorem continued to hold true when $X_1, X_2, \cdots$ were simply assumed to be in the domain of attraction of the normal law (DAN). Thus, the possibility of having the X_i's $\in DAN$ with infinite variance, will not stop Donsker's theorem holding true. This extension involves the concept of *self-normalization* and *Studentization*, as it was the case in Lemma 2.2.3 of Giné, Götze and Mason [24]. These results of CsSzW in [14] and [15] extend those of Giné, Götze and Mason to functional central limit theorems. For future reference we now state their Theorem 1 in [14] as the following lemma.

Lemma 2.3.1. *As $n \to \infty$, the following statements are equivalent:*

(a) $X_1 \in DAN$ *and* $EX_1 = \mu$;

(b) $\dfrac{\sum_{i=1}^{[nt_0]}(X_i - \mu)}{V_n} \longrightarrow_d N(0, t_0)$ *for* $t_0 \in (0, 1]$;

(c) $\dfrac{\sum_{i=1}^{[nt]}(X_i - \mu)}{V_n} \Rightarrow W(t)$ *on* $(D[0,1], \rho)$,

where ρ is the sup-norm metric for functions in $D[0,1]$ and $\{W(t), 0 \le t \le 1\}$ is a standard Wiener process;

(d) On an appropriate probability space for $X_1, X_2, \ldots$, we can construct a standard Wiener process $\{W(t), 0 \le t < \infty\}$ such that

$$\sup_{0 \le t \le 1} \left| \frac{\sum_{i=1}^{[nt]}(X_i - \mu)}{V_n} - \frac{W(nt)}{n^{\frac{1}{2}}} \right| = o_P(1).$$

Remark 2.3.2. *For the statements of Lemma 2.3.1 we have that, clearly, (d) implies (c), (c) implies (b) and, on account of Theorem 2.1 of [15] (cf. also Theorem 3.3 of [24]), (b) with $t_0 = 1$ implies (a). Hence the proof of Lemma 2.3.1 follows by showing that (a) implies (d). We refer the reader to [14] for the proof of the latter.*

Remark 2.3.3. *What we mean by saying that (d) implies (c) is the following functional central limit theorem.*

$$h\big(S_{[n.]}/V_n\big) \longrightarrow_d h\big(W(.)\big)$$

for all $h : D = D[0,1] \to \mathbb{R}$ that are $(D, \mathfrak{D})$ measurable and ρ-continuous, or ρ-continuous except for on a set of Wiener measure zero on $(D, \mathfrak{D})$, where $\mathfrak{D}$ denotes the $\sigma-$field of subsets of D generated by the finite-dimensional subsets of D.

The following Lemma 2.3.2 (cf. Proposition 2.1 in [15]), which will directly be applied in our proofs, is a restatement of Lemma 2.3.1 in view of Raikov's theorem (i.e., Lemma 2.2.2). In other words, we replace V_n by B_n, that was defined in Remark 2.2.1 of this chapter.

Lemma 2.3.2. *As $n \to \infty$, the following statements are equivalent:*
(a) $X_1 \in DAN$ and $EX_1 = \mu$;
(b) $\dfrac{\sum_{i=1}^{[nt_0]}(X_i - \mu)}{B_n} \longrightarrow_d N(0, t_0)$ for $t_0 \in (0, 1]$;
(c) $\dfrac{\sum_{i=1}^{[nt]}(X_i - \mu)}{B_n} \Rightarrow W(t)$ on $(D[0,1], \rho)$,
where ρ is the sup-norm metric for functions in $D[0,1]$ and $\{W(t), 0 \le t \le 1\}$ is a standard Wiener process;
(d) On an appropriate probability space for $X_1, X_2, \ldots,$ we can construct a standard Wiener process $\{W(t), 0 \le t < \infty\}$ such that

$$\sup_{0 \le t \le 1} \left| \frac{\sum_{i=1}^{[nt]}(X_i - \mu)}{B_n} - \frac{W(nt)}{n^{\frac{1}{2}}} \right| = o_P(1).$$

We note that $\Rightarrow$ in part (c) of Lemma 2.3.2 is as it was defined in Remark 2.3.2.

Remark 2.3.4. *We observe that the respective (d) statements of Lemmas 2.3.1 and*

2.3.2 provide weak approximations of partial sums from which appropriate weak convergence can be derived, for example like that of their respective (c) parts.

Lemma 2.3.1 and Lemma 2.3.2 were not the only contributions of CsSzW. In [15] and [16], they succeeded to show that Lemmas 2.3.1 and 2.3.2 hold true when an appropriate weighted sup-norm metric $\|/q\|$ is incorporated. In other words, they prove weighted versions of Lemmas 2.3.1 and 2.3.2 (cf. Section 2.4 for some literature on weight functions $q(.)$, and their role in the area of weak convergence).

To formally state these weighted versions of Lemmas 2.3.1 and 2.3.2 for future references in this thesis, we proceed by describing the class of weight functions of interest and some of their properties.

As in CsSzW [14] and [15], and also throughout the thesis, we let Q to be the class of functions $q(t)$, which are positive on $(0,1]$, i.e., $\inf_{\delta\le t\le 1} q(t) > 0$ for $0<\delta<1$, and non-decreasing in a neighborhood of zero. Moreover, define the integral $I(q,c)$ as follows.

$$I(q,c) := \int_{0+}^{1} t^{-1}\exp(-cq^2(t)/t)\,dt, \quad 0<c<\infty.$$

The following lemmas, which characterize the class Q, can be found in Csörgő, Csörgő, Horváth and Mason [8] (cf. also [12]).

Lemma 2.3.3. *Let $q\in Q$. If $I(q,c)<\infty$ for some $c>0$, then*

$$\lim_{t\downarrow 0}\frac{t^{\frac{1}{2}}}{q(t)} = 0.$$

Lemma 2.3.4. *Let $\{W(t), 0\le t<\infty\}$ be a standard Wiener process and $q\in Q$. Then,* *(a)* $I(q,c) < \infty$ *for all* $c > 0$ *if and only if*

$$\limsup_{t \downarrow 0} \frac{|W(t)|}{q(t)} = 0 \ a.s.$$

(b) $I(q,c) < \infty$ *for some* $c > 0$ *if and only if*

$$\limsup_{t \downarrow 0} \frac{|W(t)|}{q(t)} < \infty \ a.s.$$

The presence of the weight function q requires a definition of a proper metric on $D[0,1]$ by which weak convergence is defined. This definition is given as follows.

Definition 2.3.1. *For the functions* x, y *on* $[0,1]$ *and* $q \in Q$, *define the weighted sup-norm metric*

$$\|(x-y)/q\| = \sup_{0<t\leq 1} |x(t) - y(t)|/q(t),$$

whenever it is well defined, i.e., $\limsup_{t\downarrow 0} |x(t) - y(t)|/q(t) < \infty$. *A short hand notion for this metric will be* $\|/q\|$.

We now quote Corollary 3 of CsSzW [16] that is a consequence of their Theorem 2 in [16].

Lemma 2.3.5. *Let* $q \in Q$. *As* $n \to \infty$, *the following statements are equivalent:*

(a) $X_1 \in DAN$ *and* $EX_1 = \mu$;

(b) $\dfrac{\sum_{i=1}^{[nt]}(X_i - \mu)}{V_n} \Rightarrow W(t)$ *on* $(D, \mathfrak{D}, \|/q\|)$ *if and only if* $I(q,c) < \infty$ *for all* $c > 0$, *where* $\{W(t), 0 \leq t \leq 1\}$ *is a standard Wiener process;*

(c) On an appropriate probability space for $X_1, X_2, \ldots,$ *a standard Wiener process*

$\{W(t), 0 \leq t \leq \infty\}$ can be constructed in such a way that as $n \to \infty$,

$$\sup_{0<t\leq 1} \left| \frac{\sum_{i=1}^{[nt]}(X_i - \mu)}{V_n} - \frac{W(nt)}{\sqrt{n}} \right| / q(t) = o_P(1)$$

if and only if $I(q, c) < \infty$ for all $c > 0$.

Remark 2.3.5. *The statement (b) of Lemma 2.3.5 stands for the following functional central limit theorem on $(D, \mathfrak{D}, \|/q\|)$, where $\mathfrak{D}$ is the σ-field of subsets of $D = D[0, 1]$ generated by its finite-dimensional subsets, and $\|/q\|$ stands for the weighted sup-norm metric in $D = D[0, 1]$ with $q \in Q$ that is also càdlàg. With $\longrightarrow_d$ standing for convergence in distribution as $n \to \infty$, we have*

$$g\left(\frac{\sum_{i=1}^{[n.]} X_i - \mu}{V_n q(.)}\right) \to_d g\left(\frac{W(.)}{q(.)}\right)$$

for all càdlàg functions functions $q \in Q$, and for all $g : D = D[0, 1] \to \mathbb{R}$ that are $(D, \mathfrak{D})$ measurable and $\|/q\|$-continuous or $\|/q\|$-continuous except at points forming a set of Wiener measure zero on $(D, \mathfrak{D})$, generated by a standard Wiener process $W(.)$ on the unit interval $[0, 1]$.

The following lemma is a restatement of Lemma 2.3.5 replacing V_n by B_n in light of Raikov's theorem, i.e., Lemma 2.2.2, and it coincides with Corollary 1 of CsSzW [16] that is a direct consequence of their Theorem 1.

Lemma 2.3.6. *Let $q \in Q$. As $n \to \infty$, the following statements are equivalent:*

(a) $X_1 \in DAN$ and $EX_1 = \mu$;

(b) $\frac{\sum_{i=1}^{[nt]}(X_i - \mu)}{B_n} \Rightarrow W(t)$ on $(D, \mathfrak{D}, \|/q\|)$ if and only if $I(q, c) < \infty$ for all $c > 0$, where $\{W(t), 0 \leq t \leq 1\}$ is a standard Wiener process;

(c) On an appropriate probability space for $X_1, X_2, \ldots$, a standard Wiener process

$\{W(t), 0 \le t \le \infty\}$ *can be constructed in such a way that as* $n \to \infty$,

$$\sup_{0<t\le 1} \left| \frac{\sum_{i=1}^{[nt]}(X_i - \mu)}{B_n} - \frac{W(nt)}{\sqrt{n}} \right| / q(t) = o_P(1)$$

if and only if $I(q,c) < \infty$ *for all* $c > 0$.

We note that $\Rightarrow$ in part (b) of the above lemma is as it has already been defined in Remark 2.3.5 of this chapter.

In addition to the preceding two lemmas, in which weighted weak convergence for partial sums is derived (via weighted approximations in probability), the following Lemma 2.3.7, which is due to CsSzW [16] (cf. Corollary 4 in [16] and Corollary 3.4 in [15]), gives convergence in distribution results for weighted sup-functional of partial sums, where the weight functions q are from the larger class of weight functions characterized in part (b) of Lemma 2.3.4. This result is not implied by Lemma 2.3.5 and can not be obtained via classical methods of weak convergence. The reason is the tightness is not guaranteed by part (b) of Lemma 2.3.4.
Let $\{W(t), 0 \le t \le 1\}$ be a standard Wiener process, $X_1 \in DAN$ and $E(X_1) = 0$.

Lemma 2.3.7. *If* $q \in Q$ *and* $q(t)$ *is non decreasing on* $(0,1]$, *then, as* $n \to \infty$,

$$V_n^{-1} \sup_{0<t\le 1} \frac{|S_{[nt]}|}{q(t)} \longrightarrow_d \sup_{0<t\le 1} \frac{|W(t)|}{q(t)}$$

if and only if $I(q,c) < \infty$ *for some* $c > 0$. *Consequently, as* $n \to \infty$, *we have*

$$V_n^{-1} \sup_{0<t\le 1} \frac{|S_{[nt]}|}{(t \log\log t^{-1})^{1/2}} \longrightarrow_d \sup_{0<t\le 1} \frac{|W(t)|}{(t \log\log t^{-1})^{1/2}}.$$

Here and throughout, $\log(x) := \log(\max e, x)$.

In view of Lemma 2.2.2 (Raikov's theorem) the conclusions of Lemma 2.3.7 continues to hold true if we replace the normalizing sequence V_n by B_n. This Version of Lemma 2.3.7 is Corollary 2 in CsSzW [16] that is proved directly there.

2.4 Appendix: A short glimpse of earlier weighted approximations

Rényi, [34], in 1953 studied the asymptotic distribution of modified Kolmorgrov-Smirnov statistic. His idea was to study the statistic $|F_n(x) - F(x)|/F(x)$ instead of $|F_n(x) - F(x)|$, where $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{(X_i \le x)}$ is the empirical distribution of the first n observations of the sequence of i.i.d. random variables $X_1, X_2, \ldots$ with distribution function $F(.)$. The reason of this modification was to increase the sensitivity of of the classical statistic on the tails from the hypothesized continuous distribution $F(.)$.More precisely, he studied the the asymptotic distribution of

$$n^{1/2} \sup_{a \le F(s) \le b} (F_n(s) - F(s))/F(s), \ 0 < a < b \le 1$$

and

$$n^{1/2} \sup_{a \le F(s) \le b} (F_n(s) - F(s))/(1 - F(s)), \ 0 \le a < b < 1.$$

To proceed we note that if F is continuous then without loss of generality these statistics can be replaced respectively by

$$n^{1/2} \sup_{a \le s \le b} \alpha_n(s)/s, \ 0 < a < b \le 1$$

and

$$n^{1/2} \sup_{a \le s \le b} \alpha_n(s)/(1 - s), \ 0 \le a < b < 1,$$

where the empirical process α_n is defined as

$$\alpha_n(s) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{(U_i \le s)} - s$$

and $U_1, U_2, \cdots$, is an i.i.d. sequence of $unifrom(0,1)$.

In the same spirit in 1952 Anderson and Darling [1] studied the asymptotics of statistics like

$$n^{1/2} \sup_{a \le s \le b} \alpha_n(s)/(s(1-s))^{1/2}, \ 0 < a < b < 1.$$

A natural generalization of statistics of the latter forms lead to a weak convergence result for the process $\alpha_n(.)$ in $D[0,1]$ in $\|/q\|$-metric to a Brownian bridge . Namely, if we let Q^* denote the class of positive functions on $(0,1)$ that are bounded away from zero on $(\delta, 1-\delta)$, for all $0 < \delta < 1/2$, non-decreasing in a neighborhood of 0 and non-increasing in neighborhood of 1, then one would want to characterize the functions $q \in Q^*$ for which, as $n \to \infty$, one would have that

$$n^{1/2}\alpha_n(s)/q(s) \text{ converges weakly to } B(s)/q(s), \tag{2.4.1}$$

where $\{B(t), 0 \le t \le 1\}$ is a Brownian bridge, i.e., a mean zero Gaussian process with $E(B(t)B(s)) = \min(s,t) - st$, with $t, s \in (0,1)$. Such characterizations for $q \in Q^*$ are due to Chibisov [5], O'Reilly [33] and Csörgő, Csörgő, Horváth and Mason [CsCsHM] [8] conclude that (2.4.1) holds true if and only if

$$\limsup_{t \downarrow 0} |B(t)|/q(t) = \limsup_{t \uparrow 1} |B(t)|/q(t) = 0, \ a.s.$$

In 1986, CsCsHM in [8], derived weighted approximations for uniform empirical and quantile processes. To state one of their results for the uniform quantile process, consider again an i.i.d. sequence of $uniform(0,1)$ random variables $U_1, U_2, \cdots$, and let $U_{1,n}, \ldots, U_{n,n}$, be their first $n \geq 1$ order statistics. For each $n \geq 1$ define

$$U_n(s) = U_{k,n}, \quad (k-1)/n < s \leq k/n, \ k = 1, \ldots, n,$$

where $U_n(0) = U_{1,n}$, and the quantile process

$$u_n := s - U_n(s), \ 0 \leq s \leq 1.$$

The following Lemma 2.4.1 is Theorem 2.1 in CsCsHM [8], that is one of the main results in their Section 2 (cf. also Theorem in Csörgő, Horváth [9]).

Lemma 2.4.1. *A sequence of independent uniform*$(0,1)$ *random variables* $U_1, U_2, \cdots$ *and a sequence of Brownian bridges* $B_1, B_2, \cdots$ *can be constructed on the same probability space so that for all* $0 \leq v_1 < 1/2$

$$\sup_{1/(n+1) \leq s \leq n/(n+1)} n^{v_1} |n^{1/2} u_n(s) - B_n(s)| / (s(1-s))^{1/2 - v_1} = O_P(1).$$

The partial sums based proof of Lemma 2.4.1 (cf. [8] and [9]) inspired the study of weighted approximations of partial sums themselves. In their study of some change-point problems, Csörgő and Horváth in [10] and [11] also dealt with the tide down partial sums process

$$Z_n(t) = \begin{cases} (S_{[(n+1)t]} - \frac{[(n+1)t]}{n} S_n)/(\sigma n^{1/2}), & 0 \leq t < 1; \\ 0, & t = 1, \end{cases}$$

where $\sigma^2 = Var(X_1)$, $S_{[(n+1)t]} = \sum_{i=1}^{[(n+1)t]} X_i$, and $X_1, X_2, \ldots$ is a sequence of i.i.d.

random variables such that

$$E|X_1|^r < \infty, \quad \text{for some } r > 2.$$

Under the latter moment condition they established that there exists a sequence of Brownian bridges $\{B_n(t), 0 \le t \le 1\}$ such that, as $n \to \infty$,

$$\sup_{1/(n+1) \le t \le n/(n+1)} n^{v_1} \big| Z_n(t) - B_n(t) \big| / (t(1-t))^{1/2 - v_1} = O_P(1),$$

where $0 \le v_1 < 1/2 - 1/r$.

Given this conclusion, along the lines of the proof of Theorem 4.2.1 in [8] one arrives at the following $\| / q \|$-metric version of Donsker's theorem (cf. also Corollary 2.1 of Csörgő and Horváth in [11]).

Lemma 2.4.2. *Let $q \in Q$, $\{W(t), 0 \le t < \infty\}$ be a standard Wiener process and $X_1, X_2, \ldots$, be a sequence of i.i.d. random variables with mean zero and variance variance $\sigma^2 > 0$ such that*

$$E|X_1|^r < \infty, \; \textit{for some } r > 2,$$

then, on an appropriate probability space for $X_1, X_2, \ldots$ and $W(.)$, as $n \to \infty$, we have

$$\sup_{0 < t \le 1} \big| (n\sigma^2)^{-1/2} (S_{[nt]} - W(nt)) \big| / q(t) = o_P(1)$$

if and only if

$$\limsup_{t \downarrow 0} |W(t)| / q(t) = 0 \; a.s. \tag{2.4.2}$$

where Q is the class of positive functions on $(0, 1]$ which are bounded away from zero on $(\delta, 1]$ for all $\delta \in (0, 1)$ and non-decreasing in a neighborhood of 0.

A similar approximation result holds for $Z_n(t)$ under the condition that $E|X_1|^r < \infty$ for some $r > 2$ in $\|/q\|$-metric with $q \in Q^*$.

It was in 1992 when Szyszkowicz [36] succeeded in showing that with $q \in Q$, characterized by (2.4.2), the weighted weak approximation of Lemma 2.4.2 continues to hold true on assuming only the existence of two moments for X_1. The following Lemma 2.4.3 is a restatement of Theorem 1.2.1 of Szyszkowicz is in [36] (cf. also [38] and [37] for extensions and proofs).

Lemma 2.4.3. *Let $X_1, X_2, \cdots$ be a sequence of i.i.d. random variables such that*

$$E(X_1) = 0, \quad E(X_1^2) = 1,$$

and for each $n \geq 1$ let $S(nt) = \sum_{i=1}^{[nt]} X_i$. Then a standard Wiener process $\{W(t), 0 \leq t < \infty\}$ can be constructed in such a way that the following statements hold true.

(a) Let $q \in Q$. Then, as $n \to \infty$

$$\sup_{0<t\leq 1} \left| n^{-1/2}(S(nt) - W(nt)) \right| / q(t) = o_P(1)$$

if and only if $I(q, c) < \infty$ for all $c > 0$.

(b) Let $q \in Q$. Then, as $n \to \infty$

$$\sup_{0<t\leq 1} \left| n^{-1/2}(S(nt) - W(nt)) \right| / q(t) = O_P(1)$$

if and only if $I(q, c) < \infty$ for some $c > 0$.

From part (a) of the latter theorem of Szyszkowicz it follows that Donsker's theorem holds true in terms of $\|/q\|$-metric, when $q \in Q$ satisfies (2.4.2) and assuming only that the variance exists.

Based on CsSzW [14], [15] and [16] in Section 2.3, we summarized some recent significant progress in the area of weighted approximations of partial sums, when $X \in DAN$.

Chapter 3

Weighted approximations of Pseudo-self-normalized U-statistics and U-type statistics

3.1 Introduction

The true inspiration of this thesis comes from the work of CsSzW in [17] and the earlier work of Csörgő and Horváth in [10] in the area of changepoint analysis (cf. also [13]). We first state the so-called At Most One Change (AMOC) problem. Then, for the sake of calling the reader's attention to the motivation of this work, we briefly summarize the main results in the cited papers.

Detecting possible changes in distribution of a set of observations is referred to as a *changepoint* problem. We now state the just mentioned AMOC problem.

Let $X_1, X_2, \cdots,$ be a sequence of non-degenerate independent real valued random variables with distribution function F. We are interested in testing the null hypothesis:

$$H_0: \quad X_i, \ 1 \le i \le n, \ \textit{have the same distribution},$$

versus the one change in distribution alternative:

$$H_A: \quad \textit{there is an integer } k, \ 1 \le k < n, \ \textit{such that}$$

$$P(X_1 \le t) = \cdots = P(X_k \le t), \ P(X_{k+1} \le t) = \cdots = P(X_n \le t)$$

$$\textit{for all } 0 \le t < 1 \textit{ and } P(X_k \le t_0) \ne P(X_{k+1} \le t_0) \textit{ for some } t_0.$$

In 1988 in their work of [10] (cf. also [11]) Csörgő and Horváth introduced functionals of a U-statistics type process to test H_0 versus H_A. Letting $h(x,y)$ to be a kernel with mean θ, the Csörgő-Horváth centralized U-type process is defined by

$$Z_k = \sum_{i=1}^{k} \sum_{j=k+1}^{n} h(X_i, X_j) - \ k(k-1)\theta, \quad 1 \le k < n. \tag{3.1.1}$$

The idea is to divide the n already observed observations into the two blocks

$$\underbrace{X_1, \ldots, X_k,}\quad \underbrace{X_{k+1}, \ldots, X_n}$$

and comparing the k observation in the first block to the remaining $(n-k)$ of the second one, via an appropriate kernel of order 2, to capture the possibility of having a change in distribution at an unknown time k.

Although, Z_k is not a U-statistic itself, when h is symmetric, it can be written as a sum of three centered U-statistics as follows.

$$Z_k = \sum_{i=1}^{n} \sum_{j=i+1}^{n} h(X_i, X_j) - \binom{n}{2}\theta$$

$$-\Big\{\sum_{i=1}^{k} \sum_{j=i+1}^{n} h(X_i, X_j) - \binom{k}{2}\theta + \sum_{i=k+1}^{n} \sum_{j=i+1}^{n} h(X_i, X_j) - \binom{n-k}{2}\theta\Big\},\ 1 \le k \le n.$$

Csörgő and Horváth studied the asymptotic behavior of this functional of statistics via obtaining weak approximation of the sequence of stochastic processes in $D[0,1]$, $\{Z_{[nt]}, 0 \le t \le 1\}$ under the null hypothesis. Define the Guassian process

$$\Gamma(t) = (1-t)W(t) - t\ (W(1) - W(t)),\ 0 \le t \le 1,$$

where $\{W(t), 0 \le t < \infty\}$ is a standard Wiener process. Then under H_0, when $q \in Q^*$, $E|h(X_1, X_2)|^\nu < \infty$ for some $\nu > 2$ and $0 < \sigma^2 = E(\tilde{h}_1^2(X_1))$, Csörgő and Horváth [10] showed that a sequence of Gaussian processes $\{\Gamma_n(t), 0 \le t \le 1\}$ such that for $n \ge 1$,

$$\{\Gamma_n(t), 0 \le t \le 1\} =^d \{\Gamma(t), 0 \le t \le 1\}$$

can be constructed in a way that, as $n \to \infty$, we have

$$\sup_{0<t<1} \left| n^{-3/2}\sigma^{-1} Z_{[nt]} - \Gamma_n(t) \right| / q(t) = o_P(1) \tag{3.1.2}$$

if and only if

$$\int_{0+}^{1-} \frac{1}{t(1-t)} \exp\left(-\frac{cq^2(t)}{t(1-t)}\right) dt < \infty \text{ for all } c > 0.$$

In her PhD thesis, Szyszkowicz [36] showed that the approximation (3.1.2) continues to hold true assuming that $q \in Q^*$, $E(h^2(X_1, X_2)) < \infty$ and $0 < \sigma^2 = E(\tilde{h}_1^2(X_1))$. Clearly, the weak convergence of the process $n^{-3/2}\sigma^{-1} Z_{[nt]}$ on $(D, \mathfrak{D}, \| / q \|)$, as it is defined in Remark 2.3.5 of Chapter 2, follows from the weighted weak approximation (3.1.2).

A significant improvement in establishing weak approximations similar to (3.1.2) was made by CsSzW in [17]. The progress was made in two aspects: (i) they relaxed the condition that the second moment of the kernel exists to the one with 5/3 moments which in turn enabled them to have $\tilde{h}_1(X_1) \in DAN$ and (ii) using the method of jackknifing, they replaced σ^2 (which may not exist) by some normalizing sequence and concluded that, as $n \to \infty$, we have, under H_0,

$$\sup_{0<t<1} \left| n^{-3/2}\hat{\sigma}^{-1} \hat{Z}_{[nt]} - \Gamma_n(t) \right| / q(t) = o_P(1),$$

where q, Γ_n and Γ as before and

$$\hat{Z}_{[nt]} = \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} h^2(X_i, X_j) - n^2 t(1-t)\hat{\theta}, \quad 0 \le t < 1,$$

$$\hat{\theta} = \frac{1}{n(n-1)} \sum_{1 \le i \ne j \le n} h(X_i, X_j),$$

and

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}\Big(\frac{1}{n-1}\sum_{\substack{i=1\\ i\neq j}}^{n} h(X_i, X_j) - \frac{1}{n(n-1)}\sum_{1\le i\neq j\le n} h(X_i, X_j)\Big)^2.$$

This, in turn, constitutes the inspiration behind the work in this thesis.

In this chapter we derive a weak convergence result for processes of U-statistics and U-type statistics. To establish this result, we will introduce a normalizing sequence when the variance of the projection $\tilde{h}_1$ is not necessarily finite. Unlike the above $\hat{\sigma}^2$, this normalizing sequence relies on the distribution of the observations and can not be computed based on only the observations. To remedy this issue, in the next chapter we shall introduce another normalizing sequence, in the same spirit as the above $\hat{\sigma}^2$, based on which results similar to those of this chapter will be derived.

Employing truncation arguments and the concept of weak convergence of self-normalized and Studentized partial sums, we derive weak convergence results via approximations in probability for *pseudo-self-normalized U-statistics* and *U-statistic type* processes. Our results in this chapter require only that (i) the expected value of the kernel of the underlying U-statistic to be of order $\frac{4}{3}$ (instead of having 2 moments of the kernel), and that (ii) the conditional expected value of the kernel on each observation to be in the domain of attraction of the normal law (instead of having 2 moments). Similarly relaxed moment conditions were first used CsSzW (2008) in [17] for U-statistics type processes for changepoint problems in terms of kernels of order 2. Our results in this chapter extend their work to approximating U-statistics with higher order kernels. Our weak convergence results for U-statistics in turn extend those obtained by R.G. Miller Jr. and P.K. Sen (1972) in [30], quoted here earlier as Lemma 1.9.2.

3.2 Definitions and tools

Recall that a U-statistic based on the (symmetric) kernel h with mean θ was defined in (1.2.1) as follows.

$$U_n = \binom{n}{m}^{-1} \sum_{1 \le i_1 < \ldots < i_m \le n} h(X_{i_1}, \ldots, X_{i_m})$$

$$= [n]^{-m} \sum_{1 \le i_1 \ne \ldots \ne i_m \le n} h(X_{i_1}, \ldots, X_{i_m}).$$

In the same spirit we define a U-type statistic based on a symmetric kernel h with mean θ as follows. For $m \le K \le n$

$$\Psi_n = n^{-m} \sum_{1 \le i_1 < \ldots < i_m \le K} h(X_{i_1}, \cdots, X_{i_m})$$

$$= \frac{n^{-m}}{m!} \sum_{1 \le i_1 \ne \ldots \ne i_m \le K} h(X_{i_1}, \cdots, X_{i_m}). \tag{3.2.1}$$

Remark 3.2.1. *Despite their close relation, our notion of U-type statistics in (3.2.1) is different from the one was first defined by Csörgő and Horváth, i.e. Z_k in (3.1.1).*

With $V_n^2 := \sum_{i=1}^n \tilde{h}_1^2(X_i)$, define the pseudo self-normalized U-statistic process $U_{[nt]}^{nor}$ and the pseudo self-normalized U-type process $\Psi_{[nt]}^{nor}$, respectively as follows.

$$U_{[nt]}^{nor} = \begin{cases} 0, & 0 \le t < \dfrac{m}{n}, \\ \dfrac{[nt]}{m} \dfrac{U_{[nt]} - \theta}{V_n}, & \dfrac{m}{n} \le t \le 1. \end{cases}$$

$$\Psi_{[nt]}^{nor} = \begin{cases} 0, & 0 \le t < \frac{m}{n}, \\ \frac{n^{-m+1}}{m} \frac{\sum_{1 \le i_1 < \ldots < i_m \le [nt]} (h(X_{i_1}, \ldots, X_{i_m}) - \theta)}{V_n}, & \frac{m}{n} \le t \le 1, \end{cases}$$

where [.] denotes the greatest integer function.

The term *pseudo* refers to the fact that the normalizing sequence V_n, is not directly related to the kernel h, but to its projection $\tilde{h}_1$.

Although, as we shall see, V_n is the right normalizer for both $U_{[nt]}^{nor}$ and $\Psi_{[nt]}^{nor}$, in the sense that it leads to weak convergence to a Wiener process, it cannot be computed based on only the observations, i.e., the distribution of the observations F is still required. Therefore, the approach is not non parametric, yet. We will deal with this issue in the next chapter.

Considering that to derive our results we shall assume that $\tilde{h}_1(X_1) \in DAN$, we restate the conclusion of Raikov's theorem, cf. Lemma 2.2.2, in this context as follows. As $n \to \infty$,

$$\frac{V_n^2}{n\, \ell^2(n)} \longrightarrow_P 1. \tag{3.2.2}$$

Moreover, letting the class Q be as it was defined and characterized in Lemma 2.3.4 and $B_n = n^{1/2}\ell(n)$, when ℓ is the slowly varying function at infinity associated with $\tilde{h}_1$, also recalling that $\tilde{h}_1(X_i)$, $1 \le i \le n$, are centered i.i.d. random variables the CsSzW [16] based Lemmas 2.3.5 and 2.3.6 in this context read respectively as follows.

Lemma 3.2.1. *Let $q \in Q$. As $n \to \infty$ the following statements are equivalent:*

(a) $\tilde{h}_1(X_1) \in DAN$;

(b) $\dfrac{\sum_{i=1}^{[nt]} \tilde{h}_1(X_i)}{V_n} \Rightarrow W(t)$ *on* $(D, \mathfrak{D}, \|/q\|)$ *if and only if* $I(q, c) < \infty$ *for all* $c > 0$, *where* $\{W(t), 0 \le t \le 1\}$ *is a standard Wiener process;*

(c) On an appropriate probability space for $X_1, X_2, \ldots$, *a standard Wiener process*

$\{W(t), 0 \leq t < \infty\}$ can be constructed in such a way that as $n \to \infty$,

$$\sup_{0<t\leq 1} \Big|\frac{\sum_{i=1}^{[nt]} \tilde{h}_1(X_i)}{V_n} - \frac{W(nt)}{\sqrt{n}}\Big|/q(t) = o_P(1)$$

if and only if $I(q, c) < \infty$ for all $c > 0$.

Lemma 3.2.2. *Let $q \in Q$. As $n \to \infty$ the following statements are equivalent:*

(a) $\tilde{h}_1(X_1) \in DAN$;

(b) $\dfrac{\sum_{i=1}^{[nt]} \tilde{h}_1(X_i)}{B_n} \Rightarrow W(t)$ on $(D, \mathfrak{D}, \|/q\|)$ if and only if $I(q, c) < \infty$ for all $c > 0$, where $\{W(t), 0 \leq t \leq 1\}$ is a standard Wiener process;

(c) On an appropriate probability space for $X_1, X_2, \ldots$, a standard Wiener process $\{W(t), 0 \leq t < \infty\}$ can be constructed in such a way that as $n \to \infty$,

$$\sup_{0<t\leq 1} \Big|\frac{\sum_{i=1}^{[nt]} \tilde{h}_1(X_i)}{B_n} - \frac{W(nt)}{\sqrt{n}}\Big|/q(t) = o_P(1)$$

if and only if $I(q, c) < \infty$ for all $c > 0$.

In the following Section 3.3 we study weighted weak approximations of the process $U_{[nt]}^{nor}$. Similar results for the U-type process $\Psi_{[nt]}^{nor}$ will be derived in Section 3.4.

3.3 Pseudo self-normalized U-statistic processes

Theorem 3.3.1. *Let $q \in Q$ and assume*

$$E\big(|h(X_1, \ldots, X_m)|^{\frac{4}{3}} \log |h(X_1, \ldots, X_m)|\big) < \infty.$$

and that $\tilde{h}_1(X_1) \in DAN$. Then, as $n \to \infty$, we have

(a) $U_{[nt]}^{nor} \Rightarrow W(t)$ on $(D, \mathfrak{D}, \|/q\|)$ if and only if $I(q, c) < \infty$ for all $c > 0$, where

$\{W(t), 0 \le t \le 1\}$ is a standard Wiener process and $\|/q\|$ is the weighted sup-norm metric for functions in $D[0,1]$;

(b) On an appropriate probability space for $X_1, X_2, \ldots,$ a standard Wiener process $\{W(t), 0 \le t < \infty\}$ can be constructed in such a way that

$$\sup_{0<t\le 1} \left| U^{nor}_{[nt]} - \frac{W(nt)}{\sqrt{n}} \right| / q(t) = o_P(1),$$

if and only if $I(q,c) < \infty$ for all $c > 0$.

Theorem 3.3.2. *Let $q \in Q$ and assume*

$$E\big(\big|h(X_1, \ldots, X_m)\big|^{\frac{4}{3}} \log |h(X_1, \ldots, X_m)|\big) < \infty.$$

and that $\tilde{h}_1(X_1) \in DAN$. Then, as $n \to \infty$, we have

(a) $\frac{[nt]}{m\ell(n)\sqrt{n}}(U_{[nt]} - \theta) \Rightarrow W(t)$ on $(D, \mathfrak{D}, \|/q\|)$ if and only if $I(q,c) < \infty$ for all $c > 0$, where $\{W(t), 0 \le t \le 1\}$ is a standard Wiener process and $\|/q\|$ is the weighted sup-norm metric for functions in $D[0,1]$;

(b) On an appropriate probability space for $X_1, X_2, \ldots,$ a standard Wiener process $\{W(t), 0 \le t < \infty\}$ can be constructed in such a way that

$$\sup_{0<t\le 1} \left| \frac{[nt]}{m\ell(n)\sqrt{n}}(U_{[nt]} - \theta) - \frac{W(nt)}{\sqrt{n}} \right| / q(t) = o_P(1),$$

if and only if $I(q,c) < \infty$ for all $c > 0$.

Remark 3.3.1. *Note that taking $q(t) = 1$ results in finiteness of $I(q,c)$ for all $c > 0$, i.e., Theorems 3.3.1 and 3.3.2 remain valid for non-weighted pseudo Self-normalized U-statistic processes. Moreover, in this case, $\|/q\|$-metric will coincide with the usual* *sup-norm metric and the notion $\Rightarrow$ of part (b) of Theorems 3.3.1 and 3.3.2, as it is defined in Lemmas 2.3.5 and 2.3.6, will coincide with the the convergence in distribution of functionals definition of weak convergence on $D[0,1]$ with respect to the sup-norm metric.*

In the same spirit as that of Lemma 2.3.7, for a larger class of weight functions that imply the finiteness of $I(q,c)$ and characterized by Lemma 2.3.3 and part (b) of Lemma 2.3.4, the following Theorems 3.3.3 is a U-statistic version of Lemma 2.3.7.

Theorem 3.3.3. *Let $\{W(t), 0 \le t < \infty\}$ be a standard Wiener process, $\tilde{h}_1(X_1) \in DAN$ and*

$$E\big(\big|h(X_1,\ldots,X_m)\big|^{\frac{4}{3}} \log |h(X_1,\ldots,X_m)|\big) < \infty.$$

If $q \in Q$ and $q(t)$ is nondecreasing on $(0,1]$, then as $n \to \infty$,

$$\sup_{0<t\le 1} \big|U_{[nt]}^{nor}\big|/q(t) \longrightarrow_d \sup_{0<t\le 1} |W(t)|/q(t)$$

if and only if $I(q,c) < \infty$ for some $c > 0$. Consequently as $n \to \infty$ we have

$$\sup_{0<t\le 1} \big|U_{[nt]}^{nor}\big|/(t \log\log(\frac{1}{t}))^{\frac{1}{2}} \longrightarrow_d \sup_{0<t\le 1} |W(t)|/(t \log\log(\frac{1}{t}))^{\frac{1}{2}}.$$

Note that Theorem 3.3.3 keeps holding true if we replace V_n by B_n.

Proof of Theorems 3.3.1, 3.3.2 and 3.3.3

In view of reading of conclusion of Raikov's theorem as in (3.2.2) in this context, we have the the following equivalencies.

Lemmas 3.2.1 is equivalent to Lemma 3.2.2,

Theorem 3.3.1 is equivalent to Theorem 3.3.2.

Thus, in light of Lemma 3.2.2, we shall only give the proof of Theorem 3.3.2, and that of Theorem 3.3.3. It can be readily seen that in order to prove Theorems 3.3.2 and 3.3.3, it suffices to prove the following Theorem 3.3.4.

Theorem 3.3.4. *Let $q \in Q$ and assume*

$$E\big(|h(X_1,\ldots,X_m)|^{\frac{4}{3}}\log|h(X_1,\ldots,X_m)|\big) < \infty.$$

and that $\tilde{h}_1(X_1) \in DAN$. Then, as $n \to \infty$ we have

$$\sup_{0<t\leq 1}\Big|\frac{[nt]}{m\ell(n)\,\sqrt{n}}\,(U_{[nt]}-\theta) \;-\; \frac{\sum_{i=1}^{[nt]}\tilde{h}_1(X_i)}{\ell(n)\sqrt{n}}\Big|/q(t) = o_P(1).$$

Proof of Theorem 3.3.4

To establish Theorem 3.3.4, we first observe that

$$\sup_{0<t\leq 1}\Big|\frac{[nt]}{m\ell(n)\,\sqrt{n}}\,(U_{[nt]}-\theta) - \frac{\sum_{i=1}^{[nt]}\tilde{h}_1(X_i)}{\ell(n)\sqrt{n}}\Big|/q(t)$$

$$\leq \sup_{0<t\leq\frac{m}{n}}\Big|\frac{[nt]}{m\,\ell(n)\,\sqrt{n}}\,(U_{[nt]}-\theta) - \frac{\sum_{i=1}^{[nt]}\tilde{h}_1(X_i)}{\ell(n)\sqrt{n}}\Big|/q(t)$$

$$+\sup_{\frac{m}{n}<t\leq 1}\Big|\frac{[nt]}{m\,\ell(n)\,\sqrt{n}}\,(U_{[nt]}-\theta) \;-\; \frac{\sum_{i=1}^{[nt]}\tilde{h}_1(X_i)}{\ell(n)\sqrt{n}}\Big|/q(t). \tag{3.3.1}$$

But as $n \to \infty$, definition of $U_{[nt]}^{nor}$ and part (c) of Lemma 3.2.2 imply that

$$\sup_{0<t\leq\frac{m}{n}}\Big|\frac{[nt]}{\ell(n)\,\sqrt{n}}\,(U_{[nt]}-\theta) - \frac{\sum_{i=1}^{[nt]}\tilde{h}_1(X_i)}{\ell(n)\,\sqrt{n}}\Big|/q(t)$$

$$\leq \sup_{0<t<\frac{m}{n}}\Big|\frac{\sum_{i=1}^{[nt]}\tilde{h}_1(X_i)}{\ell(n)\sqrt{n}}\Big|/q(t)$$

$$+\Big|\frac{m}{m\,\ell(n)\,\sqrt{n}}\,(h(X_1,\ldots,X_m)-\theta) - \frac{\sum_{i=1}^{m}\tilde{h}_1(X_i)}{\ell(n)\,\sqrt{n}}\Big|/q(\frac{m}{n}) = o_P(1).$$

Therefore, according to the latter relation and (3.3.1), in order to prove Theorems 3.3.4 and 3.3.2, we need to show that

$$\sup_{\frac{m}{n}<t\le 1} \Big| \frac{[nt]}{m\,\ell(n)\,\sqrt{n}}\,(U_{[nt]}-\theta) \; - \; \frac{\sum_{i=1}^{[nt]} \tilde{h}_1(X_i)}{\ell(n)\sqrt{n}} \Big| / q(t) = o_P(1). \qquad (3.3.2)$$

Since it is true that

$$\sum_{1\le i_1<\ldots<i_m\le [nt]} (\tilde{h}_1(X_{i_1})+\ldots+\tilde{h}_1(X_{i_m})) = \frac{m}{[nt]}\binom{[nt]}{m}\sum_{i=1}^{[nt]} \tilde{h}_1(X_i),$$

it becomes clear that in order to establish (3.3.2), it will be enough to prove the following Proposition 3.3.1.

Proposition 3.3.1. *Let $q \in Q$. If*

$$E\big(|h(X_1,\ldots,X_m)|^{\frac{4}{3}}\log|h(X_1,\ldots,X_m)|\big) < \infty.$$

Then, as $n \to \infty$

$$\frac{n^{-\frac{1}{2}}}{\ell(n)} \sup_{\frac{m}{n}<t\le 1} \Big| \frac{[nt]}{\binom{[nt]}{m}} \sum_{1\le i_1<\ldots<i_m\le [nt]} (h(X_{i_1},\ldots,X_{i_m})-\theta-\tilde{h}_1(X_{i_1})-\ldots-\tilde{h}_1(X_{i_m}))\Big|/q(t)$$

$$= o_P(1).$$

Proof of Proposition 3.3.1

Without loss of generality assume that $\theta = 0$. On taking n large enough, let $\delta \in (\frac{m}{n}, 1]$ be small enough so that $q(t)$ is nondecreasing on $(0, \delta)$. Observe that

$$\frac{n^{-\frac{1}{2}}}{\ell(n)} \sup_{\frac{m}{n}<t\le 1} \Big| \frac{[nt]}{\binom{[nt]}{m}} \sum_{1\le i_1<\ldots<i_m\le [nt]} (h(X_{i_1},\ldots,X_{i_m})-\tilde{h}_1(X_{i_1})-\ldots-\tilde{h}_1(X_{i_m}))\Big|/q(t)$$

$$\leq \frac{n^{-\frac{1}{2}}}{\ell(n)} \sup_{\frac{m}{n}<t<\delta} \Big| \frac{[nt]}{\binom{[nt]}{m}} \sum_{1\leq i_1<\ldots<i_m\leq[nt]} (h(X_{i_1},\ldots,X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m}))\Big|/q(t)$$

$$+\frac{n^{-\frac{1}{2}}}{\ell(n)} \sup_{\delta\leq t\leq 1} \Big| \frac{[nt]}{\binom{[nt]}{m}} \sum_{1\leq i_1<\ldots<i_m\leq[nt]} (h(X_{i_1},\ldots,X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m}))\Big|/q(t)$$

$$:= I_1(n,t) + I_2(n,t). \tag{3.3.3}$$

To prove Proposition 3.3.1, it will be enough to show asymptotic negligibility of both $I_1(n,t)$ and $I_2(n,t)$ in probability.

To deal with $I_1(n,t)$ write

$$I_1(n,t) \leq \sup_{\frac{m}{n}<t<\delta} \frac{[nt](nt)^{\frac{-1}{2}}}{\binom{[nt]}{m}\ell(n)} \Big| \sum_{1\leq i_1<\ldots<i_m\leq[nt]} (h(X_{i_1},\ldots,X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m}))\Big|$$

$$\times \sup_{\frac{m}{n}<t<\delta} \frac{t^{\frac{1}{2}}}{q(t)}$$

$$\leq \sup_{\frac{m}{n}\leq t\leq 1} \frac{[nt]\ [nt]^{-\frac{1}{2}}}{\binom{[nt]}{m}\ell(n)} \Big| \sum_{1\leq i_1<\ldots<i_m\leq[nt]} (h(X_{i_1},\ldots,X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m}))\Big|$$

$$\times \sup_{\frac{m}{n}<t<\delta} \frac{t^{\frac{1}{2}}}{q(t)}.$$

The last relation suggests that by virtue of Lemma 2.3.3 and from the fact that $\ell(n)$ is a slowly varying function at infinity, for large n, we have that $\ell(n)\sqrt{n} \geq \sqrt{n}$. To prove $I_1(n,t) = o_P(1)$ it suffices to show that

$$\sup_{\frac{m}{n}\leq t\leq 1} \frac{[nt]^{\frac{1}{2}}}{\binom{[nt]}{m}} \Big| \sum_{1\leq i_1<\ldots<i_m\leq[nt]} (h(X_{i_1},\ldots,X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m}))\Big|$$

$$= O_P(1). \tag{3.3.4}$$

To establish (3.3.4), for $i_1 < \ldots < i_m$, consider the following truncation setup:

$$H_{im}^{tr}(X_{i_1}, \ldots, X_{i_m}) := h(X_{i_1}, \ldots, X_{i_m})\, \mathbf{1}_{(|h| \le i_m^{\frac{3}{2}})} - E(h(X_{i_1}, \ldots, X_{i_m})\, \mathbf{1}_{(|h| \le i_m^{\frac{3}{2}})}),$$

$$H_{im}^{ta}(X_{i_1}, \ldots, X_{i_m}) := h(X_{i_1}, \ldots, X_{i_m})\, \mathbf{1}_{(|h| > i_m^{\frac{3}{2}})} - E(h(X_{i_1}, \ldots, X_{i_m})\, \mathbf{1}_{(|h| > i_m^{\frac{3}{2}})}),$$

$$g_{im}^{tr}(X_{i_1}, \ldots, X_{i_m}) := H_{im}^{tr}(X_{i_1}, \ldots, X_{i_m}) - E(H_{im}^{tr}(X_{i_1}, \ldots, X_{i_m})|X_{i_1})$$

$$- \ldots - E(H_{im}^{tr}(X_{i_1}, \ldots, X_{i_m})|X_{i_m}),$$

$$g_{im}^{ta}(X_{i_1}, \ldots, X_{i_m}) := H_{im}^{ta}(X_{i_1}, \ldots, X_{i_m}) - E(H_{im}^{ta}(X_{i_1}, \ldots, X_{i_m})|X_{i_1})$$

$$- \ldots - E(H_{im}^{ta}(X_{i_1}, \ldots, X_{i_m})|X_{i_m}). \tag{3.3.5}$$

Having the above setup, to prove (3.3.4), we proceed by stating and proving the following Proposition 3.3.2.

Proposition 3.3.2. *If $E|h(X_1, \ldots, X_m)|^{\frac{4}{3}} < \infty$. Then, as $n \to \infty$, we have*

$$\max_{m \le K \le n} \frac{K^{\frac{1}{2}}}{\binom{K}{m}} \Big| \sum_{1 \le i_1 < \ldots < i_m \le K} g_{im}^{ta}(X_{i_1}, \ldots, X_{i_m}) \Big| = O_P(1) \tag{3.3.6}$$

and

$$\max_{m \le K \le n} \frac{K^{\frac{1}{2}}}{\binom{K}{m}} \Big| \sum_{1 \le i_1 < \ldots < i_m \le K} g_{im}^{tr}(X_{i_1}, \ldots, X_{i_m}) \Big| = O_P(1). \tag{3.3.7}$$

Proof of Proposition 3.3.2

For throughout use let A be a positive constant which may be different as each stage. To prove (3.3.6) we first represent the statistic $\sum_{1 \le i_1 < \ldots < i_m \le K} g_{im}^{ta}(X_{i_1}, \ldots, X_{i_m})$ in terms of $2^m - m - 1$ sums which their summand posses the property of complete degeneracy as follows.

$$\sum_{1\le i_1<\ldots<i_m\le K} g^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})$$

$$= \sum_{1\le i_1<\ldots<i_m\le K} \Big\{ \sum_{d=2}^{m}(-1)^{m-d} \sum_{1\le j_1<\ldots<j_d\le m} E(g^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})\big|X_{i_{j_1}},\ldots,X_{i_{j_d}})$$

$$+\sum_{c=2}^{m-1} \sum_{1\le k_1<\ldots<k_c\le m} \sum_{d=2}^{c}(-1)^{c-d} \sum_{1\le j_1<\ldots<j_d\le c} E(g^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})\big|X_{i_{k_{j_1}}},\ldots,X_{i_{k_{j_d}}})\Big\}$$

$$:= \sum_{1\le i_1<\ldots<i_m\le K} V^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})$$

$$+\sum_{c=2}^{m-1} \sum_{1\le k_1<\ldots<k_c\le m} \sum_{1\le i_1<\ldots<i_m\le K} V^{ta}_{i_m}(X_{i_{k_1}},\ldots,X_{i_{k_c}}).$$

In view of the latter setup to prove (3.3.6) we need to show that

$$\max_{m\le K\le n} \frac{K^{\frac{1}{2}}}{\binom{K}{m}} \Big| \sum_{1\le i_1<\ldots<i_m\le K} V^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})\Big| = O_P(1) \tag{3.3.8}$$

and for, $m\ge 3$, $c=2,\ldots,m-1$ and $1\le k_1<\ldots<k_c\le m$,

$$\max_{m\le K\le n} \frac{K^{\frac{1}{2}}}{\binom{K}{m}} \Big| \sum_{1\le i_1<\ldots<i_m\le K} V^{ta}_{i_m}(X_{i_{k_1}},\ldots,X_{i_{k_c}})\Big| = O_P(1). \tag{3.3.9}$$

Due to similarity we shall only give the proof of (3.3.8).

Noting that $\frac{K^{\frac{1}{2}}}{\binom{K}{m}}$ is decreasing in K, for $M>0$, we write,

$$P(\max_{m\le K\le n} \frac{K^{\frac{1}{2}}}{\binom{K}{m}} \Big| \sum_{1\le i_1<\ldots<i_m\le K} V^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})\Big| > M)$$

$$\le P(\max_{m\le K\le n} \sum_{i_m=m}^{K} \frac{i_m^{\frac{1}{2}}}{\binom{i_m}{m}} \Big| \sum_{1\le i_1<\ldots<i_{m-1}<i_m} V^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})\Big| > M)$$

$$\leq M^{-1}E\Big(\max_{m\leq K\leq n}\sum_{i_m=m}^{K}\frac{i_m^{\frac{1}{2}}}{\binom{i_m}{m}}\,\Big|\sum_{1\leq i_1<\ldots<i_{m-1}<i_m}V_{i_m}^{ta}(X_{i_1},\ldots,X_{i_m})\Big|\Big)$$

$$\leq M^{-1}E\Big(\sum_{i_m=m}^{\infty}\frac{i_m^{\frac{1}{2}}}{\binom{i_m}{m}}\,\Big|\sum_{1\leq i_1<\ldots<i_{m-1}<i_m}V_{i_m}^{ta}(X_{i_1},\ldots,X_{i_m})\Big|\Big)$$

$$\leq AM^{-1}\sum_{i_m=m}^{\infty}\frac{i_m^{\frac{1}{2}}}{\binom{i_m}{m}}\binom{i_m}{m-1}E\big(|h(X_{i_1},\ldots,X_{i_m})|\,\mathbf{1}_{(|h|>i_m^{\frac{3}{2}})}\big)$$

$$= AM^{-1}\sum_{i_m=m}^{\infty}\frac{i_m^{\frac{1}{2}}}{i_m-m+1}E\big(|h(X_{i_1},\ldots,X_{i_m})|\,\mathbf{1}_{(|h|>i_m^{\frac{3}{2}})}\big)$$

$$\leq AM^{-1}\sum_{i_m=m}^{\infty}i_m^{-\frac{1}{2}}E\big(|h(X_{i_1},\ldots,X_{i_m})|\,\mathbf{1}_{(|h|>i_m^{\frac{3}{2}})}\big)$$

$$= AM^{-1}\sum_{i_m=m}^{\infty}i_m^{-\frac{1}{2}}E\big(|h(X_1,\ldots,X_m)|\mathbf{1}_{(\cup_{j=i_m}^{\infty}(j^{\frac{3}{2}}<|h|\leq(j+1)^{\frac{3}{2}}))}\big)$$

$$\leq AM^{-1}\sum_{j=m}^{\infty}E\big(|h(X_1,\ldots,X_m)|\mathbf{1}_{(j^{\frac{3}{2}}<|h|\leq(j+1)^{\frac{3}{2}})}\big)\sum_{i_m=m}^{j}i_m^{-\frac{1}{2}}$$

$$\leq AM^{-1}\sum_{j=m}^{\infty}E\big(|h(X_1,\ldots,X_m)|^{\frac{4}{3}}\mathbf{1}_{(j^{\frac{3}{2}}<|h|\leq(j+1)^{\frac{3}{2}})}\big)$$

$$\leq AM^{-1}E|h(X_1,\ldots,X_m)|^{\frac{4}{3}}<\infty.$$

This completes the proof of (3.3.8) and that of (3.3.6).

Now we give the proof of (3.3.7). Similarly to what we had in the proof of (3.3.6), we write

$$\sum_{1\leq i_1<\ldots<i_m\leq K}g_{i_m}^{tr}(X_{i_1},\ldots,X_{i_m})$$

$$=\sum_{1\leq i_1<\ldots<i_m\leq K}\Big\{\sum_{d=2}^{m}(-1)^{m-d}\sum_{1\leq j_1<\ldots<j_d\leq m}E(g_{i_m}^{tr}(X_{i_1},\ldots,X_{i_m})|X_{i_{j_1}},\ldots,X_{i_{j_d}})$$

$$+\sum_{c=2}^{m-1}\sum_{1\le k_1<\ldots<k_c\le m}\sum_{d=2}^{c}(-1)^{c-d}\sum_{1\le j_1<\ldots<j_d\le c}E(g_{i_m}^{tr}(X_{i_1},\ldots,X_{i_m})|X_{i_{k_{j_1}}},\ldots,X_{i_{k_{j_d}}})\}$$

$$:=\sum_{1\le i_1<\ldots<i_m\le K}V_{i_m}^{tr}(X_{i_1},\ldots,X_{i_m})$$

$$+\sum_{c=2}^{m-1}\sum_{1\le k_1<\ldots<k_c\le m}\sum_{1\le i_1<\ldots<i_m\le K}V_{i_m}^{tr}(X_{i_{k_1}},\ldots,X_{i_{k_c}}).$$

Therefore, to establish (3.3.7) we need to show that

$$\max_{m\le K\le n}\frac{K^{\frac{1}{2}}}{\binom{K}{m}}\sum_{1\le i_1<\ldots<i_m\le K}V_{i_m}^{tr}(X_{i_1},\ldots,X_{i_m})=O_P(1) \tag{3.3.10}$$

and for, $m\ge 3$, $c=2,\ldots,m-1$ and $1\le k_1<\ldots<k_c\le m$,

$$\max_{m\le K\le n}\frac{K^{\frac{1}{2}}}{\binom{K}{m}}\sum_{1\le i_1<\ldots<i_m\le K}V_{i_m}^{tr}(X_{i_{k_1}},\ldots,X_{i_{k_c}})=O_P(1). \tag{3.3.11}$$

To prove (3.3.10), employing Proposition 1.8.3 and Lemma 1.11.1 followed by an application of Proposition 1.8.2, for $M>0$ we write

$$P(\max_{m\le K\le n}\frac{K^{\frac{1}{2}}}{\binom{K}{m}}\,|\sum_{1\le i_1<\ldots<i_m\le K}V_{i_m}^{tr}(X_{i_1},\ldots,X_{i_m})|>M)$$

$$\le M^{-2}\frac{n}{\binom{n}{m}^2}E(\sum_{1\le i_1<\ldots<i_m\le n}V_{i_m}^{tr}(X_{i_1},\ldots,X_{i_m}))^2$$

$$+M^{-2}\sum_{K=m}^{n-1}(\frac{K}{\binom{K}{m}^2}-\frac{K+1}{\binom{K+1}{m}^2})E(\sum_{1\le i_1<\ldots<i_m\le K}V_{i_m}^{tr}(X_{i_1},\ldots,X_{i_m}))^2$$

$$\le AM^{-2}\frac{n}{\binom{n}{m}^2}\sum_{i_m=m}^{n}\binom{i_m}{m-1}E(h^2(X_1,\ldots,X_{m-1},X_{i_m})\mathbf{1}_{(|h|\le i_m^{\frac{3}{2}})})$$

$$+AM^{-2}\sum_{K=m}^{n-1}\frac{2m}{\binom{K}{m}^2}\sum_{i_m=m}^{K}\binom{i_m}{m-1}E(h^2(X_1,\ldots,X_{m-1},X_{i_m})\mathbf{1}_{(|h|\le i_m^{\frac{3}{2}})})$$

$$\le AM^{-2}\frac{nm}{\binom{n}{m}(n-m+1)}\sum_{i_m=m}^{n}E(h^2(X_1,\ldots,X_{m-1},X_{i_m})\mathbf{1}_{(|h|\le i_m^{\frac{3}{2}})})$$

$$+AM^{-2}\sum_{K=m}^{n-1}E(h^2(X_1,\ldots,X_{m-1},X_{i_m})\mathbf{1}_{(|h|\le K^{\frac{3}{2}})}\frac{2m^2}{\binom{K}{m}(K-m+1)}\sum_{i_m=m}^{K}1$$

$$\le AM^{-2}\sum_{i_m=m}^{n}i_m^{-2}E(h^2(X_1,\ldots,X_{m-1},X_{i_m})\mathbf{1}_{(|h|\le i_m^{\frac{3}{2}})})$$

$$+AM^{-2}\sum_{K=m}^{n-1}K^{-2}E(h^2(X_1,\ldots,X_{m-1},X_{i_m})\mathbf{1}_{(|h|\le K^{\frac{3}{2}})}$$

$$\le AM^{-2}\sum_{K=m}^{\infty}K^{-2}E(h^2(X_1,\ldots,X_{m-1},X_{i_m})\mathbf{1}_{(|h|\le K^{\frac{3}{2}})}$$

$$= AM^{-2}\sum_{K=m}^{\infty}K^{-2}E(h^2(X_1,\ldots,X_m)\mathbf{1}_{(\cup_{i=1}^{K}(i-1)^{\frac{3}{2}}<|h|\le i^{\frac{3}{2}})})$$

$$= AM^{-2}\sum_{i=1}^{\infty}E(h^2(X_1,\ldots,X_m)\mathbf{1}_{((i-1)^{\frac{3}{2}}<|h|\le i^{\frac{3}{2}})})\sum_{K=i}^{\infty}K^{-2}$$

$$\le AM^{-2}\sum_{i=1}^{\infty}i^{-1}E(h^2(X_1,\ldots,X_m)\mathbf{1}_{((i-1)^{\frac{3}{2}}<|h|\le i^{\frac{3}{2}})})$$

$$\le AM^{-2}\,E\left(\left|h(X_1,\ldots,X_m)\right|^{\frac{4}{3}}\,\mathbf{1}_{(\cup_{i=1}^{\infty}(i-1)^{\frac{3}{2}}<|h|\le i^{\frac{3}{2}})}\right)$$

$$\le AM^{-2}\,E|h(X_1,\ldots,X_m)|^{\frac{4}{3}}<\infty.$$

Now the proof of (3.3.10) is complete.

Due to similarity, to establish (3.3.11), we shall only state the proof for $k_1 = 1,\ldots,k_c = c$, where $c = 2,\ldots,m-1$ and $m \ge 3$. But first observe that

$$\sum_{1\le i_1<\ldots<i_m\le K}V_{i_m}^{tr}(X_{i_1},\ldots,X_{i_c})$$

$$= \binom{K-m+c}{m-c-1} \sum_{i_m=m}^{K} \sum_{1 \le i_1 < \ldots < i_c < i_m} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_c})$$

$$\le \binom{K}{m-c-1} \sum_{i_m=m}^{K} \sum_{1 \le i_1 < \ldots < i_c < i_m} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_c}).$$

Once again an application of Lemma 1.11.1 followed by an application of Proposition 1.8.1, yield

$$P\Big(\max_{m \le K \le n} \frac{K^{\frac{1}{2}}}{\binom{K}{m}} \Big| \sum_{1 \le i_1 < \ldots < i_m \le K} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_c})\Big| > M\Big)$$

$$\le M^{-2}\Big\{\frac{n}{\binom{n}{m}^2} E\Big(\binom{n}{m-c-1} \sum_{i_m=m}^{n} \sum_{1 \le i_1 < \ldots < i_c < i_m} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_c})\Big)^2$$

$$+ \sum_{K=m}^{n-1} \Big(\frac{K}{\binom{K}{m}^2} - \frac{K+1}{\binom{K+1}{m}^2}\Big) E\Big(\sum_{i_m=m}^{K} \sum_{1 \le i_1 < \ldots < i_c < i_m} \binom{K}{m-c-1} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_c})\Big)^2\Big\}$$

$$\le AM^{-2} \frac{n^2}{\binom{n}{m}^2} \binom{n}{m-c-1}^2 \binom{n}{c} \sum_{i_m=m}^{n} E(h^2(X_1, \ldots, X_m) \mathbf{1}_{(|h| \le i_m^{\frac{3}{2}})})$$

$$+ AM^{-2} \sum_{K=m}^{n-1} \Big(\frac{K}{\binom{K}{m}^2} - \frac{K+1}{\binom{K+1}{m}^2}\Big) \binom{K}{m-c-1}^2 \binom{K}{c} K \sum_{i_m=m}^{K} E(h^2(X_1, \ldots, X_m) \mathbf{1}_{(|h| \le i_m^{\frac{3}{2}})})$$

$$\le AM^{-2} \sum_{i_m=m}^{n} i_m^{-2} E(h^2(X_1, \ldots, X_m) \mathbf{1}_{(|h| \le i_m^{\frac{3}{2}})})$$

$$+ AM^{-2} \sum_{K=m}^{n-1} K^{-3} E(h^2(X_1, \ldots, X_m) \mathbf{1}_{(|h| \le K^{\frac{3}{2}})}) \sum_{i_m=m}^{K} 1$$

$$\le AM^{-2} E|h(X_1, \ldots, X_m)|^{\frac{4}{3}} < \infty.$$

The latter relation completes the proof of (3.3.11) and that of Proposition 3.3.2. Hence $I_1(n,t) = o_P(1)$.

By virtue of our notion of $I_1(n,t)$ and $I_2(n,t)$, so far, we have showed that

$I_1(n,t) = o_P(1)$. To complete the proof of Proposition 3.3.1 we need to show that $I_2(n,t) = o_P(1)$. But observe that for $I_2(n,t)$ we can write

$$I_2(n,t) \le \frac{n^{-\frac{1}{2}}}{\ell(n)} \sup_{\delta \le t \le 1} \frac{[nt]}{\binom{[nt]}{m}} \Big| \sum_{1 \le i_1 < \ldots < i_m \le [nt]} (h(X_{i_1}, \ldots, X_{i_m}) - \theta - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m})) \Big|$$

$$\times \sup_{\delta \le t \le 1} \frac{1}{q(t)}.$$

And since $\sup_{\delta \le t \le 1} \dfrac{1}{q(t)} = O(1)$, in order to show that $I_2(n,t) = o_P(1)$, we only need to show that

$$\frac{n^{-\frac{1}{2}}}{\ell(n)} \sup_{\frac{m}{n} \le t \le 1} \Big| \frac{[nt]}{\binom{[nt]}{m}} \sum_{1 \le i_1 < \ldots < i_m \le [nt]} (h(X_{i_1}, \ldots, X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m})) \Big|$$

$$= o_P(1). \tag{3.3.12}$$

Remark 3.3.1. *Under the condition that*

$$E\big(\left|h(X_1, \ldots, X_m)\right|^{\frac{4}{3}} \log |h(X_1, \ldots, X_m)| \big) < \infty,$$

Nasari in [31] gives a proof for (3.3.12). Here, for the sake of consistency in our proofs, we shall establish this relation in a different way.

To prove the (3.3.12), we first consider a slightly different truncation setup than (3.3.5) as follows.

$$H_n^{tr}(X_{i_1}, \ldots, X_{i_m}) := h(X_{i_1}, \ldots, X_{i_m})\, \mathbf{1}_{(|h| \le n^{\frac{3}{2}})} - E(h(X_{i_1}, \ldots, X_{i_m})\, \mathbf{1}_{(|h| \le n^{\frac{3}{2}})}),$$

$$H_n^{ta}(X_{i_1}, \ldots, X_{i_m}) := h(X_{i_1}, \ldots, X_{i_m})\, \mathbf{1}_{(|h| > n^{\frac{3}{2}})} - E(h(X_{i_1}, \ldots, X_{i_m})\, \mathbf{1}_{(|h| > n^{\frac{3}{2}})}),$$

$$g_n^{tr}(X_{i_1}, \ldots, X_{i_m}) := H_n^{tr}(X_{i_1}, \ldots, X_{i_m}) - E(H_n^{tr}(X_{i_1}, \ldots, X_{i_m}) | X_{i_1})$$

$$-\ldots - E(H_n^{tr}(X_{i_1},\ldots,X_{i_m})|X_{i_m}),$$

$$g_n^{ta}(X_{i_1},\ldots,X_{i_m}) := H_n^{ta}(X_{i_1},\ldots,X_{i_m}) - E(H_n^{ta}(X_{i_1},\ldots,X_{i_m})|X_{i_1})$$

$$-\ldots - E(H_n^{ta}(X_{i_1},\ldots,X_{i_m})|X_{i_m}). \tag{3.3.13}$$

By virtue of the above setup to prove (3.3.12) we proceed by proving the following result.

Proposition 3.3.3. *If*

$$E\left(\left|h(X_1,\ldots,X_m)\right|^{\frac{4}{3}}\log\left|h(X_1,\ldots,X_m)\right|\right) < \infty.$$

Then, as $n \to \infty$, we have

$$n^{-\frac{1}{2}} \max_{m\le K\le n} \frac{K}{\binom{K}{m}} \Big| \sum_{1\le i_1<\ldots<i_m\le K} g_n^{ta}(X_{i_1},\ldots,X_{i_m})\Big| = o_P(1) \tag{3.3.14}$$

and

$$n^{-\frac{1}{2}} \max_{m\le K\le n} \frac{K}{\binom{K}{m}} \Big| \sum_{1\le i_1<\ldots<i_m\le K} g_n^{tr}(X_{i_1},\ldots,X_{i_m})\Big| = o_P(1). \tag{3.3.15}$$

Proof of Proposition 3.3.3

Once again to prove (3.3.14), we add and subtract terms to get:

$$\sum_{1\le i_1<\ldots<i_m\le K} g_n^{ta}(X_{i_1},\ldots,X_{i_m})$$

$$= \sum_{1\le i_1<\ldots<i_m\le k} \Big\{ \sum_{d=2}^{m}(-1)^{m-d} \sum_{1\le j_1<\ldots<j_d\le m} E(g_n^{ta}(X_{i_1},\ldots,X_{i_m})\big|X_{i_{j_1}},\ldots,X_{i_{j_d}})$$

$$+\sum_{c=2}^{m-1} \sum_{1\le k_1<\ldots<k_c\le m} \sum_{d=2}^{c}(-1)^{c-d} \sum_{1\le j_1<\ldots<j_d\le c} E(g_n^{ta}(X_{i_1},\ldots,X_{i_m})\big|X_{i_{k_{j_1}}},\ldots,X_{i_{k_{j_d}}})\Big\}$$

$$:= \sum_{1\le i_1<\ldots<i_m\le K} V_n^{ta}(X_{i_1},\ldots,X_{i_m})$$

$$+\sum_{c=2}^{m-1}\sum_{1\le k_1<\ldots<k_c\le m}\sum_{1\le i_1<\ldots<i_m\le K} V_n^{ta}(X_{i_{k_1}},\ldots,X_{i_{k_c}}).$$

(3.3.14) will follows if we show that

$$n^{-\frac{1}{2}}\max_{m\le K\le n}\frac{K}{\binom{K}{m}}\Big|\sum_{1\le i_1<\ldots<i_m\le K} V_n^{ta}(X_{i_1},\ldots,X_{i_m})\Big| = o_P(1) \tag{3.3.16}$$

and for, $m\ge 3$, $c=2,\ldots,m-1$ and $1\le k_1<\ldots<k_c\le m$,

$$n^{-\frac{1}{2}}\max_{m\le K\le n}\frac{K}{\binom{K}{m}}\Big|\sum_{1\le i_1<\ldots<i_m\le K} V_n^{ta}(X_{i_{k_1}},\ldots,X_{i_{k_c}})\Big| = o_P(1). \tag{3.3.17}$$

Having the similarity of the two we only give the proof of (3.3.16). Once again having the martingale property, we apply Lemma 1.11.1 for $\epsilon>0$ and arrive at:

$$P(n^{-\frac{1}{2}}\max_{m\le K\le n}\frac{K}{\binom{K}{m}}\Big|\sum_{1\le i_1<\ldots<i_m\le K} V_n^{ta}(X_{i_1},\ldots,X_{i_m})\Big|>\epsilon)$$

$$\le \epsilon^{-1}n^{-\frac{1}{2}}\frac{n}{\binom{n}{m}}E\big(\Big|\sum_{1\le i_1<\ldots<i_m\le n} V_n^{ta}(X_{i_1},\ldots,X_{i_m})\Big|\big)$$

$$+\epsilon^{-1}n^{-\frac{1}{2}}\sum_{K=1}^{n-1}\Big(\frac{K}{\binom{K}{m}}-\frac{K+1}{\binom{K+1}{m}}\Big)E\big(\Big|\sum_{1\le i_1<\ldots<i_m\le K} V_n^{ta}(X_{i_1},\ldots,X_{i_m})\Big|\big)$$

$$\le A\epsilon^{-1}n^{\frac{1}{2}}E\big(|h(X_1,\ldots,X_m)|\mathbf{1}_{(|h|>n^{\frac{3}{2}})}\big)$$

$$+A\epsilon^{-1}n^{-\frac{1}{2}}E\big(|h(X_1,\ldots,X_m)|\mathbf{1}_{(|h|>n^{\frac{3}{2}})}\big)\sum_{K=1}^{n-1}1$$

$$\le A\epsilon^{-1}E\big(|h(X_1,\ldots,X_m)|^{\frac{4}{3}}\mathbf{1}_{(|h|>n^{\frac{3}{2}})}\big)\to 0,\ as\ n\to\infty.$$

Now the proof of (3.3.16), (3.3.17) and that of (3.3.14).

To establish (3.3.15), similarly to the preceding case we represent the statistic in terms of complete degenerate ones as follows.

$$\sum_{1\le i_1<\ldots<i_m\le K} g_n^{tr}(X_{i_1},\ldots,X_{i_m})$$

$$= \sum_{1\le i_1<\ldots<i_m\le k} \Big\{ \sum_{d=2}^{m}(-1)^{m-d} \sum_{1\le j_1<\ldots<j_d\le m} E(g_n^{tr}(X_{i_1},\ldots,X_{i_m})|X_{i_{j_1}},\ldots,X_{i_{j_d}})$$

$$+\sum_{c=2}^{m-1} \sum_{1\le k_1<\ldots<k_c\le m} \sum_{d=2}^{c}(-1)^{c-d} \sum_{1\le j_1<\ldots<j_d\le c} E(g_n^{tr}(X_{i_1},\ldots,X_{i_m})|X_{i_{k_{j_1}}},\ldots,X_{i_{k_{j_d}}})\Big\}$$

$$:= \sum_{1\le i_1<\ldots<i_m\le K} V_n^{tr}(X_{i_1},\ldots,X_{i_m})$$

$$+\sum_{c=2}^{m-1} \sum_{1\le k_1<\ldots<k_c\le m} \sum_{1\le i_1<\ldots<i_m\le K} V_n^{tr}(X_{i_{k_1}},\ldots,X_{i_{k_c}}).$$

Therefore, (3.3.15) will follow if we show that as $n\to\infty$,

$$n^{-\frac{1}{2}} \max_{m\le K\le n} \frac{K}{\binom{K}{m}} \sum_{1\le i_1<\ldots<i_m\le K} V_n^{tr}(X_{i_1},\ldots,X_{i_m}) = o_P(1) \tag{3.3.18}$$

and for, $m\ge 3$, $c=2,\ldots,m-1$ and $1\le k_1<\ldots<k_c\le m$,

$$n^{-\frac{1}{2}} \max_{m\le K\le n} \frac{K}{\binom{K}{m}} \sum_{1\le i_1<\ldots<i_m\le K} V_n^{tr}(X_{i_{k_1}},\ldots,X_{i_{k_c}}) = o_P(1). \tag{3.3.19}$$

To establish (3.3.18), for $\epsilon>0$, we apply Lemma 1.11.1 and Proposition 1.8.1 to arrive at:

$$P(n^{-\frac{1}{2}} \max_{m\le K\le n} \frac{K}{\binom{K}{m}}\Big| \sum_{1\le i_1<\ldots<i_m\le K} V_n^{tr}(X_{i_1},\ldots,X_{i_m})\Big| > \epsilon)$$

$$\leq A\epsilon^{-2}\frac{n}{\binom{n}{m}}E\big(h^2(X_1,\ldots,X_m)\ \mathbf{1}_{(|h|\leq n^{\frac{3}{2}})}\big)$$

$$+A\epsilon^{-2}n^{-1}E\big(h^2(X_1,\ldots,X_m)\ \mathbf{1}_{(|h|\leq n^{\frac{3}{2}})}\big)\sum_{K=m}^{n-1}\Big(\frac{K^2}{\binom{K}{m}^2}-\frac{(K+1)^2}{\binom{K+1}{m}^2}\Big)\binom{K}{m}$$

$$\leq A\epsilon^{-2}n^{-1}E\big(h^2(X_1,\ldots,X_m)\ \mathbf{1}_{(|h|\leq n^{\frac{3}{2}})}\big)$$

$$+A\epsilon^{-2}n^{-1}E\big(h^2(X_1,\ldots,X_m)\ 1_{(|h|\leq n^{\frac{3}{2}})}\big)\sum_{K=m}^{n-1}\frac{K}{\binom{K}{m}} \tag{3.3.20}$$

First we show that $n^{-1}E\big(h^2(X_1,\ldots,X_m)\ 1_{(|h|\leq n^{\frac{3}{2}})}\big)=o(1)$. To do so, observe that

$$n^{-1}E\big(h^2(X_1,\ldots,X_m)\ 1_{(|h|\leq n^{\frac{3}{2}})}\big)$$

$$=n^{-1}E\big(h^2(X_1,\ldots,X_m)\ 1_{(|h|\leq n^{\frac{3}{2}})}\big)$$

$$\leq n^{-\frac{1}{3}}E\big(|h(X_1,\ldots,X_m)|^{\frac{4}{3}}1_{(|h|\leq n)}\big)+E\big(|h(X_1,\ldots,X_m)|^{\frac{4}{3}}1_{(n<|h|\leq n^{\frac{3}{2}})}\big)$$

$\longrightarrow 0$, as $n\to\infty$.

To show the asymptotic negligibility of the other term in (3.3.20) we write,

$$n^{-1}E\big(h^2(X_1,\ldots,X_m)\ \mathbf{1}_{(|h|\leq n^{\frac{3}{2}})}\big)\sum_{K=m}^{n-1}\frac{K}{\binom{K}{m}}$$

$$\leq n^{-1}\log(n)E\big(h^2(X_1,\ldots,X_m)\ \mathbf{1}_{(|h|\leq n^{\frac{3}{2}})}\big)$$

$$\leq n^{-\frac{1}{3}}\log(n)E\big(|h(X_1,\ldots,X_m)|^{\frac{4}{3}}\big)+E\big(|h(X_1,\ldots,X_m)|^{\frac{4}{3}}\log|h(X_1,\ldots,X_m)|\mathbf{1}_{(|h|>n)}\big)$$

$\longrightarrow 0$, as $n\to\infty$.

We now conclude (3.3.18). To complete the proof of (3.3.15) we now establish (3.3.19). Due to similarity we shall establish the latter only for the case $k_1=1,\ldots k_c=c$, where

$2 \leq c \leq m-1$. First note that

$$\begin{aligned}\sum_{1\leq i_1<\ldots<i_m\leq K} V_n^{tr}(X_{i_1},\ldots,X_{i_c}) &= \binom{K-m+c}{m-c} \sum_{1\leq i_1<\ldots<i_c\leq K} V_n^{tr}(X_{i_1},\ldots,X_{i_c}) \\ &\leq \binom{K}{m-c} \sum_{1\leq i_1<\ldots<i_c\leq K} V_n^{tr}(X_{i_1},\ldots,X_{i_c}).\end{aligned}$$

Now another application of Lemma 1.11.1 and Proposition 1.8.1, yield

$$P\big(n^{-\frac{1}{2}} \max_{m\leq K\leq n} \frac{K\binom{K}{m-c}}{\binom{K}{m}} \big| \sum_{1\leq i_1<\ldots<i_c\leq K} V_n^{tr}(X_{i_1},\ldots,X_{i_c})\big| > \epsilon\big)$$

$$\leq A\epsilon^{-2}\, \frac{n\binom{n}{m-c}^2\binom{n}{c}}{\binom{n}{m}^2}\; E\big(h^2(X_1,\ldots,X_m)\mathbf{1}_{(|h|\leq n^{\frac{3}{2}})}\big)$$

$$+A\epsilon^{-2}n^{-1}E\big(h^2(X_1,\ldots,X_m)\mathbf{1}_{(|h|\leq n^{\frac{3}{2}})}\big)\sum_{K=1}^{n-1}\Big(\frac{K^2\binom{K}{m-c}^2}{\binom{K}{m}^2} - \frac{(K+1)^2\binom{K+1}{m-c}^2}{\binom{K+1}{m}^2}\Big)\binom{K}{c}$$

$$\leq A\epsilon^{-2}\; n^{-c+1}E\big(h^2(X_1,\ldots,X_m)\mathbf{1}_{(|h|\leq n^{\frac{3}{2}})}\big)$$

$$+A\epsilon^{-2}n^{-1}E\big(h^2(X_1,\ldots,X_m)\mathbf{1}_{(|h|\leq n^{\frac{3}{2}})}\big)\sum_{K=1}^{n-1} K^{-c+1}$$

$$\leq A\epsilon^{-2}n^{-1}E\big(h^2(X_1,\ldots,X_m)\mathbf{1}_{(|h|\leq n^{\frac{3}{2}})}\big) + A\epsilon^{-2}n^{-1}\;\log(n)E\big(h^2(X_1,\ldots,X_m)\mathbf{1}_{(|h|\leq n^{\frac{3}{2}})}\big)$$

$\longrightarrow 0$, as $n \to \infty$.

The last relation completes the proof of (3.3.15) and that of Proposition 3.3.3. Now the proof of Theorem 3.3.4 and those of Theorems 3.3.2, 3.3.1 and 3.3.3 are complete.

□

3.4 Pseudo self-normalized U-type processes

Theorem 3.4.1. *Let $q \in Q$ and assume $\tilde{h}_1(X_1) \in DAN$ and*

$$E\big|h(X_1, \ldots, X_m)\big|^{\frac{4}{3}} < \infty.$$

Then, as $n \to \infty$, we have

(a) $\Psi_{[nt]}^{nor} \Rightarrow t^{m-1}\, W(t)$ on $(D, \mathfrak{D}, \|/q\|)$ if and only if $I(q,c) < \infty$ for all $c > 0$, where $\{W(t), 0 \le t \le 1\}$ is a standard Wiener process and $\|/q\|$ is the weighted sup-norm metric for functions in $D[0,1]$;

(b) On an appropriate probability space for $X_1, X_2, \ldots$, a standard Wiener process $\{W(t), 0 \le t < \infty\}$ can be constructed in such a way that as $n \to \infty$

$$\sup_{0<t\le 1} \Big|\Psi_{[nt]}^{nor} - \frac{\binom{[nt]}{m}}{[nt]\; n^{m-1}}\, \frac{W(nt)}{\sqrt{n}}\Big|/q(t) = o_P(1),$$

if and only if $I(q,c) < \infty$ for all $c > 0$.

Remark 3.4.1. *The statement (a) of Theorem 3.4.1 stands for the following functional central limit theorem on $(D, \mathfrak{D}, \|/q\|)$, where $\mathfrak{D}$ is the σ-field of subsets of $D = D[0,1]$ generated by its finite-dimensional subsets, and $\|/q\|$ stands for the weighted sup-norm metric in $D = D[0,1]$ with $q \in Q$ that is also càdlàg. With $\longrightarrow_d$ standing for convergence in distribution as $n \to \infty$, we have*

$$g\Big(\frac{\Psi_{[nt]}^{nor}}{V_n q(t)}\Big) \to_d g\Big(\frac{t^{m-1}\, W(t)}{q(t)}\Big)$$

for all càdlàg functions functions $q \in Q$, and for all $g : D = D[0,1] \to \mathbb{R}$ that are $(D, \mathfrak{D})$ measurable and $\|/q\|$-continuous or $\|/q\|$-continuous except at points forming a set of measure zero on $(D, \mathfrak{D})$, with respect to the measure generated by the process

$t^{m-1}\, W(t)$ *on the unit interval* $[0,1]$.

Theorem 3.4.2. *Let* $q \in Q$ *and assume* $\tilde{h}_1(X_1) \in DAN$ *and*

$$E\big|h(X_1,\ldots,X_m)\big|^{\frac{4}{3}} < \infty$$

Then, as $n \to \infty$, *we have*

$$(a)\frac{n^{-m+1}}{m\ell(n)\sqrt{n}} \sum_{1 \le i_1 < \ldots < i_m \le [nt]} (h(X_{i_1},\ldots,X_{i_m}) - \theta) \Rightarrow t^{m-1}\, W(t) \text{ on } (D, \mathfrak{D}, \|/q\|)$$

if and only if $I(q,c) < \infty$ *for all* $c > 0$, *where* $\{W(t), 0 \le t \le 1\}$ *is a standard Wiener process and* $\|/q\|$ *is the weighted sup-norm metric for functions in* $D[0,1]$*;*

(b) On an appropriate probability space for $X_1, X_2, \ldots,$ *a standard Wiener process* $\{W(t), 0 \le t < \infty\}$ *can be constructed in such a way that*

$$\sup_{0<t\le 1}\Big|\frac{n^{-m+1}}{m\ell(n)\sqrt{n}} \sum_{1 \le i_1 < \ldots < i_m \le [nt]} (h(X_{i_1},\ldots,X_{i_m}) - \theta) - \frac{\binom{[nt]}{m}}{[nt]\; n^{m-1}} \frac{W(nt)}{\sqrt{n}}\Big|/q(t) = o_P(1),$$

if and only if $I(q,c) < \infty$ *for all* $c > 0$.

We note in passing that the weak convergence in statement (a) of the preceding theorem is defined, mutatis mutandis, similar to the one is defined in Remark 3.4.1.

Theorem 3.4.3. *Let* $\{W(t), 0 \le t < \infty\}$ *be a standard Wiener process,*

$$E\big|h(X_1,\ldots,X_m)\big|^{\frac{4}{3}} < \infty$$

and $\tilde{h}_1(X_1) \in DAN$. *If* $q \in Q$ *and* $q(t)$ *is nondecreasing on* $(0,1]$, *then, as* $n \to \infty$,

$$\sup_{0<t\le 1}\big|\Psi^{nor}_{[nt]}\big|/q(t) \longrightarrow_d \sup_{0<t\le 1} t^{m-1}|W(t)|/q(t),$$

if and only if $I(q,c) < \infty$ for some $c > 0$. Consequently as $n \to \infty$ we have

$$\sup_{0<t\le 1} |\Psi^{nor}_{[nt]}|/(t\log\log(\frac{1}{t}))^{\frac{1}{2}} \longrightarrow_d \sup_{0<t\le 1} t^{m-1}|W(t)|/(t\log\log(\frac{1}{t}))^{\frac{1}{2}}.$$

Proof of Theorems 3.4.1, 3.4.2 and 3.4.3

Due to the equivalency of Theorems 3.4.1and 3.4.2, we shall only give the proof of Theorems 3.4.2 and 3.4.3. The proof of these theorems will follow by proving the following Theorem 3.4.4.

Theorem 3.4.4. *Let $q \in Q$. If $E|h(X_1,\ldots,X_m)|^{\frac{4}{3}} < \infty$ and $\tilde{h}_1(X_1) \in DAN$. Then, as $n \to \infty$, we have*

$$\sup_{0<t\le 1} |\frac{\sum_{1\le i_1<\ldots<i_m\le [nt]}(h(X_{i_1},\ldots,X_{i_m})-\theta)}{m\ n^{m-1}\ell(n)\ \sqrt{n}} - \frac{\binom{[nt]}{m}}{[nt]\ n^{m-1}}\frac{\sum_{i=1}^{[nt]}\tilde{h}_1(X_i)}{\ell(n)\sqrt{n}}|/q(t) = o_P(1).$$

Proof of Theorem 3.4.4

To prove Theorem 3.4.4, first observe that

$$\sup_{0<t\le 1} |\frac{\sum_{1\le i_1<\ldots<i_m\le [nt]}(h(X_{i_1},\ldots,X_{i_m})-\theta)}{m\ n^{m-1}\ell(n)\ \sqrt{n}} - \frac{\binom{[nt]}{m}}{[nt]\ n^{m-1}}\frac{\sum_{i=1}^{[nt]}\tilde{h}_1(X_i)}{\ell(n)\sqrt{n}}|/q(t)$$

$$\le \sup_{0<t\le \frac{m}{n}} |\frac{\sum_{1\le i_1<\ldots<i_m\le [nt]}(h(X_{i_1},\ldots,X_{i_m})-\theta)}{m\ n^{m-1}\ell(n)\ \sqrt{n}} - \frac{\binom{[nt]}{m}}{[nt]\ n^{m-1}}\frac{\sum_{i=1}^{[nt]}\tilde{h}_1(X_i)}{\ell(n)\sqrt{n}}|/q(t)$$
$$+ \sup_{\frac{m}{n}<t\le 1} |\frac{\sum_{1\le i_1<\ldots<i_m\le [nt]}(h(X_{i_1},\ldots,X_{i_m})-\theta)}{m\ n^{m-1}\ell(n)\ \sqrt{n}} - \frac{\binom{[nt]}{m}}{[nt]\ n^{m-1}}\frac{\sum_{i=1}^{[nt]}\tilde{h}_1(X_i)}{\ell(n)\sqrt{n}}|/q(t). \quad (3.4.1)$$

But as $n \to \infty$, Part (c) of Lemma 3.2.2 and some algebra imply that

$$\sup_{0<t\le \frac{m}{n}} |\frac{\sum_{1\le i_1<\ldots<i_m\le [nt]}(h(X_{i_1},\ldots,X_{i_m})-\theta)}{m\ n^{m-1}\ell(n)\ \sqrt{n}} - \frac{\binom{[nt]}{m}}{[nt]\ n^{m-1}}\frac{\sum_{i=1}^{[nt]}\tilde{h}_1(X_i)}{\ell(n)\sqrt{n}}|/q(t)$$

$$\leq \sup_{0<t<\frac{m}{n}} \Big| \frac{\sum_{i=1}^{[nt]} \tilde{h}_1(X_i)}{\ell(n)\sqrt{n}} \Big| / q(t) \times \sup_{0<t<\frac{m}{n}} \frac{\binom{[nt]}{m}}{[nt]\ n^{m-1}}$$

$$+ \Big| \frac{1}{m\ n^{m-1}\ell(n)\sqrt{n}} (h(X_1, \ldots, X_m) - \theta) - \frac{1}{m\ n^{m-1}} \frac{\sum_{i=1}^{m} \tilde{h}_1(X_i)}{\ell(n)\sqrt{n}} \Big| / q(\frac{m}{n}) = o_P(1).$$

Therefore, according to the latter relation and (3.4.1), in order to prove Theorem 3.4.4, we need to show that

$$\sup_{\frac{m}{n}<t\leq 1} \Big| \frac{\sum_{1\leq i_1<\ldots<i_m\leq [nt]} (h(X_{i_1}, \ldots, X_{i_m}) - \theta)}{m\ n^{m-1}\ell(n)\ \sqrt{n}} - \frac{\binom{[nt]}{m}}{[nt]\ n^{m-1}} \frac{\sum_{i=1}^{[nt]} \tilde{h}_1(X_i)}{\ell(n)\sqrt{n}} \Big| / q(t)$$

$$= o_P(1). \tag{3.4.2}$$

As $\ell(n)$ is a slowly varying function at infinity, for large n we have that $\ell(n)\sqrt{n} \geq \sqrt{n}$. Moreover, note that

$$\sum_{1\leq i_1<\ldots<i_m\leq [nt]} (\tilde{h}_1(X_{i_1}) + \ldots + \tilde{h}_1(X_{i_m})) = \frac{m}{[nt]} \binom{[nt]}{m} \sum_{i=1}^{[nt]} \tilde{h}_1(X_i).$$

Also observe that for $\frac{m}{n} < t \leq 1$, as $n \to \infty$, we have

$$\frac{\binom{[nt]}{m}}{[nt]\ n^{m-1}} \to t^{m-1}.$$

Now it is clear that in order to establish (3.4.2), it will be enough to prove the following Proposition 3.4.1.

Proposition 3.4.1. *Let $q \in Q$. If $E|h(X_1, \ldots, X_m)|^{\frac{4}{3}} < \infty$, then, as $n \to \infty$*

$$n^{-m+\frac{1}{2}} \sup_{\frac{m}{n}<t\leq 1} \Big| \sum_{1\leq i_1<\ldots<i_m\leq [nt]} (h(X_{i_1}, \ldots, X_{i_m}) - \theta - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m})) \Big| / q(t)$$

$= o_P(1)$.

Proof of Proposition 3.4.1

Again without loss of generality assume that $\theta = 0$ and let $\delta \in (0, 1]$, be so small such that $q(t)$ is already non-decreasing on $(0, \delta)$, then we have

$$n^{-m+\frac{1}{2}} \sup_{\frac{m}{n}<t\le 1} \Big| \sum_{1\le i_1<\ldots<i_m\le[nt]} (h(X_{i_1},\ldots,X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m}))\Big|/q(t)$$

$$\begin{aligned}
&\le n^{-m+\frac{1}{2}} \sup_{\frac{m}{n}<t<\delta} \Big| \sum_{1\le i_1<\ldots<i_m\le[nt]} (h(X_{i_1},\ldots,X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m}))\Big|/q(t) \\
&+n^{-m+\frac{1}{2}} \sup_{\delta\le t\le 1} \Big| \sum_{1\le i_1<\ldots<i_m\le[nt]} (h(X_{i_1},\ldots,X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m}))\Big|/q(t) \\
&:= T_1(n,t) + T_2(n,t) \qquad (3.4.3)
\end{aligned}$$

To prove Proposition 3.4.1 we need to show asymptotic negligibility of both $T_1(n, t)$ and $T_2(n, t)$ in probability. But first observe that for $\delta \in (0, 1]$ we have that

$$\begin{aligned}
T_1(n,t) &\le n^{-m+1} \sup_{\frac{m}{n}<t<\delta} (nt)^{\frac{-1}{2}} \Big| \sum_{1\le i_1<\ldots<i_m\le[nt]} (h(X_{i_1},\ldots,X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m}))\Big| \\
&\quad \times \sup_{\frac{m}{n}<t<\delta} \frac{t^{\frac{1}{2}}}{q(t)} \\
&\le n^{-m+1} \sup_{\frac{m}{n}\le t\le 1} [nt]^{\frac{-1}{2}} \Big| \sum_{1\le i_1<\ldots<i_m\le[nt]} (h(X_{i_1},\ldots,X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m}))\Big| \\
&\quad \times \sup_{\frac{m}{n}<t<\delta} \frac{t^{\frac{1}{2}}}{q(t)}.
\end{aligned}$$

From the latter relation and by virtue of Lemma 2.3.3, to prove $T_1(n,t) = o_P(1)$, it suffices to show that

$$n^{-m+1} \sup_{\frac{m}{n} \le t \le 1} [nt]^{\frac{-1}{2}} \Big| \sum_{1 \le i_1 < \ldots < i_m \le [nt]} (h(X_{i_1}, \ldots, X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m})) \Big|$$

$$= O_P(1). \tag{3.4.4}$$

Also for $T_2(n,t)$ we can write

$$\begin{aligned} T_2(n,t) &\le n^{-m+\frac{1}{2}} \sup_{\delta \le t \le 1} \Big| \sum_{1 \le i_1 < \ldots < i_m \le [nt]} (h(X_{i_1}, \ldots, X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m})) \Big| \\ &\qquad \times \sup_{\delta \le t \le 1} \frac{1}{q(t)} \\ &\le n^{-m+\frac{1}{2}} \sup_{\frac{m}{n} \le t \le 1} \Big| \sum_{1 \le i_1 < \ldots < i_m \le [nt]} (h(X_{i_1}, \ldots, X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m})) \Big| \\ &\qquad \times \sup_{\delta \le t \le 1} \frac{1}{q(t)}. \end{aligned}$$

But since $\inf_{\delta \le t \le 1} \frac{1}{q(t)} = O(1)$, to show that $T_2(n,t) = o_P(1)$, we only need to show that

$$n^{-m+\frac{1}{2}} \sup_{\frac{m}{n} \le t \le 1} \Big| \sum_{1 \le i_1 < \ldots < i_m \le [nt]} (h(X_{i_1}, \ldots, X_{i_m}) - \tilde{h}_1(X_{i_1}) - \ldots - \tilde{h}_1(X_{i_m})) \Big|$$

$$= o_P(1). \tag{3.4.5}$$

We first establish (3.4.4). To do so, we consider the truncation setup of (3.3.5) and we proceed by stating and proving the following proposition 3.4.2.

Proposition 3.4.2. *If* $E|h(X_1, \ldots, X_m)|^{\frac{4}{3}} < \infty$, *then, as* $n \to \infty$ *we have*

$$n^{-m+1}\max_{m\le K\le n}K^{-\frac{1}{2}}\sum_{1\le i_1<\ldots<i_m\le K}g^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})=O_P(1)\tag{3.4.6}$$

and

$$n^{-m+1}\max_{m\le K\le n}K^{-\frac{1}{2}}\sum_{1\le i_1<\ldots<i_m\le K}g^{tr}_{i_m}(X_{i_1},\ldots,X_{i_m})=O_P(1).\tag{3.4.7}$$

Proof of Proposition 3.4.2

To prove (3.4.6), we represent the statistic $\sum_{1\le i_1<\ldots<i_m\le K}g^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})$ in terms of 2^m-m-1 statistics which their kernels posses the property of complete degeneracy, as follows.

$$\begin{aligned}
&\sum_{1\le i_1<\ldots<i_m\le K}g^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})\\
&=\sum_{1\le i_1<\ldots<i_m\le K}\{\sum_{d=2}^{m}(-1)^{m-d}\sum_{1\le j_1<\ldots<j_d\le m}E(g^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})|X_{i_{j_1}},\ldots,X_{i_{j_d}})\\
&+\sum_{c=2}^{m-1}\sum_{1\le k_1<\ldots<k_c\le m}\sum_{d=2}^{c}(-1)^{c-d}\sum_{1\le j_1<\ldots<j_d\le c}E(g^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})|X_{i_{k_{j_1}}},\ldots,X_{i_{k_{j_d}}})\}\\
&:=\sum_{1\le i_1<\ldots<i_m\le K}V^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})\\
&+\sum_{c=2}^{m-1}\sum_{1\le k_1<\ldots<k_c\le m}\sum_{1\le i_1<\ldots<i_m\le K}V^{ta}_{i_m}(X_{i_{k_1}},\ldots,X_{i_{k_c}}).
\end{aligned}$$

In view of the latter setup to prove (3.4.6) we need to show that

$$n^{-m+1}\max_{m\le K\le n}K^{-\frac{1}{2}}\sum_{1\le i_1<\ldots<i_m\le K}V^{ta}_{i_m}(X_{i_1},\ldots,X_{i_m})=O_P(1)\tag{3.4.8}$$

and for, $m \geq 3$, $c = 2, \ldots, m-1$ and $1 \leq k_1 < \ldots < k_c \leq m$,

$$n^{-m+1} \max_{m \leq K \leq n} K^{-\frac{1}{2}} \sum_{1 \leq i_1 < \ldots < i_m \leq K} V_{i_m}^{ta}(X_{i_{k_1}}, \ldots, X_{i_{k_c}}) = O_P(1). \tag{3.4.9}$$

Due to similarity we shall only give the detailed proof of (3.4.8). To establish the latter write

$$\begin{aligned}
P & \ (n^{-m+1} \max_{m \leq K \leq n} K^{-\frac{1}{2}} \Big| \sum_{1 \leq i_1 < \ldots < i_m \leq K} V_{i_m}^{ta}(X_{i_1}, \ldots, X_{i_m}) \Big| > M) \\
= & \ P(n^{-m+1} \max_{m \leq K \leq n} \sum_{i_m = m}^{K} i_m^{-\frac{1}{2}} \Big| \sum_{1 \leq i_1 < \ldots < i_{m-1} < i_m} V_{i_m}^{ta}(X_{i_1}, \ldots, X_{i_m}) \Big| > M) \\
\leq & \ M^{-1} n^{-m+1} E\Big(\max_{m \leq K \leq n} \sum_{i_m = m}^{K} i_m^{-\frac{1}{2}} \Big| \sum_{1 \leq i_1 < \ldots < i_{m-1} < i_m} V_{i_m}^{ta}(X_{i_1}, \ldots, X_{i_m}) \Big| \Big) \\
\leq & \ M^{-1} n^{-m+1} E\Big(\sum_{i_m = m}^{\infty} i_m^{-\frac{1}{2}} \Big| \sum_{1 \leq i_1 < \ldots < i_{m-1} < i_m} V_{i_m}^{ta}(X_{i_1}, \ldots, X_{i_m}) \Big| \Big) \\
\leq & \ A M^{-1} n^{-m+1} \sum_{i_m = m}^{\infty} i_m^{-\frac{1}{2}} \binom{i_m}{m-1} E\Big(|h(X_1, \ldots, X_m)| \ 1_{(|h| > i_m^{\frac{3}{2}})} \Big) \\
\leq & \ A M^{-1} \sum_{i_m = m}^{\infty} i_m^{-\frac{1}{2}} E\Big(|h(X_1, \ldots, X_m)| \mathbf{1}_{(\cup_{j=i_m}^{\infty} (j^{\frac{3}{2}} < |h| \leq (j+1)^{\frac{3}{2}}))} \Big) \\
\leq & \ A M^{-1} \sum_{j=m}^{\infty} E|h(X_1, \ldots, X_m)| \mathbf{1}_{(j^{\frac{3}{2}} < |h| \leq (j+1)^{\frac{3}{2}})} \sum_{i_m = m}^{j} i_m^{-\frac{1}{2}} \\
\leq & \ A M^{-1} \sum_{j=m}^{\infty} E|h(X_1, \ldots, X_m)|^{\frac{4}{3}} \mathbf{1}_{(j^{\frac{3}{2}} < |h| \leq (j+1)^{\frac{3}{2}})} \\
= & \ A M^{-1} E\Big(|h(X_1, \ldots, X_m)|^{\frac{4}{3}} \mathbf{1}_{(\cup_{j=m}^{\infty} (j^{\frac{3}{2}} < |h| \leq (j+1)^{\frac{3}{2}}))} \Big) \\
\leq & \ A M^{-1} E|h(X_1, \ldots, X_m)|^{\frac{4}{3}} < \infty
\end{aligned}$$

The last relation completes the proof of (3.4.8) and that of (3.4.6).

To prove (3.4.7), similarly to what we had in the proof of (3.4.8), we write

$$\sum_{1\le i_1<\ldots<i_m\le K} g_{i_m}^{tr}(X_{i_1},\ldots,X_{i_m})$$

$$= \sum_{1\le i_1<\ldots<i_m\le K} \{ \sum_{d=2}^{m}(-1)^{m-d} \sum_{1\le j_1<\ldots<j_d\le m} E(g_{i_m}^{tr}(X_{i_1},\ldots,X_{i_m})|X_{i_{j_1}},\ldots,X_{i_{j_d}})$$

$$+\sum_{c=2}^{m-1} \sum_{1\le k_1<\ldots<k_c\le m} \sum_{d=2}^{c}(-1)^{c-d} \sum_{1\le j_1<\ldots<j_d\le c} E(g_{i_m}^{tr}(X_{i_1},\ldots,X_{i_m})|X_{i_{k_{j_1}}},\ldots,X_{i_{k_{j_d}}})\}$$

$$:= \sum_{1\le i_1<\ldots<i_m\le K} V_{i_m}^{tr}(X_{i_1},\ldots,X_{i_m})$$

$$+\sum_{c=2}^{m-1} \sum_{1\le k_1<\ldots<k_c\le m} \sum_{1\le i_1<\ldots<i_m\le K} V_{i_m}^{tr}(X_{i_{k_1}},\ldots,X_{i_{k_c}}).$$

Therefore, to establish (3.4.7) we need to show that

$$n^{-m+1} \max_{m\le K\le n} \sum_{1\le i_1<\ldots<i_m\le K} V_{i_m}^{tr}(X_{i_1},\ldots,X_{i_m}) = O_P(1) \tag{3.4.10}$$

and for, $m\ge 3$, $c=2,\ldots,m-1$ and $1\le k_1<\ldots<k_c\le m$,

$$n^{-m+1} \max_{m\le K\le n} \sum_{1\le i_1<\ldots<i_m\le K} V_{i_m}^{tr}(X_{i_{k_1}},\ldots,X_{i_{k_c}}) = O_P(1). \tag{3.4.11}$$

To prove (3.4.10), in view of Proposition 1.8.3, we employ Lemma 1.11.1 and Proposition 1.8.2 and we arrive at

$$
\begin{aligned}
P & \ (n^{-m+1} \max_{m \le K \le n} K^{-\frac{1}{2}} \big| \sum_{1 \le i_1 < \ldots < i_m \le K} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_m}) \big| > M) \\
\le & \ M^{-2} n^{-2m+2} \{ n^{-1} E(\sum_{1 \le i_1 < \ldots < i_m \le n} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_m}))^2 \\
+ & \ \sum_{K=m}^{n-1} (\frac{1}{K} - \frac{1}{K+1}) E(\sum_{1 \le i_1 < \ldots < i_m \le K} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_m}))^2 \} \\
\le & \ M^{-2} n^{-2m+1} \sum_{i_m=m}^{n} E\big(\sum_{1 \le i_1 < \ldots < i_{m-1} < i_m} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_m})\big)^2 \\
+ & \ M^{-2} n^{-2m+2} \sum_{K=m}^{n-1} K^{-2} \sum_{i_m=m}^{K} E\big(\sum_{1 \le i_1 < \ldots < i_{m-1} < i_m} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_m})\big)^2 \\
\le & \ A M^{-2} n^{-2m+1} \sum_{i_m=m}^{n} \binom{i_m}{m-1} E\big(h^2(X_1, \ldots, X_m) \, \mathbf{1}_{(|h| \le i_m^{\frac{3}{2}})}\big) \\
+ & \ A M^{-2} n^{-2m+2} \sum_{K=m}^{n-1} K^{-2} \sum_{i_m=m}^{K} \binom{i_m}{m-1} E\big(h^2(X_1, \ldots, X_m) \, \mathbf{1}_{(|h| \le i_m^{\frac{3}{2}})}\big) \\
\le & \ A M^{-2} n^{-m} E\big(h^2(X_1, \ldots, X_m) \, \mathbf{1}_{(|h| \le n^{\frac{3}{2}})}\big) \sum_{i_m=m}^{n} 1 \\
+ & \ A M^{-2} n^{-m+1} \sum_{K=m}^{n-1} K^{-2} E\big(h^2(X_1, \ldots, X_m) \, \mathbf{1}_{(|h| \le K^{\frac{3}{2}})}\big) \sum_{i_m=m}^{K} 1 \\
\le & \ A M^{-2} n^{-1} E\big(h^2(X_1, \ldots, X_m) \, \mathbf{1}_{(|h| \le n^{\frac{3}{2}})}\big) \\
+ & \ A M^{-2} \sum_{K=m}^{n-1} K^{-2} E\big(h^2(X_1, \ldots, X_m) \, \mathbf{1}_{(|h| \le K^{\frac{3}{2}})}\big) \\
\le & \ M^{-2} E|h(X_1, \ldots, X_m)|^{\frac{4}{3}} < \infty.
\end{aligned}
$$

Due to similarity, to establish (3.4.11) we shall only state the proof for $k_1 = 1, \ldots, k_c = c$, where $c = 2, \ldots, m-1$ and $m \ge 3$. Once again an application of Lemma 1.11.1

and Proposition 1.8.2 for $M > 0$ yield

$$\begin{aligned}
& P\ \ (n^{-m+1} \max_{m \le K \le n} K^{-\frac{1}{2}} \Big| \sum_{1 \le i_1 < \ldots < i_m \le K} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_c}) \Big| > M) \\
\le\ & P\Big(n^{-m+1} \max_{m \le K \le n} K^{-\frac{1}{2}} \binom{K}{m-c-1} \Big| \sum_{i_m=m}^{K} \sum_{i_1 < \ldots < i_c} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_c}) \Big| > M\Big) \\
\le\ & P\Big(n^{-c} \max_{m \le K \le n} \frac{K^{-\frac{1}{2}} \binom{K}{m-c-1}}{\binom{K}{m-c-1}} \Big| \sum_{i_m=m}^{K} \sum_{i_1 < \ldots < i_c} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_c}) \Big| > M\Big) \\
=\ & P\Big(n^{-c} \max_{m \le K \le n} K^{-\frac{1}{2}} \Big| \sum_{i_m=m}^{K} \sum_{i_1 < \ldots < i_c} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_c}) \Big| > M\Big) \\
\le\ & M^{-2} n^{-2c} n^{-1} n \sum_{i_m=m}^{n} E\Big(\sum_{i_1 < \ldots < i_c} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_c})\Big)^2 \\
+\ & M^{-2} n^{-2c} \sum_{K=m}^{n} \Big(\frac{1}{K} - \frac{1}{K+1}\Big) K \sum_{i_m=m}^{K} E\Big(\sum_{i_1 < \ldots < i_c} V_{i_m}^{tr}(X_{i_1}, \ldots, X_{i_c})\Big)^2 \\
\le\ & A M^{-2} n^{-2c} \sum_{i_m=m}^{n} \binom{i_m}{c} E\big(h^2(X_1, \ldots, X_m)\ 1_{(|h| \le i_m^{\frac{3}{2}})}\big) \\
+\ & A M^{-2} n^{-2c} \sum_{K=m}^{n} K^{-1} \sum_{i_m=m}^{K} \binom{i_m}{c} E\big(h^2(X_1, \ldots, X_m)\ 1_{(|h| \le i_m^{\frac{3}{2}})}\big) \\
\le\ & A M^{-2} n^{-c} E\big(h^2(X_1, \ldots, X_m)\ 1_{(|h| \le n^{\frac{3}{2}})}\big) \sum_{i_m=m}^{n} 1 \\
+\ & A M^{-2} n^{-c} E\big(h^2(X_1, \ldots, X_m)\ 1_{(|h| \le n^{\frac{3}{2}})}\big) \sum_{K=m}^{n} K^{-1} \sum_{i_m=m}^{K} 1 \\
\le\ & A M^{-2} n^{-c+1} E\big(h^2(X_1, \ldots, X_m)\ 1_{(|h| \le n^{\frac{3}{2}})}\big) \\
\le\ & A M^{-2} E|h(X_1, \ldots, X_m)|^{\frac{4}{3}} < \infty.
\end{aligned}$$

The last relation completes the proof of (3.4.11) and that of Proposition 3.4.2.

By virtue of our notion of $T_1(n,t)$ and $T_2(n,t)$, so far, we have showed that $T_1(n,t) = o_P(1)$. To complete the proof of Theorem 3.4.4, we need to show that $T_2(n,t) = o_P(1)$. But to show the latter, we only need to establish (3.4.5). To do so, first consider the truncation setup of (3.3.13). Now we need to prove the following Proposition 3.4.3.

Proposition 3.4.3. *If $E|h(X_1,\ldots,X_m|^{\frac{4}{3}}<\infty$, then, as $n\to\infty$ we have*

$$n^{-m+\frac{1}{2}}\max_{m\le K\le n}\sum_{1\le i_1<\ldots<i_m\le K} g_n^{ta}(X_{i_1},\ldots,X_{i_m})=o_P(1) \tag{3.4.12}$$

and

$$n^{-m+\frac{1}{2}}\max_{m\le K\le n}\sum_{1\le i_1<\ldots<i_m\le K} g_n^{tr}(X_{i_1},\ldots,X_{i_m})=o_P(1). \tag{3.4.13}$$

Proof of Proposition 3.4.3

To prove (3.4.12) we observe that

$$\sum_{1\le i_1<\ldots<i_m\le K} g_n^{ta}(X_{i_1},\ldots,X_{i_m})$$

$$=\sum_{1\le i_1<\ldots<i_m\le K}\{\sum_{d=2}^{m}(-1)^{m-d}\sum_{1\le j_1<\ldots<j_d\le m}E(g_n^{ta}(X_{i_1},\ldots,X_{i_m})|X_{i_{j_1}},\ldots,X_{i_{j_d}})$$

$$+\sum_{c=2}^{m-1}\sum_{1\le k_1<\ldots<k_c\le m}\sum_{d=2}^{c}(-1)^{c-d}\sum_{1\le j_1<\ldots<j_d\le c}E(g_n^{ta}(X_{i_1},\ldots,X_{i_m})|X_{i_{k_{j_1}}},\ldots,X_{i_{k_{j_d}}})\}$$

$$:=\sum_{1\le i_1<\ldots<i_m\le K}V_n^{ta}(X_{i_1},\ldots,X_{i_m})$$

$$+\sum_{c=2}^{m-1}\sum_{1\le k_1<\ldots<k_c\le m}\sum_{1\le i_1<\ldots<i_m\le K}V_n^{ta}(X_{i_{k_1}},\ldots,X_{i_{k_c}}).$$

Therefore, (3.4.12) will follow if we show

$$n^{-m+\frac{1}{2}}\max_{m\le K\le n}\sum_{1\le i_1<\ldots<i_m\le K} V_n^{ta}(X_{i_1},\ldots,X_{i_m})=o_P(1) \tag{3.4.14}$$

and for, $m \geq 3$, $c = 2, \ldots, m-1$ and $1 \leq k_1 < \ldots < k_c \leq m$,

$$n^{-m+\frac{1}{2}} \max_{m \leq K \leq n} \sum_{1 \leq i_1 < \ldots < i_m \leq K} V_n^{ta}(X_{i_{k_1}}, \ldots, X_{i_{k_c}}) = o_P(1). \tag{3.4.15}$$

Again due to similarity we shall only give the proof of (3.4.14). Due to martingale property of $\sum_{1 \leq i_1 < \ldots < i_m \leq K} V_n^{ta}(X_{i_1}, \ldots, X_{i_m})$, cf. Proposition 1.8.3, an application of Lemma 1.11.2, for $\epsilon > 0$ yields,

$$\begin{aligned}
P & \left(n^{-m+\frac{1}{2}} \max_{m \leq K \leq n} \Big| \sum_{1 \leq i_1 < \ldots < i_m \leq K} V_n^{ta}(X_{i_1}, \ldots, X_{i_m}) \Big| > \epsilon\right) \\
\leq & \ \epsilon^{-1} n^{-m+\frac{1}{2}} E\Big(\Big| \sum_{1 \leq i_1 < \ldots < i_m \leq n} V_n^{ta}(X_{i_1}, \ldots, X_{i_m}) \Big|\Big) \\
\leq & \ A \epsilon^{-1} n^{-m+\frac{1}{2}} \binom{n}{m} E(|h(X_1, \ldots, X_m)| \mathbf{1}_{|h| > n^{\frac{3}{2}}}) \\
\leq & \ A \epsilon^{-1} \frac{\binom{n}{m}}{n^m} E(|h(X_1, \ldots, X_m)|^{\frac{4}{3}} \mathbf{1}_{|h| > n^{\frac{3}{2}}}) \\
\leq & \ A \epsilon^{-1} E(|h(X_1, \ldots, X_m)|^{\frac{4}{3}} \mathbf{1}_{|h| > n^{\frac{3}{2}}}) \to 0, \text{ as } \to \infty.
\end{aligned}$$

Now the proof of (3.4.14) is complete.

To establish (3.4.13) we write

$$\begin{aligned}
& \sum_{1 \leq i_1 < \ldots < i_m \leq K} g_n^{tr}(X_{i_1}, \ldots, X_{i_m}) \\
= & \sum_{1 \leq i_1 < \ldots < i_m \leq k} \Big\{ \sum_{d=2}^{m} (-1)^{m-d} \sum_{1 \leq j_1 < \ldots < j_d \leq m} E(g_n^{tr}(X_{i_1}, \ldots, X_{i_m}) | X_{i_{j_1}}, \ldots, X_{i_{j_d}}) \\
& + \sum_{c=2}^{m-1} \sum_{1 \leq k_1 < \ldots < k_c \leq m} \sum_{d=2}^{c} (-1)^{c-d} \sum_{1 \leq j_1 < \ldots < j_d \leq c} E(g_n^{tr}(X_{i_1}, \ldots, X_{i_m}) | X_{i_{k_{j_1}}}, \ldots, X_{i_{k_{j_d}}}) \Big\} \\
:= & \sum_{1 \leq i_1 < \ldots < i_m \leq K} V_n^{tr}(X_{i_1}, \ldots, X_{i_m})
\end{aligned}$$

$$+\sum_{c=2}^{m-1} \sum_{1\le k_1<\ldots<k_c\le m} \sum_{1\le i_1<\ldots<i_m\le K} V_n^{tr}(X_{i_{k_1}},\ldots,X_{i_{k_c}}).$$

Thus, the proof of (3.4.13) will follow if we show that the following (3.4.16) and (3.4.17) are true.

$$n^{-m+\frac{1}{2}} \max_{m\le K\le n} \sum_{1\le i_1<\ldots<i_m\le K} V_n^{tr}(X_{i_1},\ldots,X_{i_m}) = o_P(1) \tag{3.4.16}$$

and for, $m\ge 3$, $c=2,\ldots,m-1$ and $1\le k_1<\ldots<k_c\le m$,

$$n^{-m+\frac{1}{2}} \max_{m\le K\le n} \sum_{1\le i_1<\ldots<i_m\le K} V_n^{tr}(X_{i_{k_1}},\ldots,X_{i_{k_c}}) = o_P(1). \tag{3.4.17}$$

To establish (3.4.16) again using the maximal inequality of Lemma 1.11.2 followed by an application of Proposition 1.8.1 for $\epsilon>0$ we arrive at

$$\begin{aligned}
P & \left(\left|n^{-m+\frac{1}{2}} \max_{m\le K\le n} \sum_{1\le i_1<\ldots<i_m\le K} V_n^{tr}(X_{i_1},\ldots,X_{i_m})\right| > \epsilon\right) \\
\le & \ \epsilon^{-2} n^{-2m+1} E\Big(\sum_{1\le i_1<\ldots<i_m\le n} V_n^{tr}(X_{i_1},\ldots,X_{i_m})\Big)^2 \\
\le & \ A\epsilon^{-2} n^{-2m+1} \binom{n}{m} E(h^2(X_1,\ldots,X_m)1_{(|h|\le n^{\frac{3}{2}})}) \\
\le & \ A\epsilon^{-2} n^{-\frac{1}{3}} E(|h(X_1,\ldots,X_m)|^{\frac{4}{3}}) + E(|h(X_1,\ldots,X_m)|^{\frac{4}{3}}\mathbf{1}_{(|h|>n)}) \\
& \longrightarrow 0, \text{ as } n\to\infty.
\end{aligned}$$

The proof of (3.4.16) is now complete. Due to similarity of the two, we only give the proof of (3.4.17) for $k_1=1,\ldots,k_c=c$. Once again we make use of martingale property and apply Proposition 1.8.1 to our statistic and for $\epsilon > 0$ we write

$$\begin{aligned}
P\ & \Big(\big|n^{-m+\frac{1}{2}} \max_{m\le K\le n} \sum_{1\le i_1<\ldots<i_m\le K} V_n^{tr}(X_{i_1},\ldots,X_{i_c})\big| > \epsilon\Big) \\
\le\ & \epsilon^{-2} n^{-2m+1} E\Big(\sum_{1\le i_1<\ldots<i_m\le n} V_n^{tr}(X_{i_1},\ldots,X_{i_c})\big|\Big)^2 \\
\le\ & \epsilon^{-2} A n^{-2m+1} \binom{n}{m-c}^2 \binom{n}{c} E(h^2(X_1,\ldots,X_m)\mathbf{1}_{(|h|\le \frac{3}{2})}) \\
\le\ & \epsilon^{-2} A n^{-c+1} E(h^2(X_1,\ldots,X_m)\mathbf{1}_{(|h|\le n^{\frac{3}{2}})}) \\
\le\ & \epsilon^{-2} A n^{-c+2} E(|h(X_1,\ldots,X_m)|^{\frac{4}{3}}\mathbf{1}_{(|h|>n)}) + n^{-\frac{1}{3}} E(|h(X_1,\ldots,X_m)|^{\frac{4}{3}})
\end{aligned}$$

$$\longrightarrow 0, \text{ as, } n\rightarrow\infty,$$

And this completes the proof of (3.4.17) and those of (3.4.6) and Proposition 3.4.3. Now the Theorem 3.4.4 is complete. □

3.5 Examples

Due to the nonexistence of the second moment of the kernel of the underlying U-statistic in the following example, the weak convergence result of Theorem 1.9.2 fails to apply. However, using Theorems 3.3.1 and 3.4.1, one can still derive weak convergence results for the underlying U-statistic.

Examples. Let $X_1, X_2, \ldots,$ be a sequence of i.i.d. random variables with the density function

$$f(x) = \begin{cases} |x-a|^{-3}, & |x-a|\ge 1,\ a\ne 0, \\ 0, & \text{elsewhere.} \end{cases}$$

Consider the parameter $\theta = E^m(X_1) = a^m$, where $m \ge 1$ is a positive integer, and the kernel $h(X_1,\ldots,X_m) = \prod_{i=1}^m X_i$. Then with m, n satisfying $n \ge m$, the corresponding U-statistic is

$$U_n = \binom{n}{m}^{-1} \sum_{1 \le i_1 < \ldots < i_m \le n} \prod_{j=1}^{m} X_{i_j}.$$

Simple calculation shows that $\tilde{h}_1(X_1) = X_1 \; a^{m-1} \; - \; a^m$.

It is easy to check that $E\big(|h(X_1, \ldots, X_m)|^{\frac{4}{3}} \log |h(X_1, \ldots, X_m)|\big) \; < \; \infty$ and that $\tilde{h}_1(X_1) \in DAN$. In order to apply Theorems 3.3.1 and 3.4.1, define

$$U^{nor}_{[nt]} = \begin{cases} 0, & 0 \le t < \dfrac{m}{n}, \\ \dfrac{\binom{[nt]}{m}^{-1} \sum_{1 \le i_1 < \ldots < i_m \le [nt]} (\prod_{j=1}^{m} X_{i_j} \; - \; a^m)}{m(\sum_{i=1}^{n} (X_i \; a^{m-1} \; - \; a^m)^2)^{\frac{1}{2}}}, & \dfrac{m}{n} \le t \le 1 \end{cases}$$

and

$$\Psi^{nor}_{[nt]} = \begin{cases} 0, & 0 \le t < \dfrac{m}{n}, \\ \dfrac{n^{-m+1} \sum_{1 \le i_1 < \ldots < i_m \le [nt]} (\prod_{j=1}^{m} X_{i_j} \; - \; a^m)}{m(\sum_{i=1}^{n} (X_i \; a^{m-1} \; - \; a^m)^2)^{\frac{1}{2}}}, & \dfrac{m}{n} \le t \le 1. \end{cases}$$

Then, based on (a) of Theorems 3.3.1 and 3.4.1, when $q = 1$, as $n \to \infty$, we have

$$U^{nor}_{[nt]} \Rightarrow W(t) \;\; \text{on} \; (D[0,1], \rho)$$

and

$$\Psi^{nor}_{[nt]} \Rightarrow t^{m-1} \; W(t) \;\; \text{on} \; (D[0,1], \rho),$$

where ρ is the sup-norm metric for functions in $D[0,1]$ and $\{W(t), \; 0 \le t \le 1\}$ is a standard Wiener process.

Chapter 4

Studentized U-statistics and Studentized U-type statistics

4.1 Introduction

In Theorems 3.3.1 and 3.4.1 of the preceding chapter, we introduced the pseudo self-normalizing sequence $V_n = \sqrt{\sum_{i=1}^n \tilde{h}_1^2(X_i)}$ and derived weighted weak approximations results for both U-statistics and U-type statistics. We also showed, in case of possibly infinite variance of the projection $\tilde{h}_1$, V_n remains the right substitute for $n^{1/2}\sqrt{Var\tilde{h}_1(X_1)}$ in the sense that incorporating V_n leads to a Wiener limit in distribution. Along these lines, we used the asymptotic equivalency of V_n and $B_n = n^{1/2}\ell(n)$ in view of Lemma 2.2.2, to establish the proofs (cf. Theorems 3.3.2 and 3.4.2). However, when it comes to requiring the distribution of the original observations $X_1, X_2, \ldots$, V_n offers no more help than B_n. In other words, computing both of them requires that we know the distribution function F of X_i's. But we should highlight the importance of V_n in the sense that it constitutes a significant first step toward Studentizing U-statistics and U-type statistics. In fact, as we shall see, estimating V_n in order to have a completely data based computable statistic, leads to having studentized U-statistics and U-type statistics which are the subject of this chapter. An immediate application of this Studentization is establishing asymptotic confidence intervals for the parameter of interest $\theta = \theta(F) = E_F(h(X_1, \ldots, X_m))$, $m \leq n$, in a nonparametric manner.

4.2 Statement of the results

For $i = 1, \ldots, n$, let U_{n-1}^i be the *jackknifed* version of U_n based on $X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n$, defined as follows.

$$U_{n-1}^i = \frac{1}{\binom{n-1}{m}} \sum_{\substack{1 \leq j_1 < \ldots < j_m \leq n \\ j_1, \ldots, j_m \neq i}} h(X_{j_1}, \ldots, X_{j_m}).$$

Also define the *Studentized* U-statistic process, $U^{stu}_{[nt]}$ and the *Studentized* U-type process, $\Psi^{stu}_{[nt]}$, respectively as follows.

$$U^{stu}_{[nt]} = \begin{cases} 0, & 0 \le t < \dfrac{m}{n}, \\ \dfrac{[nt]\left(U_{[nt]} - \theta\right)}{\sqrt{(n-1)\sum_{i=1}^{n}(U^{i}_{n-1} - U_n)^2}}, & \dfrac{m}{n} \le t \le 1, \end{cases}$$

$$\Psi^{stu}_{[nt]} = \begin{cases} 0, & 0 \le t < \dfrac{m}{n}, \\ \dfrac{n^{-m+1}\sum_{1\le i_1<\ldots<i_m\le[nt]}(h(X_{i_1},\ldots,X_{i_m}) - \theta)}{\sqrt{(n-1)\sum_{i=1}^{n}(U^{i}_{n-1} - U_n)^2}}, & \dfrac{m}{n} \le t \le 1. \end{cases}$$

Remark 4.2.1. *Unlike the U-processes and U-type processes in Theorems 3.3.1 and 3.4.1, apart from the unknown parameter of interest θ, $U^{\rm stu}_{[nt]}$ and $\Psi^{\rm stu}_{[nt]}$ are completely computable, based on the observations $X_1, \ldots, X_n$.*

Under a slightly stronger moment condition, which is the price we pay for the normalization involved in $U^{stu}_{[nt]}$, the Studentized companions of Theorems 3.3.1 and 3.3.2 read as follows.

Theorem 4.2.1. *Let $q \in Q$. If*

(a) $E|h(X_1,\ldots,X_m)|^{\frac{5}{3}} < \infty$ and $\tilde{h}_1(X_1) \in DAN$,

then, as $n \to \infty$, we have

(b) $U^{\rm stu}_{[nt]} \Rightarrow W(t)$ on $(D, \mathfrak{D}, \|/q\|)$ if and only if $I(q,c) < \infty$ for all $c > 0$, where $\{W(t), 0 \le t \le 1\}$ is a standard Wiener process and $\|/q\|$ is the weighted sup-norm metric for functions in $D[0,1]$;

(c) On an appropriate probability space for $X_1, X_2, \ldots$, we can construct a standard

Wiener process $\{W(t), 0 \le t < \infty\}$ *such that*

$$\sup_{0\le t\le 1}\left| U^{\mathrm{stu}}_{[nt]} - \frac{W(nt)}{n^{\frac{1}{2}}}\right| / q(t) = o_P(1),$$

if and only if $I(q,c) < \infty$.

Similarly, the Studentized U-type companions of Theorems 3.4.1 and 3.4.2, read as follows.

Theorem 4.2.2. *Let* $q \in Q$. *If*

(a) $E|h(X_1,\ldots,X_m)|^{\frac{5}{3}} < \infty$ *and* $\tilde{h}_1(X_1) \in DAN$,

then, as $n \to \infty$, *we have*

(b) $\Psi^{\mathrm{stu}}_{[nt]} \Rightarrow t^{m-1}\, W(t)$ *on* $(D, \mathfrak{D}, \|/q\|)$ *if and only if* $I(q,c) < \infty$ *for all* $c > 0$, *where* $\{W(t), 0 \le t \le 1\}$ *is a standard Wiener process and* $\|/q\|$ *is the weighted sup-norm metric for functions in* $D[0,1]$;

(c) On an appropriate probability space for $X_1, X_2, \ldots$, *we can construct a standard Wiener process* $\{W(t), 0 \le t < \infty\}$ *such that*

$$\sup_{0\le t\le 1}\left| \Psi^{\mathrm{stu}}_{[nt]} - \frac{\binom{[nt]}{m}}{[nt]\; n^{m-1}}\frac{W(nt)}{n^{\frac{1}{2}}}\right| / q(t) = o_P(1),$$

if and only if $I(q,c) < \infty$.

Also the following Theorems 4.2.3 and 4.2.4 are respectively Studentized versions of Theorems 3.3.3 and 3.4.3.

Theorem 4.2.3. *Let* $\{W(t), 0 \le t < \infty\}$ *be a standard Wiener process,* $E\big|h(X_1,\ldots,X_m)\big|^{\frac{5}{3}} < \infty$ *and* $\tilde{h}_1(X_1) \in DAN$. *If* $q \in Q$ *and* $q(t)$ *is nondecreasing on*

$(0,1]$, *then as* $n \to \infty$,

$$\sup_{0<t\leq 1} |U^{\mathrm{stu}}_{[nt]}|/q(t) \longrightarrow_d \sup_{0<t\leq 1} |W(t)|/q(t)$$

if and only if $I(q,c) < \infty$ *for some* $c > 0$. *Consequently as* $n \to \infty$ *we have*

$$\sup_{0<t\leq 1} |U^{\mathrm{stu}}_{[nt]}|/(t \log\log(\frac{1}{t}))^{\frac{1}{2}} \longrightarrow_d \sup_{0<t\leq 1} |W(t)|/(t\log\log(\frac{1}{t}))^{\frac{1}{2}}.$$

Theorem 4.2.4. *Let* $\{W(t), 0 \leq t < \infty\}$ *be a standard Wiener process,*

$$E\big|h(X_1,\ldots,X_m)\big|^{\frac{5}{3}} < \infty$$

and $\tilde{h}_1(X_1) \in DAN$. *If* $q \in Q$ *and* $q(t)$ *is nondecreasing on* $(0,1]$, *then as* $n \to \infty$,

$$\sup_{0<t\leq 1} |\Psi^{\mathrm{stu}}_{[nt]}|/q(t) \longrightarrow_d \sup_{0<t\leq 1} t^{m-1}|W(t)|/q(t),$$

if and only if $I(q,c) < \infty$ *for some* $c > 0$. *Consequently, as* $n \to \infty$, *we have*

$$\sup_{0<t\leq 1} |\Psi^{\mathrm{stu}}_{[nt]}|/(t\log\log(\frac{1}{t}))^{\frac{1}{2}} \longrightarrow_d \sup_{0<t\leq 1} t^{m-1}|W(t)|/(t\log\log(\frac{1}{t}))^{\frac{1}{2}}.$$

In view of Theorems 3.3.1 and 3.4.1 and on account of the conclusion of Raikov's theorem in this context, i.e. (3.2.2), in order to prove Theorems 4.2.1, 4.2.2, 4.2.3 and 4.2.4, it suffices to prove the following result.

Theorem 4.2.5. *If* $E|h(X_1,\ldots,X_m)|^{\frac{5}{3}} < \infty$ *and* $\tilde{h}_1(X_1) \in DAN$, *then, as* $n \to \infty$,

$$\left| \frac{(n-1)}{m^2\ \ell^2(n)} \sum_{i=1}^{n} (U^i_{n-1} - U_n)^2 - \frac{1}{n\ \ell^2(n)} \sum_{i=1}^{n} \tilde{h}_1^2(X_i) \right| = o_P(1).$$

Consequently, the latter approximation combined with (3.2.2), the conclusion of Raikov's theorem, yields a Raikov type result for the distribution free jackkifed version of U-statistics which is of interest on its own (cf. Remark 4.2.1 of this chapter), and it reads as follows.

Corollary 4.2.1. *If* $E|h(X_1,\ldots,X_m)|^{\frac{5}{3}} < \infty$ *and* $\tilde{h}_1(X_1) \in DAN$, *then, as* $n \to \infty$,

$$\frac{(n-1)}{m^2\ \ell^2(n)} \sum_{i=1}^{n} (U_{n-1}^i - U_n)^2 \longrightarrow_P 1.$$

Combining now Corollary 4.2.1 with Theorems 3.3.1, 3.4.1, 3.3.3 and 3.4.3 we arrive respectively at Theorems 4.2.1, 4.2.2, 4.2.3 and 4.2.4 of this paper.

Remark 4.2.1. *When* $Eh^2(X_1,\ldots,X_m) < \infty$, *which in turn implies that* $E\tilde{h}_1^2(X_1) < \infty$, *then* $\ell^2(n) = E\tilde{h}_1^2(X_1) > 0$ *and, as* $n \to \infty$, *Corollary 4.2.1 implies that*

$$\frac{(n-1)}{m^2} \sum_{i=1}^{n} (U_{n-1}^i - U_n)^2 \longrightarrow_P E\tilde{h}_1^2(X_1).$$

The latter version of Corollary 4.2.1 coincides with one of the results obtained by Arvesen [2] who extended the idea of the so-called (by Tukey) pseudo- values to U-statistics and studied the asymptotic distribution of nondegenerate U-statistics via jackknifing.

Remark 4.2.2. *When* $m = 1$, *the projection* $\tilde{h}_1(X_1)$ *will coincide with* $h(X_1)-\theta$, *and Theorems 4.2.1 and 4.2.2, correspond to Corollary 5 of CsSzW (2008 [16]) on taking the weight function* $q = 1$ *for the therein studied Studentized process* $T_{n,t}(X-\mu)$, *i.e., when* $m = 1$, *the studentized U-process* $U_{[nt]}^{\text{stu}}$ *coincides with* $T_{n,t}(X-\mu)$. *Hence in this exposition we shall state our proofs for* $m \geq 2$. *Also, when* $m = 2$, *the two conditions in (a) of Theorem 4.2.2 as well as the idea of its proof by truncation, coincide with the* *corresponding ones of Theorem 2 of CsSzW (2008 [17]) on weighted approximations for Studentized U-type processes.*

4.3 Proof of Theorem 4.2.5

To prove Theorem 4.2.5, it suffices to show that as $n \to \infty$,

$$\left|(n-1)\ \sum_{i=1}^{n}(U_{n-1}^{i}-U_n)^2-\frac{m^2}{n}\sum_{i=1}^{n}\tilde{h}_1^2(X_i)\right|=o_P(1). \tag{4.3.1}$$

Before giving the proof of (4.3.1) we do some simplifications as follows.

$(n-1)\ \sum_{i=1}^{n}(U_{n-1}^{i}-U_n)^2$

$$=(n-1)\sum_{i=1}^{n}\left[\frac{\binom{n}{m}}{\binom{n-1}{m}}U_n-\frac{1}{\binom{n-1}{m}}\sum_{\substack{1\le j_1<\ldots<j_{m-1}\le n\\ j_1,\ldots,j_{m-1}\neq i}}h(X_i,X_{j_1}\ldots,X_{j_{m-1}})-U_n\right]^2$$

$$=(n-1)\sum_{i_1=1}^{n}\left[\frac{m}{n-m}\Big(\frac{1}{\binom{n-1}{m-1}}\sum_{\substack{1\le i_2<\ldots<i_m\le n\\ i_2,\ldots,i_m\neq i_1}}h(X_{i_1},X_{i_2},\ldots,X_{i_m})-U_n\Big)\right]^2$$

$$=\frac{m^2(n-1)}{(n-m)^2}\sum_{i_1=1}^{n}\left[\frac{1}{\binom{n-1}{m-1}}\sum_{\substack{1\le i_2<\ldots<i_m\le n\\ i_2,\ldots,i_m\neq i_1}}h(X_{i_1},X_{i_2},\ldots,X_{i_m})-U_n\right]^2 \tag{4.3.2}$$

$$=\frac{m^2(n-1)}{(n-m)^2}\sum_{i_1=1}^{n}\left[\frac{1}{\binom{n-1}{m-1}}\sum_{\substack{1\le i_2<\ldots<i_m\le n\\ i_2,\ldots,i_m\neq i_1}}h(X_{i_1},X_{i_2},\ldots,X_{i_m})\right]^2+\frac{m^2n(n-1)}{(n-m)^2}\,U_n^2$$

$$- 2\, \frac{m^2(n-1)}{(n-m)^2}\, U_n\, \frac{1}{\binom{n-1}{m-1}} \sum_{i_1=1}^{n} \sum_{\substack{1\le i_2<\ldots<i_m\le n\\ i_2,\ldots,i_m\ne i_1}} h(X_{i_1}, X_{i_2}, \ldots, X_{i_m})$$

$$= \frac{m^2(n-1)}{(n-m)^2} \sum_{i_1=1}^{n} \left[\frac{1}{\binom{n-1}{m-1}} \sum_{\substack{1\le i_2<\ldots<i_m\le n\\ i_2,\ldots,i_m\ne i_1}} h(X_{i_1}, X_{i_2}, \ldots, X_{i_m}) \right]^2 - \frac{m^2 n(n-1)}{(n-m)^2}\, U_n^2 \tag{4.3.3}$$

Remark 4.3.1. *In view of (4.3.2) in what will follow without loss of generality we may and shall assume that $\theta = 0$.*

By virtue of (4.3.3) to prove (4.3.1) it will be enough to prove the following two propositions.

Proposition 4.3.1. *If $E|h(X_1, \ldots, X_m)|^{\frac{5}{3}} < \infty$, then, as $n \to \infty$,*

$$U_n^2 \longrightarrow 0 \ \ a.s.$$

Proof of Proposition 4.3.1

The proof this theorem follows from the SLLN for U-statistics, i.e. Lemma 1.10.1 .

Proposition 4.3.2. *If $E|h(X_1, \ldots, X_m)|^{\frac{5}{3}} < \infty$ and $\tilde{h}_1(X_1) \in DAN$, then, as $n \to \infty$,*

$$\left| \frac{(n-1)}{(n-m)^2} \sum_{i_1=1}^{n} \left[\frac{1}{\binom{n-1}{m-1}} \sum_{\substack{1\le i_2<\ldots<i_m\le n\\ i_2,\ldots,i_m\ne i_1}} h(X_{i_1}, X_{i_2}, \ldots, X_{i_m}) \right]^2 - \frac{1}{n} \sum_{i=1}^{n} \tilde{h}_1^2(X_i) \right|$$

$$= o_P(1).$$

Proof of Proposition 4.3.2

As before, let $a_n \sim b_n$ stand for the asymptotic equivalency of the numerical sequences $(a_n)_n$ and $(b_n)_n$, i.e., as $n \to \infty$, $\dfrac{a_n}{b_n} \to 1$.

To prove Proposition 4.3.2 observe that

$$
\begin{aligned}
&\frac{(n-1)}{(n-m)^2} \sum_{i_1=1}^{n} \left[\frac{1}{\binom{n-1}{m-1}} \sum_{\substack{1 \le i_2 < \ldots < i_m \le n \\ i_2, \ldots, i_m \neq i_1}} h(X_{i_1}, X_{i_2}, \ldots, X_{i_m}) \right]^2 \\
&= \frac{(n-1)}{(n-m)^2} \sum_{i_1=1}^{n} \left[[n-1]^{-m+1} \sum_{\substack{1 \le i_2 \neq \ldots \neq i_m \le n \\ i_2, \ldots, i_m \neq i_1}} h(X_{i_1}, X_{i_2}, \ldots, X_{i_m}) \right]^2 \\
&\sim [n]^{-2m+1} \sum_{i_1=1}^{n} \left[\sum_{\substack{1 \le i_2 \neq \ldots \neq i_m \le n \\ i_2, \ldots, i_m \neq i_1}} h(X_{i_1}, X_{i_2}, \ldots, X_{i_m}) \right]^2 \\
&= [n]^{-2m+1} \sum_{1 \le i_1 \neq \ldots \neq i_m \le n} h^2(X_{i_1}, \ldots, X_{i_m}) \\
&\quad + [n]^{-2m+1} \sum_{j=2}^{m-1} \sum_{1 \le i_1 \neq \ldots \neq i_{2m-j} \le n} h(X_{i_1}, \ldots, X_{i_j}, X_{i_{j+1}}, \ldots, X_{i_m}) \\
&\qquad\qquad \times h(X_{i_1}, \ldots, X_{i_j}, X_{i_{m+1}}, \ldots, X_{i_{2m-j}}) \\
&\quad + [n]^{-2m+1} \sum_{1 \le i_1 \neq \ldots \neq i_{2m-1} \le n} h(X_{i_1}, X_{i_2}, \ldots, X_{i_m})\, h(X_{i_1}, X_{i_{m+1}}, \ldots, X_{i_{2m-1}}).
\end{aligned}
$$

The first term and the second one, which obviously does not appear when $m = 2$, in the latter equality will be seen to be negligible in probability (cf. Propositions 4.3.3 and 4.3.4), thus the third term becomes the main term that will play the main role in establishing Proposition 4.3.2.

To complete the proof of Proposition 4.3.2 we shall state and prove the next three results, namely Propositions 4.3.3, 4.3.4 and Theorem 4.3.1.

Proposition 4.3.3. *If $E|h(X_1,\ldots,X_m)|^{\frac{5}{3}} < \infty$, then, as $n \to \infty$,*

$$[n]^{-2m+1} \sum_{1 \le i_1 \ne \ldots \ne i_m \le n} h^2(X_{i_1},\ldots,X_{i_m}) \to 0 \quad a.s.$$

Proof of Proposition 4.3.3

From the fact that for $m \ge 2$, $\dfrac{2m}{2m-1} < \dfrac{5}{3}$, it follows that

$$E\left|h^2(X_1,\ldots,X_m)\right|^{\frac{m}{2m-1}} = E\left|h(X_1,\ldots,X_m)\right|^{\frac{2m}{2m-1}} < \infty.$$

By this the proof of Proposition 4.3.3 follows from Lemma 1.10.2.

Proposition 4.3.4. *For $m \ge 3$, If $E|h(X_1,\ldots,X_m)|^{\frac{5}{3}} < \infty$, then, as $n \to \infty$,*

$$[n]^{-2m+1} \sum_{j=2}^{m-1} \sum_{1 \le i_1 \ne \ldots \ne i_{2m-j} \le n} h(X_{i_1},\ldots,X_{i_j},X_{i_{j+1}},\ldots,X_{i_m}) \\ \times h(X_{i_1},\ldots,X_{i_j},X_{i_{m+1}},\ldots,X_{i_{2m-j}}) = o_P(1).$$

Proof of Proposition 4.3.4

In order to prove Proposition 4.3.4 it suffices to show that as $n \to \infty$, for $j = 2,\ldots,m-1$, we have

$$[n]^{-2m+1} \sum_{1 \le i_1 \ne \ldots \ne i_{2m-j} \le n} h(X_{i_1},\ldots,X_{i_j},X_{i_{j+1}},\ldots,X_{i_m}) \\ \times h(X_{i_1},\ldots,X_{i_j},X_{i_{m+1}},\ldots,X_{i_{2m-j}}) = o_P(1).$$

Since the proof of the latter relation can be done by modifying, mutatis mutandis, that of the next theorem, i.e., Theorem 4.3.1, hence the detailed proof is presented in Appendix of this chapter.

Theorem 4.3.1. *If $E|h(X_1,\ldots,X_m)|^{\frac{5}{3}} < \infty$ and $\tilde{h}_1(X_1) \in DAN$, then, as $n \to \infty$,*

$$\Big| \; [n]^{-2m+1} \sum_{1\le i_1\ne\ldots\ne i_{2m-1}\le n} h(X_{i_1},X_{i_2},\ldots,X_{i_m})h(X_{i_1},X_{i_{m+1}},\ldots,X_{i_{2m-1}})$$

$$-\frac{1}{n}\sum_{i=1}^{n}\tilde{h}_1^2(X_i) \; \Big| = o_P(1).$$

Proof of Theorem 4.3.1

Before stating the proof of Theorem 4.3.1 we note that the concept of complete degeneracy defined in Section 1.6 will play the key role in the proof of Theorem 4.3.1. Here we make use of the property of complete degeneracy of functions (summands) even though, here they will not be symmetric.

For further use in this proof, we consider the following setup:

$$\begin{aligned}
h_{1\ldots m} &:= h(X_1,\ldots,X_m),\\
h^{(m)}_{1\ldots m} &:= h_{1\ldots m}\,\mathbf{1}_{(|h|\le n^{\frac{3m}{5}})},\\
h^{*}_{12\ldots 2m-1} &:= h^{(m)}_{12\ldots m}\,h^{(m)}_{1m+1\ldots 2m-1},\\
\tilde{h}^{(m)}_1(x) &:= E\,(\,h^{(m)}_{1\ldots m}|X_1=x\,),\\
h^{(j)}_{1\ldots m} &:= h^{(m)}_{1\ldots m}\,\mathbf{1}_{(|h^{(m)}|\le n^{\frac{3j}{5}})},\ j=1,\ldots,m-1,\\
h^{(0)}_{1\ldots m} &:= h^{(m)}_{1\ldots m}\,\mathbf{1}_{(|h^{(m)}|\le \log(n))},\\
h^{(\ell)}_{1\ldots m} &:= h^{(m)}_{1\ldots m}\,\mathbf{1}_{(|\tilde{h}^{(m)}_1(x)|\le n^{1/2}\,\ell(n))},
\end{aligned}$$

where, again, $\mathbf{1}_A$ denotes the indicator function of the set A and $\ell(.)$ is a slowly varying function at infinity associated to $\tilde{h}_1(X_1)$.

In view of the above set up, observe that as $n\to\infty$

$$P\Big(\sum_{1\le i_1\ne\ldots\ne i_{2m-1}\le n} h_{i_1i_2\ldots i_m}\,h_{i_1i_{m+1}\ldots i_{2m-1}} \ne \sum_{1\le i_1\ne\ldots\ne i_{2m-1}\le n} h^{*}_{i_1\ldots i_{2m-1}}\Big)$$

$$\leq \; n^m \; P\,(\; |h_{1\dots m}| > \; n^{\frac{3m}{5}}\;)$$

$$\leq E\,[\,|h_{1\dots m}|^{\frac{5}{3}}\, 1_{(|h_{1\dots m}|>n^{\frac{3m}{5}})}\,] \longrightarrow 0.$$

Hence the asymptotic equivalency of $\sum_{1\leq i_1\neq\dots\neq i_{2m-1}\leq n} \; h_{i_1 i_2\dots i_m} \; h_{i_1 i_{m+1}\dots i_{2m-1}}$ and its truncated version i.e., $\sum_{1\leq i_1\neq\dots\neq i_{2m-1}\leq n} \; h^*_{i_1\dots i_{2m-1}}$ in probability.

Having the asymptotic equivalency of the original statistic and its truncated version, to prove Theorem 4.3.1, we shall proceed by working with the truncated version. Noting that due to lack of symmetry, the statistic of our interest, i.e., $\sum_{1\leq i_1\neq\dots\neq i_{2m-1}\leq n} \; h^*_{i_1\dots i_{2m-1}}$, is not a U-statistic, once again here, we extend the idea of Hoeffding procedure to represent U-statistics in terms of complete degenerate ones (cf. Section 1.7.1), in our context. This extension shall be done by creating complete degenerate statistics by adding and subtracting required terms. Then by employing proper new truncations and applying Lemma 1.8.1 we conclude the asymptotic negligibility of all of these *complete degenerate* statistics in probability (cf. Propositions 4.3.1, 4.3.2 and 4.3.3) except for the last group of them which are of the form of sums of i.i.d. random variables (cf. Remark 4.3.1 of this chapter). Among those the latter mentioned just one (cf. part (b) of Proposition 4.3.4) will asymptotically in probability coincide with $\dfrac{1}{n}\sum_{i=1}^{n}\tilde{h}_1^2(X_i)$ and that will complete the proof of Theorem 4.3.1.

Now as it was already mentioned, by adding and subtracting required terms we write

$$\sum_{1\leq i_1\neq\dots\neq i_{2m-1}\leq n} \; h^*_{i_1\dots i_{2m-1}}$$

$$= \sum_{1\le i_1 \ne \ldots \ne i_{2m-1}\le n} \Big\{ \sum_{d=1}^{2m-1} (-1)^{2m-1-d} \sum_{1\le j_1 < \ldots < j_d \le 2m-1} E(h^*_{i_1 \ldots i_{2m-1}} - E(h^*_{i_1 \ldots i_{2m-1}})|\ X_{i_{j_1}}, \ldots X_{i_{j_d}})$$

$$+ \sum_{c=1}^{2m-2} \sum_{1\le k_1 < \ldots < k_c \le 2m-1} \sum_{d=1}^{c} (-1)^{c-d} \sum_{1\le j_1 < \ldots < j_d \le c} E(h^*_{i_1 \ldots i_{2m-1}} - E(h^*_{i_1 \ldots i_{2m-1}})|\ X_{i_{k_{j_1}}}, \ldots X_{i_{k_{j_d}}})$$

$$+ E\ (h^*_{i_1 \ldots i_{2m-1}}) \Big\}$$

$$:= \sum_{1\le i_1 \ne \ldots \ne i_{2m-1}\le n} V(i_1, \ldots, i_{2m-1}) + \sum_{c=1}^{2m-2} \sum_{1\le k_1 < \ldots < k_c \le 2m-1} \sum_{1\le i_1 \ne \ldots \ne i_{2m-1}\le n} V(i_{k_1}, \ldots, i_{k_c})$$

$$+ \sum_{1\le i_1 \ne \ldots \ne i_{2m-1}\le n} E(h^*_{i_1 \ldots i_{2m-1}}).$$

Proposition 4.3.1. *If* $E\ |h_{1\ldots m}|^{\frac{5}{3}} < \infty$, *then, as* $n \to \infty$,

$$[n]^{-2m+1} \sum_{1\le i_1 \ne \ldots \ne i_{2m-1}\le n} V(i_1, \ldots, i_{2m-1})\ = o_P(1).$$

Proof of Proposition 4.3.1

Since $V(i_1, \ldots, i_{2m-1})$ posses the property of degeneracy we can apply Lemma 1.8.1 for the associated statistics and write, for $\epsilon > 0$,

$$\begin{aligned}
& P\ \Big(\ |\ [n]^{-2m+1} \sum_{1\le i_1 \ne \ldots \ne i_{2m-1}\le n} V(i_1, \ldots, i_{2m-1})| > \epsilon\Big) \\
\le\ & \epsilon^{-2}\ E\ [\ [n]^{-2m+1} \sum_{1\le i_1 \ne \ldots \ne i_{2m-1}\le n} V(i_1, \ldots, i_{2m-1})\ ]^2 \\
\le\ & \epsilon^{-2}\ [n]^{-2m+1}\ E\ [\ V(1, \ldots, 2m-1)\ ]^2 \\
\le\ & A\ \epsilon^{-2}\ [n]^{-2m+1}\ n^{2m-1}\ n^{-2m+1}\ E\ [\ h^{(m)}_{12\ldots m}\ h^{(m)}_{1m+1\ldots 2m-1}\ ]^2 \\
\le\ & A\ \epsilon^{-2}\ [n]^{-2m+1}\ n^{2m-1}\ n^{-2m+1}\ n^{\frac{7m}{5}}\ E\ |\ h_{12\ldots m}\ |^{\frac{5}{3}}
\end{aligned}$$

$$\longrightarrow 0, \text{ as } n \to \infty.$$

The estimation for $m \geq 3$ that occurs in our next proposition does not appear, and hence not needed, when $m = 2$.

Proposition 4.3.2. *For $m \geq 3$, if $E \mid h_{12\ldots m} \mid^{\frac{5}{3}} < \infty$, then, as $n \to \infty$*

$$[n]^{-2m+1} \sum_{1 \leq i_1 \neq \ldots \neq i_{2m-1} \leq n} V(i_{k_1}, \ldots, i_{k_c}) \; = o_P(1),$$

where $c = 3, \ldots, 2m-2$ and $1 \leq k_1 < \ldots < k_c \leq 2m-1$.

Proof of Proposition 4.3.2

Based on the way $i_{k_1}, \ldots, i_{k_c}$ are distributed between $h^{(m)}_{i_1 i_2 \ldots i_m}$ and $h^{(m)}_{i_1 i_{m+1} \ldots i_{2m-1}}$ in two different cases when $k_1 = 1$ and $k_1 \neq 1$, the proof is stated as follows.

Case $k_1 = 1$

Let s and t be respectively the number of elements of the sets

$$\{i_{k_1}, \ldots, i_{k_c}\} \cap \{i_1, i_2, \ldots, i_m\} \text{ and } \{i_{k_1}, \ldots, i_{k_c}\} \cap \{i_1, i_{m+1}, \ldots, i_{2m-1}\}.$$

It is clear that in this case, i.e., $k_1 = 1$, we have that $s, t \geq 1$ and $s + t = c + 1$. Now define

$$V^T(i_{k_1}, \ldots, i_{k_c}) = \sum_{d=1}^{c} (-1)^{c-d} \sum_{1 \leq j_1 < \ldots < j_d \leq c} E(h^{*^T}_{i_1 \ldots i_{2m-1}} - E(h^{*^T}_{i_1 \ldots i_{2m-1}}) \mid X_{i_{k_{j_1}}}, \ldots, X_{i_{k_{j_d}}}), \tag{4.3.4}$$

$$V^{T'}(i_{k_1}, \ldots, i_{k_c}) = \sum_{d=1}^{c} (-1)^{c-d} \sum_{1 \leq j_1 < \ldots < j_d \leq c} E(h^{*^{T'}}_{i_1 \ldots i_{2m-1}} - E(h^{*^{T'}}_{i_1 \ldots i_{2m-1}}) \mid X_{i_{k_{j_1}}}, \ldots, X_{i_{k_{j_d}}}), \tag{4.3.5}$$

where $h^{*^T}_{i_1 \ldots i_{2m-1}} = h^{(s)}_{i_1 i_2 \ldots i_m} h^{(m)}_{i_1 i_{m+1} \ldots i_{2m-1}}$ and $h^{*^{T'}}_{i_1 \ldots i_{2m-1}} = h^{(s)}_{i_1 i_2 \ldots i_m} h^{(t)}_{i_1 i_{m+1} \ldots i_{2m-1}}$.

Now observe that as, $n \to \infty$,

$$P\Big(\sum_{1\le i_1\ne \ldots \ne i_{2m-1}\le n} V(i_{k_1},\ldots,i_{k_c}) \ne \sum_{1\le i_1\ne \ldots \ne i_{2m-1}\le n} V^{T'}(i_{k_1},\ldots,i_{k_c}) \Big)$$

$$\le P\Big(\sum_{1\le i_1\ne \ldots \ne i_{2m-1}\le n} V(i_{k_1},\ldots,i_{k_c}) \ne \sum_{1\le i_1\ne \ldots \ne i_{2m-1}\le n} V^{T}(i_{k_1},\ldots,i_{k_c}) \Big)$$

$$+ P\Big(\sum_{1\le i_1\ne \ldots \ne i_{2m-1}\le n} V^{T}(i_{k_1},\ldots,i_{k_c}) \ne \sum_{1\le i_1\ne \ldots \ne i_{2m-1}\le n} V^{T'}(i_{k_1},\ldots,i_{k_c}) \Big)$$

$$\le \; n^s \, P\,(\, |h^{(m)}_{12\ldots m}| > \; n^{\frac{3s}{5}} \,) + \; n^t \, P\,(\, |h^{(m)}_{1m+1\ldots 2m-1}| > \; n^{\frac{3t}{5}} \,)$$

$$\le \; E\,[\, |h_{12\ldots m}|^{\frac{5}{3}} \, \mathbf{1}_{(|h|>n^{\frac{3s}{5}})} \,] + \; E\,[\, |h_{1m+1\ldots 2m-1}|^{\frac{5}{3}} \, \mathbf{1}_{(|h|>n^{\frac{3t}{5}})} \,] \longrightarrow 0.$$

The latter relation suggests that $\sum_{1\le i_1\ne \ldots \ne i_{2m-1}\le n} V(i_{k_1},\ldots,i_{k_c})$ and $\sum_{1\le i_1\ne \ldots \ne i_{2m-1}\le n} V^{T'}(i_{k_1},\ldots,i_{k_c})$ are asymptotically equivalent in probability.

Since $V^{T'}(i_{k_1},\ldots,i_{k_c})$ is complete degenerate, Markov's inequality followed by an application of Lemma 1.8.1 yields,

$$\begin{aligned}
P \quad & \Big(\, \Big| \, [n]^{-2m+1} \sum_{1\le i_1\ne \ldots \ne i_{2m-1}\le n} V^{T'}(i_{k_1},\ldots,i_{k_c}) \,\Big| \, > \epsilon \,\Big) \\
\le \quad & \epsilon^{-2} \, E\Big(\, [n]^{-2m+1} \sum_{1\le i_1\ne \ldots \ne i_{2m-1}\le n} V^{T'}(i_{k_1},\ldots,i_{k_c}) \,\Big)^2 \\
\le \quad & A\epsilon^{-2} \, [n-(2m-1-c)]^{-c} \, E \,\Big(\, h^{(s)}_{12\ldots m} \, h^{(t)}_{1m+1\ldots 2m-1} \,\Big)^2 \\
\le \quad & A\epsilon^{-2} \, [n-(2m-1-c)]^{-c} \, n^c \, n^{-c} \, n^{\frac{7(t+s)}{10}} \; E \,|\, h_{12\ldots m} \,|^{\frac{5}{3}}
\end{aligned}$$

$$\longrightarrow 0, \; \text{ as } n \to \infty.$$

The latter relation is true since when $c \ge 3$, we have $-c + \dfrac{7(t+s)}{10} < 0$.

Case $k_1 \ne 1$

Similarly to the previous case let s and t be respectively the number of elements of the sets $\{i_{k_1},\dots,i_{k_c}\}\cap\{i_1,i_2,\dots,i_m\}$ and $\{i_{k_1},\dots,i_{k_c}\}\cap\{i_1,i_{m+1},\dots,i_{2m-1}\}$. Clearly here we have $s,t\geq 0$ and $s+t=c$. It is obvious that in this case s,t can be zero but not simultaneously. More specifically, either $(s=c,t=0)$ or $(s=0,t=c)$ can happen and due to their similarity we shall only treat $(s=c,t=0)$.

Let $V^T(i_{k_1},\dots,i_{k_c})$ and $V^{T'}(i_{k_1},\dots,i_{k_c})$ be of the forms respectively (4.3.4) and (4.3.5), where

$$h^{*T}_{i_1\dots i_{2m-1}} = h^{(s)}_{i_1i_2\dots i_m}\ h^{(m)}_{i_1i_{m+1}\dots i_{2m-1}}$$

and

$$h^{*T'}_{i_1\dots i_{2m-1}} = h^{(s)}_{i_1i_2\dots i_m}\ h^{(t)}_{i_1i_{m+1}\dots i_{2m-1}}.$$

Observe that as $n\to\infty$,

$$P\Big(\sum_{1\leq i_1\neq\dots\neq i_{2m-1}\leq n} V(i_{k_1},\dots,i_{k_c}) \neq \sum_{1\leq i_1\neq\dots\neq i_{2m-1}\leq n} V^{T'}(i_{k_1},\dots,i_{k_c})\Big)$$

$$\leq \begin{cases} n^s\ P\big(|h^{(m)}_{12\dots m}| > n^{\frac{3s}{5}}\big) + n^t\ P\big(|h^{(m)}_{1m+1\dots 2m-1}| > n^{\frac{3t}{5}}\big), & \text{s,t}>0,\ \text{s+t=c} \\ n^c\ P\big(|h^{(m)}_{12\dots m}| > n^{\frac{3c}{5}}\big) + P\big(|h^{(m)}_{1m+1\dots 2m-1}| > \log(n)\big), & \text{s=c,t=0} \end{cases}$$

$$\leq \begin{cases} E\big(|h_{12\dots m}|^{\frac{5}{3}}\ \mathbf{1}_{(|h|>n^{\frac{3s}{5}})}\big) + E\big(|h_{1m+1\dots 2m-1}|^{\frac{5}{3}}\ \mathbf{1}_{(|h|>n^{\frac{3t}{5}})}\big), & \text{s,t} > 0,\ \text{s+t=c} \\ E\big[\,|h_{12\dots m}|^{\frac{5}{3}}\ \mathbf{1}_{(|h|>n^{\frac{3c}{5}})}\big] + P\big(|h^{(m)}_{1m+1\dots 2m-1}| > \log(n)\big), & \text{s=c,t=0} \end{cases}$$

$\longrightarrow 0.$

Applying Markov's inequality followed by an application of Lemma 1.8.1 once again yields,

$$P\Big(\Big|\,[n]^{-2m+1}\sum_{1\leq i_1\neq\dots\neq i_{2m-1}\leq n} V^{T'}(i_{k_1},\dots,i_{k_c})\Big| > \epsilon\Big)$$

$$\leq A\epsilon^{-2}\,[n-(2m-1-c)]^{-c}\ n^c\ n^{-c}\ E\Big(h^{(s)}_{12\dots m}\ h^{(t)}_{1m+1\dots 2m-1}\Big)^2$$

$$\leq \begin{cases} A\epsilon^{-2}\,[n-(2m-1-c)]^{-c}\; n^{c}\; n^{-c}\; n^{\frac{7c}{10}}\; E|h_{12\ldots m}|^{\frac{5}{3}}, & \text{s,t} > 0,\ \text{s+t=c} \\ A\epsilon^{-2}\,[n-(2m-1-c)]^{-c}\; n^{c}\; n^{-c}\; n^{\frac{7c}{10}}\; \log^{\frac{7}{6}}(n)\; E|h_{12\ldots m}|^{\frac{5}{3}}, & \text{s=c,t=0} \end{cases}$$

$\longrightarrow 0$, as $n \to \infty$.

This completes the proof of Proposition 4.3.2.

Proposition 4.3.3. *If $E\ |h_{12\ldots m}|^{\frac{5}{3}} < \infty$ and $\tilde{h}_1(X_1) \in DAN$, then, as $n \to \infty$,*

$$[n]^{-2m+1} \sum_{1 \leq i_1 \neq \ldots \neq i_{2m-1} \leq n} V(i_{k_1}, i_{k_2}) = o_P(1),$$

where, $1 \leq k_1 < k_2 \leq 2m-1$.

Proof of Proposition 4.3.3

As it was the case in the proof of the last proposition, we shall state the proof for two cases $k_1 = 1$ and $k_1 \neq 1$ separately.

Case $k_1 = 1$

Again let s and t be respectively the number of elements of the sets $\{i_{k_1}, i_{k_2}\} \cap \{i_1, i_2, \ldots, i_m\}$ and $\{i_{k_1}, i_{k_2}\} \cap \{i_1, i_{m+1}, \ldots, i_{2m-1}\}$. It is clear that in this case we either have $(s = 2, t = 1)$ or $(s = 1, t = 2)$ which due to their similarity only $(s = 2, t = 1)$ will be treated as follows.

Define

$$V^{T}(i_{k_1}, i_{k_2}) = \sum_{d=1}^{2} (-1)^{2-d} \sum_{1 \leq j_1 < \ldots < j_d \leq 2} E(h^{*T}_{i_1 \ldots i_{2m-1}} - E(h^{*T}_{i_1 \ldots i_{2m-1}}) \mid X_{i_{k_{j_1}}}, \ldots, X_{i_{k_{j_d}}}),$$

$$V^{T'}(i_{k_1}, i_{k_2}) = \sum_{d=1}^{2} (-1)^{2-d} \sum_{1 \leq j_1 < \ldots < j_d \leq 2} E(h^{*T'}_{i_1 \ldots i_{2m-1}} - E(h^{*T'}_{i_1 \ldots i_{2m-1}}) \mid X_{i_{k_{j_1}}}, \ldots, X_{i_{k_{j_d}}}),$$

where $h^{*T}_{i_1\ldots i_{2m-1}} = h^{(2)}_{i_1 i_2\ldots i_m}\ h^{(m)}_{i_1 i_{m+1}\ldots i_{2m-1}}$ and $h^{*T'}_{i_1\ldots i_{2m-1}} = h^{(2)}_{i_1 i_2\ldots i_m}\ h^{(\ell)}_{i_1 i_{m+1}\ldots i_{2m-1}}$.

Having the above setup, as $n \to \infty$, we have

$$\begin{aligned}
P & \Big(\sum_{1\le i_1\neq\ldots\neq i_{2m-1}\le n} V(i_{k_1}, i_{k_2}) \neq \sum_{1\le i_1\neq\ldots\neq i_{2m-1}\le n} V^{T'}(i_{k_1}, i_{k_2}) \Big) \\
\le & \ P\Big(\sum_{1\le i_1\neq\ldots\neq i_{2m-1}\le n} V(i_{k_1}, i_{k_2}) \neq \sum_{1\le i_1\neq\ldots\neq i_{2m-1}\le n} V^{T}(i_{k_1}, i_{k_2}) \Big) \\
+ & \ P\Big(\sum_{1\le i_1\neq\ldots\neq i_{2m-1}\le n} V^{T}(i_{k_1}, i_{k_2}) \neq \sum_{1\le i_1\neq\ldots\neq i_{2m-1}\le n} V^{T'}(i_{k_1}, i_{k_2}) \Big) \\
\le & \ n^2\ P(\ |h^{(m)}_{12\ldots m}| > n^{6/5}\ ) + \ n\ P(\ |\tilde{h}^{(m)}_1(X_1)| > n^{1/2}\ \ell(n)\ ) \\
\le & \ E\ \big(\ |h_{12\ldots m}|^{\frac{5}{3}}\ \mathbf{1}_{(|h|>n^{6/5})}\ \big) + \ n\ P(\ |\tilde{h}^{(m)}_1(X_1)| > n^{1/2}\ \ell(n)\ ) \\
:= & \ I_1(n) + \ I_2(n).
\end{aligned}$$

It can be easily seen that as n tends to infinity $I_1(n) \to 0$.

To deal with $I_2(n)$ we write

$$\begin{aligned}
& n\ P\ (\ |\tilde{h}^{(m)}_1(X_1)| > n^{1/2}\ \ell(n)\ ) \\
& \le\ n\ P\ (\ |\tilde{h}_1(X_1)| > \frac{n^{1/2}\ \ell(n)}{2}\ ) \\
& \quad + n\ P\ (\ |\ E(h_{1m+1\ldots 2m-1}\ \mathbf{1}_{(|h|>n^{\frac{3m}{5}})}\ |\ X_1)\ | > \frac{n^{1/2}\ \ell(n)}{2}) \\
& \le\ n\ P\ (\ |\tilde{h}_1(X_1)| > \frac{n^{1/2}\ \ell(n)}{2}\ ) \\
& \quad + 2\ n^{1/2}\ \ell^{-1}(n)\ E\ \big(\ |\ h_{1m+1\ldots 2m-1}\ |\ \mathbf{1}_{(|h|>n^{\frac{3m}{5}})}\ \big) \\
& \le\ n\ P\ (\ |\tilde{h}_1(X_1)| > \frac{n^{1/2}\ \ell(n)}{2}\ ) \\
& \quad + 2\ n^{1/2}\ n^{-\frac{2m}{5}}\ \ell^{-1}(n)\ E|\ h_{1m+1\ldots 2m-1}\ |^{\frac{5}{3}} \\
& \longrightarrow 0,\ \text{ as } n \to \infty.
\end{aligned}$$

The latter relation is true since $\tilde{h}_1(X_1) \in DAN$ and $m \geq 2$, and it means that $I_2(n) = o(1)$. Hence the asymptotic equivalency of $\sum_{1\leq i_1\neq \ldots \neq i_{2m-1}\leq n} V(i_{k_1}, i_{k_2})$ and $\sum_{1\leq i_1\neq \ldots \neq i_{2m-1}\leq n} V^{T'}(i_{k_1}, i_{k_2})$ in probability.

Before applying Proposition 1.8.1 for $[n]^{-2m+1} \sum_{1\leq i_1\neq \ldots \neq i_{2m-1}\leq n} V^{T'}(i_{k_1}, i_{k_2})$, since we know that $k_1 = 1$ and $s = 2$, due to symmetry of $h_{i_1 i_2 \ldots i_m}$, without loss of generality we assume that $k_2 = 2$.

Now for $\epsilon > 0$, Markov's inequality and Lemma 1.8.1 lead to

$$P(\,|\,[n]^{-2m+1} \sum_{1\leq i_1\neq \ldots \neq i_{2m-1}\leq n} V^{T'}(i_1, i_2)\,| > \epsilon\,)$$

$$\leq\; A\epsilon^{-2}[n-(2m-3)]^{-2} E\Big(E\big(h^{(2)}_{12\ldots m} h^{(\ell)}_{1m+1\ldots 2m-1} - E(h^{(2)}_{12\ldots m} h^{(\ell)}_{1m+1\ldots 2m-1})|X_1, X_2\big)\Big)^2$$

$$+A\,\epsilon^{-2}\,[n-(2m-3)]^{-2}\;\; E\,\Big(E\big(h^{(2)}_{12\ldots m} h^{(\ell)}_{1m+1\ldots 2m-1} - E(h^{(2)}_{12\ldots m} h^{(\ell)}_{1m+1\ldots 2m-1})|X_1\big)\Big)^2$$

$$+A\,\epsilon^{-2}\,[n-(2m-3)]^{-2}\;\; E\,\Big(E\big(h^{(2)}_{12\ldots m} h^{(\ell)}_{1m+1\ldots 2m-1} - E(h^{(2)}_{12\ldots m} h^{(\ell)}_{1m+1\ldots 2m-1})|X_2\big)\Big)^2$$

$$:=\; A\,\epsilon^{-2}\,[n-(2m-3)]^{-2}\,n^2\,J_1(n)$$

$$+\;A\,\epsilon^{-2}\,[n-(2m-3)]^{-2}\,n^2\,J_2(n)$$

$$+\;A\,\epsilon^{-2}\,[n-(2m-3)]^{-2}\,n^2\,J_3(n).$$

Considering that as $n \to \infty$, $[n-(2m-3)]^{-2}\,n^2 \to 1$, we will show that

$$J_1(n), J_2(n), J_3(n) = o(1).$$

To deal with $J_1(n)$ write

$$J_1(n) \leq\; n^{-2}\,E\big(\;E(\;h^{(2)}_{12\ldots m} h^{(\ell)}_{1m+1\ldots 2m-1} \mid X_1, X_2)\;\big)^2$$

$$= n^{-2}\, E\big(\, E^2(h^{(2)}_{12\ldots m}|\; X_1, X_2)\; E^2(h^{(\ell)}_{1m+1\ldots 2m-1}|\; X_1)\;\big)$$

$$= n^{-2}\, E\big(\, E^2(h^{(2)}_{12\ldots m}|\; X_1, X_2)\; E^2(h^{(m)}_{1m+1\ldots 2m-1}|\; X_1)\; \mathbf{1}_{(|\tilde{h}^{(m)}_1(X_1)| \le n^{1/2}\; \ell(n))}\;\big)$$

$$\le n^{-1}\, \ell^2(n)\; E\big(\; h^{(2)}_{12\ldots m}\big)^2$$

$$\le n^{-\frac{3}{5}}\, \ell^2(n)\; E|\; h_{12\ldots m}|^{\frac{5}{3}}$$

$$\longrightarrow 0, \text{ as } n \to \infty,$$

i.e., $J_1(n) = o(1)$. A similar argument yields, $J_2(n) = o(1)$, hence the details are omitted.

As for $J_3(n)$ we write

$$\begin{aligned} J_3(n) &\le n^{-2}\, E\big(\; E(h^{(2)}_{12\ldots m} h^{(\ell)}_{1m+1\ldots 2m-1} \mid X_2)\;\big)^2 \\ &= n^{-2}\, E\{\; E\left(E(h^{(2)}_{12\ldots m} h^{(\ell)}_{1m+1\ldots 2m-1}|X_1, \ldots, X_m) \mid X_2\;\right)\}^2 \\ &= n^{-2}\, E\{\; E(h^{(2)}_{12\ldots m}|\; X_2)\; E(h^{(\ell)}_{1m+1\ldots 2m-1}|\; X_1)\}^2 \\ &\le n^{-\frac{3}{5}}\, \ell^2(n)\; E|\; h_{12\ldots m}|^{\frac{5}{3}} \end{aligned}$$

$$\longrightarrow 0, \text{ as } n \to \infty.$$

The latter relation means that $J_3(n) = o(1)$. By this the proof of Proposition 4.3.3, when $k_1 = 1$, is complete.

At this stage we state the proof of Proposition 4.3.3, when $k_1 \neq 1$.

Case $k_1 \neq 1$

Once again let s and t be respectively the number of elements of the sets $\{i_{k_1}, i_{k_2}\} \cap \{i_1, i_2, \ldots, i_m\}$ and $\{i_{k_1}, i_{k_2}\} \cap \{i_1, i_{m+1}, \ldots, i_{2m-1}\}$. It is obvious that in this case the possibilities are $s = t = 1$ and when $m \ge 3$, $(s = 2, t = 0)$ or $(s = 0, t = 2)$. We shall treat the cases $s = t = 1$ and when $m \ge 3$, $(s = 2, t = 0)$, separately as follows.

Case $k_1 \neq 1$*:* $s = t = 1$

We note that here we have $k_1 \in \{2, \ldots, m\}$ and $k_2 \in \{m+1, \ldots, 2m-1\}$.

Now define

$$V^T(i_{k_1}, i_{k_2}) = \sum_{d=1}^{2} (-1)^{2-d} \sum_{1 \le j_1 < \ldots < j_d \le 2} E(h^{*T}_{i_1 \ldots i_{2m-1}} - E(h^{*T}_{i_1 \ldots i_{2m-1}}) \mid X_{i_{k_{j_1}}}, \ldots, X_{i_{k_{j_d}}}),$$

$$V^{T'}(i_{k_1}, i_{k_2}) = \sum_{d=1}^{2} (-1)^{2-d} \sum_{1 \le j_1 < \ldots < j_d \le 2} E(h^{*T'}_{i_1 \ldots i_{2m-1}} - E(h^{*T'}_{i_1 \ldots i_{2m-1}}) \mid X_{i_{k_{j_1}}}, \ldots, X_{i_{k_{j_d}}}),$$

where $h^{*T}_{i_1 \ldots i_{2m-1}} = h^{(1)}_{i_1 i_2 \ldots i_m}\ h^{(m)}_{i_1 i_{m+1} \ldots i_{2m-1}}$ and $h^{*T'}_{i_1 \ldots i_{2m-1}} = h^{(1)}_{i_1 i_2 \ldots i_m}\ h^{(1)}_{i_1 i_{m+1} \ldots i_{2m-1}}$. Now observe that as $n \to \infty$, we have

$$\begin{aligned}
P &\Big(\sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V(i_{k_1}, i_{k_2}) \ne \sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V^{T'}(i_{k_1}, i_{k_2}) \Big) \\
\le & \ P\Big(\sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V(i_{k_1}, i_{k_2}) \ne \sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V^{T}(i_{k_1}, i_{k_2}) \Big) \\
+ & \ P\Big(\sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V^T(i_{k_1}, i_{k_2}) \ne \sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V^{T'}(i_{k_1}, i_{k_2}) \Big) \\
\le & \ 2\, n\, P\big(|h^{(m)}_{12\ldots m}| > n^{3/5} \big) \\
\le & \ 2\, E\big[\, |h_{12\ldots m}|^{\frac{5}{3}}\, \mathbf{1}_{(|h| > n^{3/5})} \,\big]
\end{aligned}$$

$$\longrightarrow 0.$$

In view of the latter relation we apply Proposition 1.8.1 to the sum

$$[n]^{-2m+1} \sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V^{T'}(i_{k_1}, i_{k_2})$$

and we get

$$P\big([n]^{-2m+1} \sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} \big| V^{T'}(i_{k_1}, i_{k_2}) \big| > \epsilon \big)$$

$$\leq A\,\epsilon^{-2}\,[n-(2m-3)]^{-2}\,n^{2}\,n^{-2}\,E(h^{(1)}_{12\ldots m}h^{(1)}_{1m+1\ldots 2m-1})^{2}$$

$$\leq A\,\epsilon^{-2}\,[n-(2m-3)]^{-2}\,n^{2}\,n^{-2}\,n^{7/5}E|h_{12\ldots m}|^{\frac{5}{3}}$$

$$\longrightarrow 0,\ \text{as } n\to\infty.$$

This completes the proof of Proposition 4.3.3 for the Case $k_1\neq 1$ when $s=t=1$.

Case $k_1\neq 1$: ($m\geq 3$) $s=2, t=0$

In this case we first note that $k_1,k_2\in\{2,\ldots,m\}$. Now define

$$V^{T}(i_{k_1},i_{k_2})=\sum_{d=1}^{2}(-1)^{2-d}\sum_{1\leq j_1<\ldots<j_d\leq 2}E(h^{*^{T}}_{i_1\ldots i_{2m-1}}-E(h^{*^{T}}_{i_1\ldots i_{2m-1}})\mid X_{i_{k_{j_1}}},\ldots,X_{i_{k_{j_d}}}),$$

$$V^{T'}(i_{k_1},i_{k_2})=\sum_{d=1}^{2}(-1)^{2-d}\sum_{1\leq j_1<\ldots<j_d\leq 2}E(h^{*^{T'}}_{i_1\ldots i_{2m-1}}-E(h^{*^{T'}}_{i_1\ldots i_{2m-1}})\mid X_{i_{k_{j_1}}},\ldots,X_{i_{k_{j_d}}}),$$

where $h^{*^{T}}_{i_1\ldots i_{2m-1}}=h^{(2)}_{i_1i_2\ldots i_m}\,h^{(m)}_{i_1i_{m+1}\ldots i_{2m-1}}$ and $h^{*^{T'}}_{i_1\ldots i_{2m-1}}=h^{(2)}_{i_1i_2\ldots i_m}\,h^{(0)}_{i_1i_{m+1}\ldots i_{2m-1}}$. Now observe that as $n\to\infty$

$$\begin{aligned}
&P\ \Big(\sum_{1\leq i_1\neq\ldots\neq i_{2m-1}\leq n}V(i_{k_1},i_{k_2})\neq\sum_{1\leq i_1\neq\ldots\neq i_{2m-1}\leq n}V^{T'}(i_{k_1},i_{k_2})\Big)\\
&\leq\ P\Big(\sum_{1\leq i_1\neq\ldots\neq i_{2m-1}\leq n}V(i_{k_1},i_{k_2})\neq\sum_{1\leq i_1\neq\ldots\neq i_{2m-1}\leq n}V^{T}(i_{k_1},i_{k_2})\Big)\\
&+\ P\Big(\sum_{1\leq i_1\neq\ldots\neq i_{2m-1}\leq n}V^{T}(i_{k_1},i_{k_2})\neq\sum_{1\leq i_1\neq\ldots\neq i_{2m-1}\leq n}V^{T'}(i_{k_1},i_{k_2})\Big)\\
&\leq\ n^{2}\,P(\,|h^{(m)}_{12\ldots m}|>n^{6/5}\,)+\ P(\,|h^{(m)}_{1m+1\ldots 2m-1}|>\log(n)\,)\\
&\leq\ E\big(\,|\,h_{12\ldots m}\,|^{\frac{5}{3}}\,\mathbf{1}_{(|h|>n^{6/5})}\,\big)+\ P(\,|h_{1m+1\ldots 2m-1}|>\log(n)\,)
\end{aligned}$$

$$\longrightarrow 0.$$

The latter relation together with degeneracy of $V^{T'}(i_{k_1},i_{k_2})$ enable us to use Lemma

1.8.1 once again and arrive at

$$P\big([n]^{-2m+1} \sum_{1\le i_1\ne\ldots\ne i_{2m-1}\le n} \big|V^{T'}(i_{k_1}, i_{k_2})\big| > \epsilon\big)$$

$$\le A\, \epsilon^{-2}\, [n-(2m-3)]^{-2}\, n^2\, n^{-2}\, E(h^{(2)}_{12\ldots m}\, h^{(0)}_{1m+1\ldots 2m-1})^2$$

$$\le A\, \epsilon^{-2}\, [n-(2m-3)]^{-2}\, n^2\, n^{-\frac{3}{5}}\, \log^{7/6}(n)\, E|h_{12\ldots m}|^{\frac{5}{3}}$$

$$\longrightarrow 0, \text{ as } n \to \infty.$$

Now the proof of Proposition 4.3.3 is complete. □

Remark 4.3.1. *Before stating our next result we note in passing that when* $k_1 = 1$ *then,* $[n]^{-2m+1} \sum_{1\le i_1\ne\ldots\ne i_{2m-1}\le n} V(i_{k_1})$ *is of the form*

$$[n-(2m-2)]^{-1} \sum_{i_1\in\{1,\ldots,n\}/\{2,\ldots,2m-1\}}^{n} E(h^*_{i_1 2\ldots 2m-1} - E(h^*_{i_1 2\ldots 2m-1})\ |X_{i_1}),$$

otherwise, i.e., when for example $k_1 = 2$ *it has the following form*

$$[n-(2m-2)]^{-1} \sum_{i_2\in\{1,\ldots,n\}/\{1,3,\ldots,2m-1\}}^{n} E(h^*_{1i_2 3\ldots 2m-1} - E(h^*_{1i_2 3\ldots 2m-1})\ |X_{i_2}),$$

and so on for $k_1 \in \{2,\ldots,2m-1\}$.

Proposition 4.3.4. *If* $E|h_{1\ldots m}|^{\frac{5}{3}} < \infty$ *and* $\tilde{h}_1(X_1) \in DAN$*, then, as* $n \to \infty$

(a) $[n]^{-2m+1} \sum_{1\le\ne i_1\ne\ldots\ne i_{2m-1}\le n} V(i_{k_1}) = o_P(1),$ *for* $k_1 \in \{2,\ldots,2m-1\}$,

$$(b)\ \Big|\ [n-(2m-2)]^{-1} \sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-1\}}^{n} E(h^*_{i2\ldots 2m-1} - E(h^*_{i2\ldots 2m-1})\ |X_i)$$

$$+ E(h^*_{12\ldots 2m-1}) - \frac{1}{n}\sum_{i=1}^{n} \tilde{h}_1^2(X_i)\ \Big| = o_P(1).$$

Proof of Proposition 4.3.4

First we give the proof of part (a). Due to similarities, we state the proof only for $k_1 = 2$.

Define

$$V^T(i_2) = E(h^{*T}_{i_1 i_2 \ldots i_{2m-1}} - E(h^{*T}_{i_1 i_2 \ldots i_{2m-1}}) |\, X_{i_2}),$$

$$V^{T'}(i_2) = E(h^{*T'}_{i_1 i_2 \ldots i_{2m-1}} - E(h^{*T'}_{i_1 i_2 \ldots i_{2m-1}}) |\, X_{i_2}),$$

where $h^{*T}_{i_1 i_2 \ldots i_{2m-1}} = h^{(1)}_{i_1 i_2 \ldots i_m}\, h^{(m)}_{i_1 i_{m+1} \ldots i_{2m-1}}$ and $h^{*T'}_{i_1 i_2 \ldots i_{2m-1}} = h^{(1)}_{i_1 i_2 \ldots i_m}\, h^{(0)}_{i_1 i_{m+1} \ldots i_{2m-1}}$. Again observe that as $n \to \infty$,

$$P(\sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V(i_2) \ne \sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V^{T'}(i_2))$$

$$\le P(\sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V(i_2) \ne \sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V^{T}(i_2))$$

$$+ P(\sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V^T(i_2) \ne \sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V^{T'}(i_2))$$

$$\le n\ P(|h^{(m)}_{12\ldots m}| > n^{3/5}) + P(|h^{(m)}_{1m+1\ldots 2m-1}| > \log(n))$$

$$\le E(\ |h_{12\ldots m}|^{\frac{5}{3}}\ \mathbf{1}_{(|h| > n^{3/5})}\ ) + P(|h_{1m+1\ldots 2m-1}| > \log(n))$$

$$\longrightarrow 0.$$

An application of Markov's inequality yields

$$P\ (\ |\ [n]^{-2m+1} \sum_{1 \le i_1 \ne \ldots \ne i_{2m-1} \le n} V^{T'}(i_2)\ | > \epsilon)$$

$$\le A\epsilon^{-2}\ [n - (2m-2)]^{-1}\ n\ n^{-1}\ E(h^{(1)}_{12\ldots m}\ h^{(0)}_{1m+1\ldots 2m-1})^2$$

$$\le A\epsilon^{-2}\ [n - (2m-2)]^{-1}\ n\ n^{-\frac{3}{10}}\ \log^{7/6}(n)\ E|h_{12\ldots m}|^{\frac{5}{3}}$$

$\longrightarrow 0, \text{ as } n \to \infty.$

This complete the proof of part (a).

In the final stage of our proofs, to prove part (b) first define

$$\tilde{h^*}(x) = E(h_{12\ldots m}\, \mathbf{1}_{(|h|>n^{\frac{3m}{5}})} |X_1 = x)$$

and write

$$\Big|\frac{1}{n-2m+2} \sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-1\}}^{n} E(h^*_{i2\ldots 2m-1} - E(h^*_{i2\ldots 2m-1})\ |X_i)$$

$$+ E(h^*_{12\ldots 2m-1}) - \frac{1}{n}\sum_{i=1}^{n} \tilde{h}_1^2(X_i)\ \Big|$$

$$= \Big|\frac{1}{n-2m+2} \sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-1\}}^{n} E(h^*_{i2\ldots 2m-1}\ |X_i) - \frac{1}{n}\sum_{i=1}^{n} \tilde{h}_1^2(X_i)\ \Big|$$

$$\leq \Big|\frac{1}{n-2m+2} \sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-1\}}^{n} E(h^*_{i2\ldots 2m-1}\ |X_i)$$

$$- \frac{1}{n-2m+2}\sum_{i=1}^{n} \tilde{h}_1^2(X_i)\ \Big| + \frac{2m-2}{n(n-2m+2)}\sum_{i=1}^{n} \tilde{h}_1^2(X_i)$$

$$\leq \Big|\frac{1}{n-2m+2} \sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-1\}}^{n} E(h^*_{i2\ldots 2m-1}\ |X_i)$$

$$- \frac{1}{n-2m+2} \sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-1\}} \tilde{h}_1^2(X_i)\ \Big| + \frac{1}{n-2m+2}\sum_{i=2}^{2m-1} \tilde{h}_1^2(X_i)$$

$$+ \frac{2m-2}{n(n-2m+2)}\sum_{i=1}^{n} \tilde{h}_1^2(X_i)$$

$$= \frac{1}{n-2m+2}\Big| \sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-1\}} (-\tilde{h^*}(X_i)\ )\ (2\tilde{h}_1^{(m)}(X_i) + \tilde{h^*}(X_i))\ \Big|$$

$$+\frac{1}{n-2m+2}\sum_{i=2}^{2m-1}\tilde{h}_1^2(X_i)+\frac{2m-2}{n(n-2m+2)}\sum_{i=1}^{n}\tilde{h}_1^2(X_i)$$

$$\leq \frac{1}{n-2m+2}\Big(\sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-1\}}\tilde{h}_1^2(X_i)\Big)^{1/2}\Big(\sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-1\}}\tilde{h^*}^2(X_i)\Big)^{1/2}$$

$$+\frac{1}{n-2m+2}\sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-1\}}\tilde{h^*}^2(X_i)\;+\;\frac{1}{n-2m+2}\sum_{i=2}^{2m-1}\tilde{h}_1^2(X_i)$$

$$+\frac{2m-2}{n(n-2m+2)}\sum_{i=1}^{n}\tilde{h}_1^2(X_i). \tag{4.3.6}$$

It is easy to see that as $n\to\infty$, we have $\frac{1}{n-2m+2}\sum_{i=2}^{2m-1}\tilde{h}_1^2(X_i)=o_P(1)$. Also in view of (3.2.2), i.e., Raikov theorem in the present context, we have

$$\frac{2m-2}{n(n-2m+2)}\sum_{i=1}^{n}\tilde{h}_1^2(X_i)=o_P(1),\ \text{ as } n\to\infty.$$

Hence, in view of (4.3.6), in order to complete the proof of part (b), it suffices to show that as $n\to\infty$,

$$\frac{1}{n-2m+2}\sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-1\}}\tilde{h^*}^2(X_i)=o_P(1).$$

To prove the latter relation we first use Markov's inequality and conclude that

$$
\begin{aligned}
P \ (& \sum_{i \in \{1,\dots,n\}/\{2,\dots,2m-1\}} \tilde{h^*}^2(X_i) > \ \epsilon \ (n-2m+2) \ ) \\
\leq \ & \epsilon^{-\frac{1}{2}} \ (n-2m+2)^{-\frac{1}{2}} \sum_{i \in \{1,\dots,n\}/\{2,\dots,2m-1\}} E \ | \ \tilde{h^*}^2(X_i) \ |^{\frac{1}{2}} \\
\leq \ & \epsilon^{-\frac{1}{2}} \ (n-2m+2)^{\frac{1}{2}} \ E \ | \ \tilde{h^*}(X_1) \ | \\
\leq \ & \epsilon^{-\frac{1}{2}} \ (n-2m+2)^{\frac{1}{2}} \ n^{-\frac{1}{2}} \ n^{\frac{1}{2}} \ E\big(\ |h_{12\dots m}| \ \mathbf{1}_{(|h|>n^{\frac{3m}{5}})} \ \big) \\
\leq \ & \epsilon^{-\frac{1}{2}} \ (n-2m+2)^{\frac{1}{2}} \ n^{-\frac{1}{2}} \ E\big(\ |h_{12\dots m}|^{\frac{5}{6m}+1} \ \mathbf{1}_{(|h|>n^{\frac{3m}{5}})} \ \big)
\end{aligned}
$$

$\longrightarrow 0$, as $n \to \infty$.

The last relation is true since for $m \geq 2$, we have that $\dfrac{5}{6m} + 1 < \dfrac{5}{3}$, and this completes the proof of part (b) and those of Proposition 4.3.4 and Theorem 4.3.1. $\square$

4.4 Examples

Let $X_1, X_2, \dots,$ be a sequence of i.i.d. random variables with the density function

$$
f(x) = \begin{cases} |x-a|^{-3}, & |x-a| \geq 1, \ a \neq 0, \\ 0 \qquad\quad , & \text{elsewhere.} \end{cases}
$$

Consider the parameter $\theta = E^m(X_1) = a^m$, where $m \geq 1$ is a positive integer, and the kernel $h(X_1, \dots, X_m) = \prod_{i=1}^m X_i$. Then with m, n satisfying $n \geq m$, the corresponding U-statistic is

$$
U_n = \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \dots < i_m \leq n} \prod_{j=1}^{m} X_{i_j}.
$$

Simple calculation shows that $\tilde{h}_1(X_1) = X_1\ a^{m-1}\ -\ a^m$.

It is easy to check that $E|h(X_1, \ldots, X_m)|^{\frac{5}{3}} < \infty$ and that $\tilde{h}_1(X_1) \in DAN$ (cf. Gut, [25]).

For the pseudo-selfnormalized processes

$$U_{[nt]}^{nor} = \begin{cases} 0, & 0 \le t < \dfrac{m}{n}, \\ \dfrac{[nt]\left(\binom{[nt]}{m}^{-1} \sum_{1 \le i_1 < \ldots < i_m \le [nt]} \prod_{j=1}^{m} X_{i_j}\ -\ a^m\right)}{m\left(\sum_{i=1}^{n} (X_i a^{m-1}\ -\ a^m)^2\right)^{\frac{1}{2}}}, & \dfrac{m}{n} \le t \le 1 \end{cases}$$

and

$$\Psi_{[nt]}^{nor} = \begin{cases} 0, & 0 \le t < \dfrac{m}{n}, \\ \dfrac{n^{-m+1} \sum_{1 \le i_1 < \ldots < i_m \le [nt]} \left(\prod_{j=1}^{m} X_{i_j}\ -\ a^m\right)}{m\left(\sum_{i=1}^{n} (X_i a^{m-1}\ -\ a^m)^2\right)^{\frac{1}{2}}}, & \dfrac{m}{n} \le t \le 1. \end{cases}$$

Theorems 3.3.1 and 3.4.1, with $q(t) = 1$, conclude that

$$U_{[nt]}^{nor} \Rightarrow W(t) \quad \text{on } (D[0,1], \rho)$$

and

$$\Psi_{[nt]}^{nor} \Rightarrow t^{m-1}\ W(t) \quad \text{on } (D[0,1], \rho),$$

where ρ is the sup-norm for functions in $D[0,1]$ and $\{W(t), 0 \le t \le 1\}$ is a standard Wiener process.

The studentized U-process and U-type processes here are defined as follows.

$$U_{[nt]}^{stu} = \begin{cases} 0, & 0 \le t < \dfrac{m}{n}, \\ \dfrac{[nt]\left(\binom{[nt]}{m}^{-1} \sum_{1 \le i_1 < \ldots < i_m \le [nt]} \prod_{j=1}^{m} X_{i_j} - \theta\right)}{\sqrt{(n-1) \sum_{i=1}^{n} (U_{n-1}^{i} - U_n)^2}}, & \dfrac{m}{n} \le t \le 1 \end{cases}$$

and

$$\Psi^{stu}_{[nt]} = \begin{cases} 0, & 0 \le t < \dfrac{m}{n}, \\ \dfrac{n^{m-1} \sum_{1\le i_1<\ldots<i_m\le[nt]} \prod_{j=1}^{m} X_{i_j} - \theta)}{\sqrt{(n-1)\sum_{i=1}^{n}(U^i_{n-1} - U_n)^2}}, & \dfrac{m}{n} \le t \le 1. \end{cases}$$

where, by (4.3.3),

$$(n-1) \sum_{i=1}^{n} (U^i_{n-1} - U_n)^2$$

$$= \frac{m^2(n-1)}{(n-m)^2} \{ \sum_{i=1}^{n} X_i^2 [\binom{n-1}{m-1}^{-1} \sum_{\substack{1\le i_2<\ldots<i_m\le n \\ i_2,\ldots,i_m\neq i}} \prod_{j=2}^{m} X_{i_j}]^2 - n [\binom{n}{m}^{-1} \sum_{C(n,m)} \prod_{j=1}^{m} X_{i_j}]^2 \}.$$

In view of $U^{stu}_{[nt]}$ and $\Psi^{stu}_{[nt]}$, Our Theorems 4.2.1 and 4.2.2 are applicable for $U^{stu}_{[nt]}$ and $\Psi^{stu}_{[nt]}$ provided Theorem 4.2.5 continues hold true in this case. Hence, part (b) of Theorems 4.2.1 and 4.2.2, with $q(t) = 1$, imply that

$$U^{stu}_{[nt]} \Rightarrow W(t) \text{ on } (D[0,1], \rho)$$

and

$$\Psi^{stu}_{[nt]} \Rightarrow t^{m-1}\, W(t) \text{ on } (D[0,1], \rho),$$

where ρ is the sup-norm for functions in $D[0,1]$ and $\{W(t), 0 \le t \le 1\}$ is a standard Wiener process. Also applying Theorems 4.2.3 and 4.2.4, as $n \to \infty$, we have that

$$\sup_{0<t\le1} U^{stu}_{[nt]}/(t \log\log(\frac{1}{t})) \longrightarrow_d \sup_{0<t\le1} |W(t)|/(t\log\log(\frac{1}{t}))$$

and

$$\sup_{0<t\le1} \Psi^{stu}_{[nt]}/(t \log\log(\frac{1}{t})) \longrightarrow_d \sup_{0<t\le1} t^{m-1}|W(t)|/(t\log\log(\frac{1}{t})).$$

4.5 Appendix: Proof of Proposition 4.3.4

As it was mentioned before, the proof of this proposition can be done by modifying that of Theorem 4.3.1, except for that some of the steps are not required. This is due to the presence of the extra term of n with negative power i.e., n^{-j+1} in this proposition, where $j = 2, \ldots, m-1$, and $m \geq 3$. It is clear that among the statistics in Proposition 4.3.4 the one associated to $j = 2$ has the largest extra term of n^{-1}. Hence, we shall only show that as $n \to \infty$,

$$[n]^{-2m+1} \sum_{1 \leq i_1 \neq \ldots \neq i_{2m-2} \leq n} h(X_{i_1}, X_{i_2}, X_{i_3}, \ldots, X_{i_m}) \times h(X_{i_1}, X_{i_2}, X_{i_{m+1}}, \ldots, X_{i_{2m-2}}) = o_P(1). \tag{4.5.1}$$

To establish (4.5.1), consider the following setup:

$$h_{1\ldots m} := h(X_1, \ldots, X_m),$$
$$h^{(m)}_{1\ldots m} := h_{1\ldots m} \, \mathbf{1}_{(|h| \leq n^{\frac{3m}{5}})},$$
$$h^{**}_{12\ldots 2m-2} := h^{(m)}_{123\ldots m} \, h^{(m)}_{12\; m+1 \ldots 2m-2},$$
$$h^{(j)}_{1\ldots m} := h^{(m)}_{1\ldots m} \, \mathbf{1}_{(|h^{(m)}| \leq n^{\frac{3j}{5}})}, \; j = 1, \ldots, m-1,$$
$$h^{(0)}_{1\ldots m} := h^{(m)}_{1\ldots m} \, \mathbf{1}_{(|h^{(m)}| \leq \log(n))},$$

where $\mathbf{1}_A$ is the indicator function of the set A.

Now observe that as $n \to \infty$

$$P\Big(\sum_{1 \leq i_1 \neq \ldots \neq i_{2m-2} \leq n} h_{i_1 i_2 i_3 \ldots i_m} \, h_{i_1 i_2 i_{m+1} \ldots i_{2m-2}} \neq \sum_{1 \leq i_1 \neq \ldots \neq i_{2m-2} \leq n} h^{(m)}_{i_1 i_2 i_3 \ldots i_m} \, h^{(m)}_{i_1 i_2 i_{m+1} \ldots i_{2m-2}} \Big)$$

$$\leq n^m \, P\big(\, |h_{1\ldots m}| > n^{\frac{3m}{5}} \, \big)$$

$$\leq E\big[\, |h_{1\ldots m}|^{\frac{5}{3}} \, 1_{(|h_{1\ldots m}| > n^{\frac{3m}{5}})} \, \big] \longrightarrow 0.$$

In view of the latter asymptotic equivalency and our setup, in order to prove (4.5.1), we need to show that as $n \to \infty$,

$$[n]^{-2m+1} \sum_{1 \le i_1 \neq \ldots \neq i_{2m-2} \le n} h^{**}_{i_1, i_2, \ldots, i_{2m-2}} = o_P(1). \tag{4.5.2}$$

To prove (4.5.2), similarly to what we had in the proof of Theorem 4.3.1 we write

$$\sum_{1 \le i_1 \neq \ldots \neq i_{2m-2} \le n} h^{**}_{i_1 \ldots i_{2m-2}}$$

$$= \sum_{1 \le i_1 \neq \ldots \neq i_{2m-2} \le n} \Big\{ \sum_{d=1}^{2m-2} (-1)^{2m-2-d} \sum_{1 \le j_1 < \ldots < j_d \le 2m-2} E(h^{**}_{i_1 \ldots i_{2m-2}} - E(h^{**}_{i_1 \ldots i_{2m-2}}) |\, X_{i_{j_1}}, \ldots X_{i_{j_d}})$$

$$+ \sum_{c=1}^{2m-3} \sum_{1 \le k_1 < \ldots < k_c \le 2m-2} \sum_{d=1}^{c} (-1)^{c-d} \sum_{1 \le j_1 < \ldots < j_d \le c} E(h^{**}_{i_1 \ldots i_{2m-1}} - E(h^{**}_{i_1 \ldots i_{2m-2}}) |\, X_{i_{k_{j_1}}}, \ldots X_{i_{k_{j_d}}})$$

$$+ E\,(h^{**}_{i_1 \ldots i_{2m-2}}) \Big\}$$

$$:= \sum_{1 \le i_1 \neq \ldots \neq i_{2m-2} \le n} V^*(i_1, \ldots, i_{2m-2}) + \sum_{c=1}^{2m-3} \sum_{1 \le k_1 < \ldots < k_c \le 2m-2} \sum_{1 \le i_1 \neq \ldots \neq i_{2m-2} \le n} V^*(i_{k_1}, \ldots, i_{k_c})$$

$$+ \textstyle\sum_{1 \le i_1 \neq \ldots \neq i_{2m-2} \le n} E(h^{**}_{i_1 \ldots i_{2m-2}}).$$

In order to prove (4.5.2), we shall show the asymptotic negligibility of all of the above terms in the next three propositions.

Proposition 4.5.1. *If $E\,|h_{1 \ldots m}|^{\frac{5}{3}} < \infty$, then, as $n \to \infty$*

$$[n]^{-2m+1} \sum_{1 \le i_1 \neq \ldots \neq i_{2m-2} \le n} V^*(i_1, \ldots, i_{2m-2}) = o_P(1).$$

Proof of Proposition 4.5.1

For throughout use A will be a positive constant that may be different at each stage.

Since $V^*(i_1, \ldots, i_{2m-2})$ posses the property of complete degeneracy we can apply Lemma 1.8.1 after an application of Markov's inequality for the associated statistic and write, for $\epsilon > 0$,

$$P\ (\ |\ [n]^{-2m+2} \sum_{1 \leq i_1 \neq \ldots \neq i_{2m-2} \leq n} V^*(i_1, \ldots, i_{2m-2})| > \epsilon\ (n - 2m + 2)\ )$$

$$\leq \epsilon^{-2}\ \ (n - 2m + 2)^{-2}\ E\ [\ [n]^{-2m+2} \sum_{1 \leq i_1 \neq \ldots \neq i_{2m-1} \leq n} V^*(i_1, \ldots, i_{2m-2})\ ]^2$$

$$\leq \epsilon^{-2}\ \ (n - 2m + 2)^{-2}\ [n]^{-2m+2}\ E\ [\ V^*(1, \ldots, 2m-2)\ ]^2$$

$$\leq\ A\ \epsilon^{-2}\ (n - 2m + 2)^{-2}\ [n]^{-2m+2}\ n^{2m}\ n^{-2m}\ E\ [\ h^{(m)}_{12\ 3 \ldots m}\ h^{(m)}_{12\ m+1 \ldots 2m-2}\ ]^2$$

$$\leq\ A\ \epsilon^{-2}\ (n - 2m + 2)^{-2}\ [n]^{-2m+2}\ n^{2m}\ n^{-2m}\ n^{\frac{7m}{5}}\ E\ |\ h_{12 \ldots m}\ |^{\frac{5}{3}}$$

$$\longrightarrow 0,\ \text{ as } n \to \infty.$$

Proposition 4.5.2. . *If* $E\ |h_{1 \ldots m}|^{\frac{5}{3}} < \infty$, *then, as* $n \to \infty$,

$$[n]^{-2m+1} \sum_{1 \leq i_1 \neq \ldots \neq i_{2m-2} \leq n} V(i_{k_1}, \ldots, i_{k_c})\ = o_P(1),$$

where $c = 2, \ldots, 2m - 3$ *and* $1 \leq k_1 < \ldots < k_c \leq 2m - 2$.

Proof of Proposition 4.5.2

The proof will be stated in three cases according to the values of k_1 and k_2 as follows.

Case $k_1 = 1$ *and* $k_2 = 2$

Let s and t be respectively the number of elements of the sets

$$\{i_{k_1},\ldots,i_{k_c}\}\cap\{i_1,i_2,i_3,\ldots,i_m\} \text{ and } \{i_{k_1},\ldots,i_{k_c}\}\cap\{i_1,i_2,i_{m+1},\ldots,i_{2m-1}\}.$$

It is clear that in this case, i.e., $k_1 = 1$ and $k_2 = 2$, we have that $s,t \geq 2$ and $s+t = c+2$. Now define

$$V^{*^T}(i_{k_1},\ldots,i_{k_c}) = \sum_{d=1}^{c}(-1)^{c-d}\sum_{1\leq j_1<\ldots<j_d\leq c} E(h^{**^T}_{i_1\ldots i_{2m-2}} - E(h^{**^T}_{i_1\ldots i_{2m-2}}) \mid X_{i_{k_{j_1}}},\ldots,X_{i_{k_{j_d}}}),$$

$$V^{*T'}(i_{k_1},\ldots,i_{k_c}) = \sum_{d=1}^{c}(-1)^{c-d}\sum_{1\leq j_1<\ldots<j_d\leq c} E(h^{**^{T'}}_{i_1\ldots i_{2m-2}} - E(h^{**^{T'}}_{i_1\ldots i_{2m-2}}) \mid X_{i_{k_{j_1}}},\ldots,X_{i_{k_{j_d}}}),$$

where $h^{**^T}_{i_1\ldots i_{2m-2}} = h^{(s)}_{i_1i_2\ i_3\ldots i_m}\ h^{(m)}_{i_1i_2\ i_{m+1}\ldots i_{2m-2}}$ and $h^{**^{T'}}_{i_1\ldots i_{2m-2}} = h^{(s)}_{i_1i_2\ i_3\ldots i_m}\ h^{(t)}_{i_1i_2\ i_{m+1}\ldots i_{2m-2}}$.

Now observe that as $n\to\infty$

$$\begin{aligned}
& P\ \Big(\sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^*(i_{k_1},\ldots,i_{k_c}) \neq \sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*^{T'}}(i_{k_1},\ldots,i_{k_c})\Big) \\
\leq\ & P\Big(\sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^*(i_{k_1},\ldots,i_{k_c}) \neq \sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*^T}(i_{k_1},\ldots,i_{k_c})\Big) \\
+\ & P\Big(\sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*^T}(i_{k_1},\ldots,i_{k_c}) \neq \sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*^{T'}}(i_{k_1},\ldots,i_{k_c})\Big) \\
\leq\ & n^s\ P\ (\ |h^{(m)}_{12\ 3\ldots m}| > n^{\frac{3s}{5}}\ ) + n^t\ P\ (\ |h^{(m)}_{12\ m+1\ldots 2m-1}| > n^{\frac{3t}{5}}\ ) \\
\leq\ & E\ [\ |h_{12\ 3\ldots m}|^{\frac{5}{3}}\ \mathbf{1}_{(|h|>n^{\frac{3s}{5}})}\ ] + E\ [\ |h_{12\ m+1\ldots 2m-2}|^{\frac{5}{3}}\ \mathbf{1}_{(|h|>n^{\frac{3t}{5}})}\ ] \longrightarrow 0.
\end{aligned}$$

The latter relation suggests that

$$\sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^*(i_{k_1},\ldots,i_{k_c}) \text{ and } \sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*^{T'}}(i_{k_1},\ldots,i_{k_c})$$

are asymptotically equivalent in probability.

Since $V^{*^{T'}}(i_{k_1}, \ldots, i_{k_c})$ is complete degenerate, Markov's inequality followed by an application of Lemma 1.8.2 yields,

$$P\ \Big(\ [n]^{-2m+2} \sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} \left|V^{*^{T'}}(i_{k_1}, \ldots, i_{k_c})\right| > \epsilon\ (n-2m+2)\Big)$$

$$\leq\ \epsilon^{-2}\ (n-2m+2)^{-2}\ E\ [\ [n]^{-2m+2} \sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*^{T'}}(i_{k_1}, \ldots, i_{k_c})\ ]^2$$

$$\leq\ A\epsilon^{-2}\ (n-2m+2)^{-2}\ [n-(2m-2-c)]^{-c}\ E\ [\ h^{(s)}_{123\ldots m}\ h^{(t)}_{12\ m+1\ldots 2m-1}\ ]^2$$

$$\leq\ A\epsilon^{-2}\ (n-2m+2)^{-2}\ [n-(2m-1-c)]^{-c}\ n^{c+2}\ n^{-c-2}\ n^{\frac{7(t+s)}{10}}\ E\ |\ h_{1\ldots m}\ |^{\frac{5}{3}}$$

$$\longrightarrow 0,\ \text{ as } n\rightarrow\infty.$$

This completes the proof of Proposition 4.5.2 when $k_1 = 1$ and $k_2 = 2$.

Case either $k_1 \neq 1$ or $k_2 \neq 2$

Let s, t be as were defined in the previous case and note that here we have that $s, t \geq 1$ and $s + t = c + 1$. The proof of Proposition 4.5.2 in this case results from a similar argument to what was given for the previous case, hence the details are omitted.

Case $k_1 \neq 1, k_2 \neq 2$

Let s, t be as what were defined in the previous two cases and note that in this case we have $s, t \geq 0$ and $s + t = c$. Also let V^{*^T} and $V^{*^{T'}}$ as they were defined in the case $k_1 = 1, k_2 = 2$ and observe that as $n \rightarrow \infty$ we have

$$P\ (\sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*}(i_{k_1}, \ldots, i_{k_c})\ \neq \sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*^{T'}}(i_{k_1}, \ldots, i_{k_c})\ )$$

$$\leq P\ (\sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*}(i_{k_1}, \ldots, i_{k_c})\ \neq \sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*^{T}}(i_{k_1}, \ldots, i_{k_c})\ )$$

$$+ P\Big(\sum_{1\le i_1\ne\ldots\ne i_{2m-2}\le n} V^{*^T}(i_{k_1},\ldots,i_{k_c}) \ne \sum_{1\le i_1\ne\ldots\ne i_{2m-2}\le n} V^{*^{T'}}(i_{k_1},\ldots,i_{k_c}) \Big)$$

$$\le \begin{cases} n^s P(\ |h^{(m)}_{12\ 3\ldots m}| > n^{\frac{3s}{5}}\ ) + n^t P(\ |h^{(m)}_{12\ m+1\ldots 2m-2}| > n^{\frac{3t}{5}}\ ), & \text{s,t} > 0,\ \text{s+t=c} \\ n^c P(\ |h^{(m)}_{12\ 3\ldots m}| > n^{\frac{3c}{5}}\ ) + P(\ |h^{(m)}_{12\ m+1\ldots 2m-2}| > \log(n)\ ), & \text{s=c,t=0} \end{cases}$$

$$\le \begin{cases} E[\ |h_{12\ 3\ldots m}|^{\frac{5}{3}}\ \mathbf{1}_{(|h|>n^{\frac{3s}{5}})}\ ] + E[\ |h_{12\ m+1\ldots 2m-2}|^{\frac{5}{3}}\ \mathbf{1}_{(|h|>n^{\frac{3t}{5}})}\ ], & \text{s,t} > 0,\ \text{s+t=c} \\ E[\ |h_{12\ldots m}|^{\frac{5}{3}}\ \mathbf{1}_{(|h|>n^{\frac{3c}{5}})}\ ] + P(\ |h^{(m)}_{1m+1\ldots 2m-2}| > \log(n)\ ), & \text{s=c,t=0} \end{cases}$$

$$\longrightarrow 0.$$

Applying Markov's inequality followed by an application Lemma 1.8.1 once again yields

$$P(\big|[n]^{-2m+2} \sum_{1\le i_1\ne\ldots\ne i_{2m-2}\le} V^{*T'}(i_{k_1},\ldots,i_{k_c})\big| > \epsilon n(n-2m+2))$$

$$\le A\epsilon^{-2}\,(n-2m+2)^{-2}\,[n-(2m-2-c)]^{-c}\,n^{c+2}\,n^{-c-2}\,E\,[\,h^{(s)}_{123\ldots m}\,h^{(t)}_{12\ m+1\ldots 2m-2}\,]^2$$

$$\le \begin{cases} A\epsilon^{-2}\dfrac{(n-2m+2)^{-2}}{[n-(2m-2-c)]^c}\ n^{c+2}\ n^{-c-2}\ n^{\frac{7c}{10}}\ E|h_{12\ldots m}|^{\frac{5}{3}}, & \text{s,t} > 0,\ \text{s+t=c} \\ A\epsilon^{-2}\dfrac{(n-2m+2)^{-2}}{[n-(2m-2-c)]^c}\ n^{c+2}\ n^{-c-2}\ n^{\frac{7c}{10}}\ \log^{\frac{7}{6}}(n)\ E|h_{12\ldots m}|^{\frac{5}{3}}, & \text{s=c,t=0} \end{cases}$$

$$\longrightarrow 0,\ \text{as } n\to\infty.$$

This completes the proof of Proposition 4.5.2. □

As the last step of the proof of Proposition 4.3.4, in the next result we deal with terms of the form of sums of i.i.d. random variables (cf. Remark 4.3.1 of this chapter).

Proposition 4.5.3. . *If* $E|h_{1\ldots m}|^{\frac{5}{3}} < \infty$, *then, as* $n\to\infty$

(a) $[n]^{-2m+1}\sum_{1\leq\neq i_1\neq\ldots\neq i_{2m-2}\leq n} V^*(i_{k_1}) = o_P(1),\ \ k_1\in\{3,\ldots,2m-2\}$,

$$(b)\ \frac{1}{(n-2m+2)(n-2m+3)}\sum_{i\in\{1,\ldots,n\}/\{1,3,\ldots,2m-2\}}^{n} E(h^{**}_{1i3\ldots 2m-2}-E(h^{**}_{1i3\ldots 2m-2})\ |X_i)$$

$$=o_P(1),$$

$$(c)\frac{1}{(n-2m+2)(n-2m+3)}\sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-2\}}^{n} E(h^{**}_{i2\ldots 2m-2}-E(h^{**}_{i2\ldots 2m-2})\ |X_i)$$

$$+\ \frac{1}{n-2m+2}\ E(h^{**}_{12\ldots 2m-2})\ \ =o_P(1).$$

Proof of Proposition 4.5.3

First we give the proof of part (a). Due to similarities, we shall state the proof only for the case that $k_1=3$.

Define

$$V^{*^{T}}(i_3) =\ E(h^{**^{T}}_{i_1i_2\ldots i_{2m-2}}-E(h^{**^{T}}_{i_1i_2\ldots i_{2m-2}})|\ X_{i_3}),$$

$$V^{*^{T'}}(i_3) =\ E(h^{**^{T'}}_{i_1i_2\ldots i_{2m-2}}-E(h^{**^{T'}}_{i_1i_2\ldots i_{2m-2}})|\ X_{i_3}),$$

where $h^{**^{T}}_{i_1i_2\ldots i_{2m-2}}=h^{(1)}_{i_1i_2i_3\ldots i_m}\,h^{(m)}_{i_1i_2i_{m+1}\ldots i_{2m-2}}$ and $h^{**^{T'}}_{i_1i_2\ldots i_{2m-2}}=h^{(1)}_{i_1i_2i_3\ldots i_m}\,h^{(0)}_{i_1i_2i_{m+1}\ldots i_{2m-2}}$. Again observe that as $n\to\infty$

$$P(\sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^*(i_3)\neq \sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*^{T'}}(i_3)\ )$$

$$\leq\ P(\sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^*(i_3)\neq \sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*^{T}}(i_3)\ )$$

$$+\ P(\sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*^{T}}(i_3)\neq \sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*^{T'}}(i_3)\ )$$

$$\leq\ n\ P(|h^{(m)}_{123\ldots m}|>n^{3/5})+\ P(|h^{(m)}_{12\ m+1\ldots 2m-2}|>\log(n))$$

$\longrightarrow 0.$

Applying Markov's inequality we arrive at

$$P(\ |\frac{1}{n-2m+3}\sum_{1\leq i_1\neq\ldots\neq i_{2m-2}\leq n} V^{*^{T'}}(i_3)\ |>\ \epsilon\ (n-2m+2)\ )$$

$$\leq\ A\ \epsilon^{-2}\ (n-2m+2)^{-2}\ (n-2m+3)^{-1}\ E(h^{(1)}_{123\ldots m}h^{(0)}_{12\ m+1\ldots 2m-2})^2$$

$\longrightarrow 0,$ as $n \to \infty$.

This completes the proof of part (a).

Next to prove part (b) define

$h^{**^T} = h^{(1)}_{123\ldots m}h^{(m)}_{12\ m+1\ldots 2m-2},$

$h^{**^{T'}} = h^{(1)}_{123\ldots m}h^{(1)}_{12\ m+1\ldots 2m-2},$

and observe that as $n \to \infty$ we have

$$P(\sum_{i\in\{1,\ldots,n\}/\{1,3,\ldots,2m-2\}}^{n} E(h^{**}_{1i3\ldots 2m-2} - E(h^{**}_{1i3\ldots 2m-2})\ |X_i)$$

$$\neq \sum_{i\in\{1,\ldots,n\}/\{1,3,\ldots,2m-2\}}^{n} E(h^{**^{T'}}_{1i3\ldots 2m-2} - E(h^{**^{T'}}_{1i3\ldots 2m-2})\ |X_i)\ )$$

$$\leq P(\sum_{i\in\{1,\ldots,n\}/\{1,3,\ldots,2m-2\}}^{n} E(h^{**}_{1i3\ldots 2m-2} - E(h^{**}_{1i3\ldots 2m-2})\ |X_i)$$

$$\neq \sum_{i\in\{1,\ldots,n\}/\{1,3,\ldots,2m-2\}}^{n} E(h^{**^{T}}_{1i3\ldots 2m-2} - E(h^{**^{T}}_{1i3\ldots 2m-2})\ |X_i)\ )$$

$$+ P(\sum_{i\in\{1,\ldots,n\}/\{1,3,\ldots,2m-2\}}^{n} E(h^{**^{T}}_{1i3\ldots 2m-2} - E(h^{**^{T}}_{1i3\ldots 2m-2})\ |X_i)$$

$$\neq \sum_{i\in\{1,\ldots,n\}/\{1,3,\ldots,2m-2\}}^{n} E(h^{**^{T'}}_{1i3\ldots 2m-2} - E(h^{**^{T'}}_{1i3\ldots 2m-2})\ |X_i)\ )$$

$$\leq n\ P(|h^{(m)}_{123\ldots m}| > n^{3/5}) + \ n\ P(|h^{(m)}_{12\ m+1\ldots 2m-2}| > n^{3/5})$$

$\longrightarrow 0.$

Hence another application of Markov's inequality yields,

$$P(\ |\frac{1}{n-2m+3} \sum_{i\in\{1,\ldots,n\}/\{1,3,\ldots,2m-2\}}^{n} E(h^{**T'}_{1i3\ldots 2m-2} - E(h^{**T'}_{1i3\ldots 2m-2})\ |X_i)| > \epsilon\ (n-2m+2))$$

$$\leq\ A\ \epsilon^{-2}\ (n-2m+2)^{-2}\ (n-2m+3)^{-1}\ E(h^{(1)}_{123\ldots m} h^{(1)}_{12\ m+1\ldots 2m-2})^2$$

$\longrightarrow 0,$ as $n \to \infty$.

Now the proof of part (b) is complete.

To prove part (c) we only need to observe that

$$\frac{1}{(n-2m+2)(n-2m+3)} \sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-2\}}^{n} E(h^{**}_{i2\ldots 2m-2} - E(h^{**}_{i2\ldots 2m-2})\ |X_i)$$

$$+\frac{1}{n-2m+2}\ E(h^{**}_{12\ldots 2m-2})$$

$$= \frac{1}{(n-2m+2)(n-2m+3)} \sum_{i\in\{1,\ldots,n\}/\{2,\ldots,2m-2\}}^{n} E(h^{**}_{i2\ldots 2m-2}\ |X_i).$$

The rest of the proof is similar to that of part (b), hence the details are omitted. Now the proof of Proposition 4.5.3 and those of (4.5.2), (4.5.1) and Proposition 4.3.4 is complete. □

Bibliography

[1] Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Annals of Mathematical Statatistics.*, **23** 193-212.

[2] Arvesen, J. N. (1969). Jackknifing U-statistics. *Annals of Mathematical Statistics.* **40** , 2076-2100

[3] Billingsley, P. (1971). *Weak convergence of measures: Applications in Probability.* Siam.

[4] Borovskikh, Y. V. (1996). *U-statistics in Banach Spaces.* VSP, Utrecht.

[5] Chibisov, D. (1964). Some theorems on the limiting behaviour of empirical distribution functions. *Selected Translations in Mathematical Statistics and Probability.*, **6** 147-156.

[6] Chow, Y. S. (1960). A Martingale Inequality and the Law of Large Numbers. *Proceedings of the American Mathematical Society*, Vol. 11, **1**, 107-111.

[7] Csörgő, M. (2004). A glimpse of the impact of Pál Erdős on probability and statitistics. *The Canadian Journal of Statistics* **4**, 493-556.

[8] Csörgő, M., Csörgő, S., Horváth, L. and Mason D., (1986). Weighted empirical and of quantile processes. *The Annals of Probability* **14**, 31-85.

[9] Csörgő, M., Horváth, L. (1986). Approximations of weighted empirical and quantile processes. *Statistics and Probability Letters*, **4** 275-280.

[10] Csörgő, M., Horváth, L. (1988). Invariance principles for changepoint problems. *Journal of Multivariate Analysis* **1**, 151-168.

[11] Csörgő, M., Horváth, L. (1988). Nonparametric methods for changepoint problems. *Handbook of Statistics*, Vol. **7** 403-425, Elsevier Science Publishers B.V. (North-Holland).

[12] Csörgő, M., Horváth, L. (1993). *Weighted approximations in probability and statistics.* John Wiley and Sons, Ltd., Chichester.

[13] Csörgő, M., Horváth, L. (1997). *Limit Theorems in Change-point Analysis.* Wiley, Chichester.

[14] Csörgő, M., Szyszkowicz, B. and Wang, Q. (2003). Donsker's theorem for self-normalized parial sums processes. *The Annals of Probability* **31**, 1228-1240.

[15] Csörgő, M., Szyszkowicz, B. and Wang, Q. (2004). On Weighted Approximations and Strong Limit Theorems for Self-normalized Partial Sums Processes. In *Asymptotic methods in Stochastics*, 489-521, Fields Inst. Commun.44, Amer. Math. Soc., Providence RI.

[16] Csörgő, M., Szyszkowicz, B. and Wang, Q. (2008). On weighted approximations in $D[0,1]$ with application to self-normalized partial sum processes. *Acta Mathematica Hungarica* **121 (4)**, 307-332.

[17] Csörgő, M., Szyszkowicz, B. and Wang, Q. (2008). Asymptotics of studentized U-type processes for changepoint problems. *Acta Mathematica Hungarica* **121 (4)**, 333-357.

[18] Donsker, M. (1951). An invariance principle for certain probability limit theorems. *Memoirs of American Mathematical Society* **6** (four papers on probability).

[19] Donsker, M. (1952). Justification and extenstion of Doob's heuristic approach to the Kolmororov-Smirnov theorems. *The Annals of Mathematical Statistics* **23** 277-281.

[20] Doob, J. L. (1990). *Stochastic processes.* John Wiley and Sons, Inc., New York.

[21] Erdős, P. and Kac, M. (1946). On certain limit theorems of the theory of probability. *Bulltien of the American Mathematical Society.* **52**, 292-302.

[22] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. 2. Wiley.

[23] Giné, E. and Zinn, J. (1992). Marcinkiewicz type laws of large numbers and convergence of moments for U-statistics. In *Probability in Banach Spaces* (R. Dudley, M. Hahn and J. Kuelbs, eds) **8**, 273-291, Birkhauser, Boston.

[24] Giné, E., Götze, F. and Mason D. M. (1997). When is the student t-statistic asymptotically Normal? *The Annals of Probability* **25**, 1514-1531.

[25] Gut, A. (2005). *Probability: A Graduate Course.* Springer.

[26] Hall, P. and Hyde, C. C. (1980). *Martingale limit theory and its application.* Academic Press, Inc.

[27] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statatistics.* **19**, 293-325.

[28] Kac, M. (1946). On the average of a certain Wiener functional and a related limit theorem in calculus of probability. *In Selected Translations in Mathematical Statistics and Probability.* **59** 401-414.

[29] Loève, M. (1977). *Probability Theory I.* Springer, Berlin.

[30] Miller, R. G. Jr. and Sen, P. K. (1972). Weak convergence of U-statistics and Von Mises' differentiable statistical functions. *Annals of Mathematical Statistics* **43**, 31-41.

[31] Nasari, M. M. (2009). On weak approximations of U-statistics. *Statistics and Probability Letters* **79**, 1528-1535.

[32] Nasari M. M. (2009). Studentized processes of U-statistics. In *Technical Reports in LRSP.* No. **445** Carleton U.-U. of Ottawa.

[33] O'Reilly, N. (1974). On weak convergence of empirical processes in sup-norm metric. *The Annals of Probability,* **2** 642-651.

[34] Rényi, A. (1953). On the theory of order statistics. *Acta Mathematica Academiae Scientiarum Hungarica.,* **4** 191-232.

[35] Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

[36] Szyszkowicz, B. (1992). Weak convergence of stochastic processes in weighted metrics and their applications to contiguous changepoint analysis. Doctoral dissertion, Carleton University, Ottawa.

[37] Szyszkowicz, B. (1996). Weighted approximations of partial sum processes in $D[0, \infty]$, II. *Studia Scientiarum Mathematicarum Hungarica*, **31**, 323-353.

[38] Szyszkowicz, B. (1997). Weighted approximations of partial sum processes in $D[0, \infty]$, II. *Studia Scientiarum Mathematicarum Hungarica*, **33**, 305-320.

[39] Wilks S. S. (1962). *Mathematical Statistics.* John Wiley and Sons, Inc.