AN EDGEWORTH EXPANSION FOR SYMMETRIC FINITE POPULATION STATISTICS

M. Bloznelis[1], F. Götze

Vilnius University and the Institute of
Mathematics and Informatics, Bielefeld University

2000 (Revised 2001)

Abstract. Let T be a symmetric statistic based on sample of size n drawn without replacement from a finite population of size N, where $N > n$. Assuming that the linear part of Hoeffding's decomposition of T is nondegenerate we construct a one term Edgeworth expansion for the distribution function of T and prove the validity of the expansion with the remainder $O(1/n^*)$ as $n^* \to \infty$, where $n^* = \min\{n, N-n\}$.

1. Introduction and results

1.1. Introduction. Given a set $\mathcal{X} = \{x_1, \dots, x_N\}$, let $(X_1, \dots, X_N)$ be a random permutation of the ordered set $(x_1, \dots, x_N)$. We assume that the random permutation is uniformly distributed over the class of permutations. Let

$$T = t(X_1, \dots, X_n)$$

denote a symmetric statistic of the first n observations $X_1, \dots, X_n$, where $n < N$. That is, t is a real function defined on the class of subsets $\{x_{i_1}, \dots, x_{i_n}\} \subset \mathcal{X}$ of size n and we assume that $t(x_{i_1}, \dots, x_{i_n})$ is invariant under permutations of its arguments. Since $X_1, \dots, X_n$ represents a sample drawn without replacement from the population $\mathcal{X}$, we call T a symmetric finite population statistic.

We shall consider symmetric finite population statistics which are asymptotically normal when n^* and N tend to ∞, where $n^* = \min\{n, N-n\}$. In the simplest case of linear statistics the asymptotic normality was established by Erdős and Rényi

[1]Research supported by the A.v. Humboldt Foundation

1991 *Mathematics Subject Classification.* Primary 62E20; secondary 60F05.

Key words and phrases. Edgeworth expansion, finite population, Hoeffding decomposition, sampling without replacement.

(1959) under fairly general conditions. The rate in the Erdős-Rényi central limit theorem was studied by Bikelis (1972). Höglund (1978) proved the Berry–Esseen bound. An Edgeworth expansion was established by Robinson (1978), see also Bickel and van Zwet (1978), Schneller (1989), Babu and Bai (1996).

Asymptotic normality of nonlinear finite population statistics was studied by Nandi and Sen (1963), who proved a central limit theorem for U–statistics. The accuracy of the normal approximation of U–statistics was studied by Zhao and Chen (1987, 1990), Kokic and Weber (1990). A general Berry–Esseen bound for combinatorial multivariate sampling statistics (including finite population U–statistics) was established by Bolthausen and Götze (1993). Rao and Zhao (1994), Bloznelis (1999) constructed Berry-Esseen bounds for Student's t statistic. One term asymptotic expansions of nonlinear statistics, which can be approximated by smooth functions of (multivariate) sample means have been shown by Babu and Singh (1985), see also Babu and Bai (1996). For U–statistics of degree two one term Edgeworth expansions were constructed by Kokic and Weber (1990). Bloznelis and Götze (1999, 2000) established the validity of one term Edgeworth expansion for U–statistics of degree two with remainders $o(1/\sqrt{n^*})$ and $O(1/n^*)$.

A second order asymptotic theory for general asymptotically normal symmetric statistics of *independent and identically distributed* observations was developed in a recent paper by Bentkus, Götze and van Zwet (1997), which concludes a number of previous investigations of particular statistics: Bickel (1974), Callaert and Janssen (1978), Götze (1979), Callaert, Janssen and Veraverbeke (1980), Serfling (1980), Helmers (1982), Helmers and van Zwet (1982), van Zwet (1984), Bickel, Götze and van Zwet (1986), Lai and Wang (1993), etc. This theory is based on the representation of symmetric statistics by sums of U–statistics of increasing order via Hoeffding's decomposition. Another approach, see, e.g., Chibisov (1972), Pfanzagl (1973), Bhattacharya and Ghosh (1978), which is based on Taylor expansions of statistics in powers of the underlying i.i.d. observations, focuses on smooth functions of observations.

In view of important classes of applications (jackknife histogram, see, Wu (1990), Shao (1989), Booth and Hall (1993) and subsampling, see, Politis and Romano (1994), Bertail (1997), Bickel, Götze and van Zwet (1997)) we want to develop in this paper a second order asymptotic theory similar to that of Bentkus, Götze and van Zwet (1997) for simple random samples drawn without replacement from finite populations.

The starting point of our asymptotic analysis is the Hoeffding decomposition

$$T = \mathbf{E}T + \sum_{1\le i\le n} g_1(X_i) + \sum_{1\le i<j\le n} g_2(X_i, X_j) + \dots \tag{1.1}$$

which expands T into the sum of mutually uncorrelated U-statistics of increasing order

$$U_k = \sum_{1\le i_1<\cdots<i_k\le n} g_k(X_{i_1},\dots,X_{i_k}), \qquad k=1,\dots,n. \tag{1.2}$$

Here the symmetric kernels g_k, $k=1,\dots,n$, are centered, $\mathbf{E}g_k(X_1,\dots,X_k)=0$, and satisfy the orthogonality condition

$$\mathbf{E}(g_k(X_1,\dots,X_k)|X_1,\dots,X_{k-1})=0 \quad \text{almost surely.} \tag{1.3}$$

It follows from (1.3) that $U_1,\dots,U_n$ are orthogonal in L_2 (i.e., $\mathbf{E}U_kU_r=0$, for $k\neq r$). Furthermore, the condition (1.3) ensures the uniqueness of the decomposition 1.1) in the following sense: given another decomposition like 1.1) with symmetric kernels, say g_k', satisfying (1.3), we always have $g_k=g_k'$.

Let us mention briefly that given k, the function $g_k(x_{i_1},\dots,x_{i_k})$ can be expressed by a linear combination of conditional expectations $\mathbf{E}(T-\mathbf{E}T|X_1=x_{r_1},\dots,X_j=x_{r_j})$ where $j=1,\dots,k$ and $\{x_{r_1},\dots,x_{r_j}\}\subset\{x_{i_1},\dots,x_{i_k}\}$. The expressions for g_1 and g_2 are provided by (1.5) below. For larger $k=3,\dots,n$, the expressions are more complex and we refer to Bloznelis and Götze (2001) where a general formula for g_k is derived.

We shall assume that the linear part $U_1=\sum g_1(X_i)$ is nondegenerate. That is, $\sigma^2>0$, where $\sigma^2=\mathbf{Var}g_1(X_1)$. In the case where, for large n^*, the linear part dominates the statistic we can approximate the distribution of T by a normal distribution, using the central limit theorem. Furthermore, the sum of the linear and quadratic term,

$$U=\mathbf{E}T+\sum_{1\le i\le n} g_1(X_i)+\sum_{1\le i<j\le n} g_2(X_i,X_j)$$

typically provides a sufficiently precise approximation to T so that one term Edgeworth expansions for the distribution functions of T and U are, in fact, the same. Therefore, in order to construct a one term Edgeworth expansion of T it suffices to find such expansion for U. In particular, we do not need to evaluate all the summands of the decomposition 1.1), but (moments of) the first two terms only, cf. (1.4) below. Similarly, the two term Edgeworth expansion for the distribution function of T could be constructed using the approximation $T\approx\mathbf{E}T+U_1+U_2+U_3$, etc. An advantage of such an approach is that it provides (at least formal) Edgeworth expansion for an *arbitrary* symmetric finite population statistic T no matter whether it is a smooth function of observations or not. In the present paper we construct the one term Edgeworth expansion for the distribution function of T and prove the validity of the expansion with the remainder $O(1/n^*)$.

A simple calculation shows that the variance of the linear part satisfies

$$\mathbf{Var} \sum_{1\le i\le n} g_1(X_i) = \sigma^2\tau^2 \frac{N}{N-1}, \qquad \tau^2 = Npq, \qquad p = n/N, \qquad q = 1-p.$$

Note that $n^*/2 \le \tau^2 \le n^*$. We approximate the distribution function

$$F(x) = \mathbf{P}\{T \le \mathbf{E}T + \sigma\tau x\},$$

by the one term Edgeworth expansion

$$G(x) = \Phi(x) - \frac{(q-p)\,\alpha + 3\,\kappa}{6\,\tau}\,\Phi^{(3)}(x) \tag{1.4}$$

and provide an explicit bound for the remainder

$$\Delta = \sup_{x\in R} |F(x) - G(x)|,$$

where

$$\alpha = \sigma^{-3}\mathbf{E}g_1^3(X_1) \qquad \text{and} \qquad \kappa = \sigma^{-3}\tau^2\mathbf{E}g_2(X_1,X_2)g_1(X_1)g_1(X_2)$$

and where

$$g_1(X_i) = \frac{N-1}{N-n}\,\mathbf{E}(T' \mid X_i), \qquad T' = T - \mathbf{E}T, \tag{1.5}$$

$$g_2(X_i,X_j) = \frac{N-2}{N-n}\,\frac{N-3}{N-n-1}\,\Big(\mathbf{E}(T' \mid X_i, X_j) - \frac{N-1}{N-2}\,(\mathbf{E}(T'|X_i) + \mathbf{E}(T'|X_j))\Big).$$

Furthermore, Φ denotes the standard normal distribution function, and $\Phi^{(3)}$ denotes the third derivative of Φ.

Before to formulate our main results, Theorems 1.1 and 1.2, we introduce the smoothness conditions, which together with the moment conditions, ensure the validity of the expansion (1.4).

1.2. Smoothness conditions. Given a general symmetric statistic T we approximate it by a U-statistic via Hoeffding's decomposition. In order to control the accuracy of such an approximation we use moments of finite differences of T. Introduce the difference operation

$$D^jT = t(X_1,\dots,X_j,\dots,X_n) - t(X_1,\dots,X_j',\dots,X_n), \quad X_j' = X_{n+j},$$

where X_j is replaced by X_j' in the second summand, for $j \le n^*$. Higher order difference operations are defined recursively:

$$D^{j_1,j_2}T = D^{j_2}\big(D^{j_1}T\big), \qquad D^{j_1,j_2,j_3}T = D^{j_3}\big(D^{j_2}(D^{j_1}T)\big), \dots .$$

It is easy to see that the difference operations are symmetric, i.e., $D^{j_1,j_2}T = D^{j_2,j_1}T$, etc. Given $k < n^*$ write

$$\delta_j = \delta_j(T) = \mathbf{E}\big(\tau^{2(j-1)}\mathbb{D}_jT\big)^2, \qquad \mathbb{D}_jT = D^{1,2,\dots,j}T, \qquad 1 \le j \le k.$$

Bounds for the accuracy of the approximation of T by the sum of the first few terms of the decomposition 1.1) are provided by the following theorem.

Theorem A. *(Bloznelis and Götze (2001)) For $1 \le k < n^*$, we have*

$$T = \mathbf{E}T + U_1 + \cdots + U_k + R_k, \qquad \textit{with} \qquad \mathbf{E}R_k^2 \le \tau^{-2(k-1)}\delta_{k+1}. \tag{1.6}$$

In typical situations (U-statistics, smooth functions of sample means, Student's t and many others) for a properly standardized statistic T we have $U_j = O_P\big(\tau^{1-j}\big)$, for $j = 1, \dots, k$, and

$$\delta_{k+1} = O(\tau^{-2}) \qquad \text{as} \qquad n^*, N \to \infty, \tag{1.7}$$

for some k. Note that (1.7) can be viewed as a smoothness condition. For instance, given a statistic which is a function of the sample mean this condition is satisfied in the case where the function (defining the statistic) is $k+1$ times differentiable, see Bloznelis and Götze (2001).

Assuming that (1.7) holds for $k = 2$ we obtain from Theorem A that $T = U + O_P(\tau^{-2})$ thus, showing that up to an error $O(\tau^{-2})$ the statistics T and U are asymptotically equivalent. Finally, we remark that (1.7) holds if δ_{k+1}/σ^2 and the variance of the linear part $\mathbf{Var}U_1$ remain bounded as $n^*, N \to \infty$.

Another smoothness condition we are going to use is a Cramér type condition. Recall Cramér's (C) condition for the distribution F_Z of a random variable Z,

$$\sup_{|t|>\delta} |\mathbf{E}\exp\{itZ\}| < 1, \qquad \text{for some} \qquad \delta > 0. \tag{C}$$

In the classical theory of sums of independent random variables this condition together with moment conditions ensures the validity of Edgeworth expansions with remainders $O(n^{-k/2})$ and $o(n^{-k/2})$, $k = 2, 3, \dots$, for the distribution function of the sum of n independent observations from the distribution F_Z, see Petrov (1975).

In our situation the condition like (C) is too stringent. We shall use a modification of (C) which is applicable to random variables assuming a finite number of values only. For the summand of the linear part $Z = \sigma^{-1}g_1(X_1)$, we assume that $\rho > 0$, where

$$\rho := 1 - \sup\{|\mathbf{E}\exp\{itZ\}| : b_1/\beta_3 \le |t| \le \tau\}.$$

Here b_1 is a small absolute constant (one may choose, e.g., $b_1 = 0.001$) and $\beta_3 = \sigma^{-3}\mathbf{E}|g_1(X_1)|^3$. Other modifications of Cramér's (C) condition which are applicable to discrete random variables were considered by Albers, Bickel and van Zwet (1976), Robinson (1978) and Bloznelis and Götze (2000), where in the latter paper relations between various conditions are discussed.

1.3. Results. Write $\zeta = \sigma^{-2}\tau^8\mathbf{E}g_3^2(X_1, X_2, X_3)$ and denote

$$\beta_k = \sigma^{-k}\mathbf{E}|g_1(X_1)|^k, \qquad \gamma_k = \sigma^{-k}\tau^{2k}\mathbf{E}|g_2(X_1, X_2)|^k, \qquad k = 2,\ 3,\ 4.$$

Theorem 1.1. *There exists an absolute constant $c > 0$ such that*

$$\Delta \le \frac{c}{\tau^2}\,\frac{\beta_4+\gamma_4+\zeta}{\rho^2} + \frac{c}{\tau^2}\,\frac{\delta_4}{\sigma^2\rho^2}\,. \tag{1.8}$$

For U–statistics of arbitrary but fixed degree k

$$\sum_{1\le i_1<\dots<i_k\le n} h(X_{i_1},\dots,X_{i_k}), \tag{1.9}$$

where h is a real symmetric function defined on k-subsets of $\mathcal{X}$, we have the following bound.

Theorem 1.2. *There exist an absolute constant $c > 0$ and a constant $c(k) > 0$ depending only on k such that*

$$\Delta \le \frac{c}{\tau^2}\,\frac{\beta_4+\gamma_4}{\rho^2} + \frac{c(k)}{\tau^2}\,\frac{\delta_3}{\sigma^2\rho^2}\,. \tag{1.10}$$

Since the absolute constants are not specified Theorems 1.1 and 1.2 should be viewed as asymptotic results. Assume that the population size $N \to \infty$ and the sample size n increases so that $n^* \to \infty$. In particular, $\tau \to \infty$. In those models where β_4, γ_4 and $\zeta + \delta_4/\sigma^2$ (respectively δ_3/σ^2) remain bounded and

$$\liminf \rho > 0 \tag{1.11}$$

Theorem 1.1 (respectively Theorem 1.2) provides the bound $\Delta = O(\tau^{-2})$. Since $n^*/2 \le \tau^2 \le n^*$ we obtain $\Delta = O(1/n^*)$.

Remark 1. Note that the bounds of Theorems 1.1 and 1.2 are established without any additional assumption on p and q. This fact is important for applications, like subsampling, where p or q may tend to zero as $N \to \infty$.

Remark 2. The bound of order $O(\tau^{-2})$ for the remainder is unimprovable, because the next term of the Edgeworth expansion, at least for linear statistics, is of order $O(\tau^{-2})$, see Robinson (1978).

Remark 3. An expansion of the probability $P\{T \le \mathbf{E}T + \sigma\tau x\}$ in powers of τ^{-1} would be the most natural choice of asymptotics. We invoke two simple arguments supporting this choice. Firstly, τ^2 is proportional to the variance of the linear part. Secondly, the number of observations n does not longer determine the scale of T in the case where samples are drawn without replacement since the statistic effectively depends on $n^*(\approx \tau^2)$ observations. Indeed, it was shown in Bloznelis and Götze (2001) that, for $n > N - n$, we have almost surely

$$T = T^*, \qquad T^* = \mathbf{E}T + U_1^* + \dots + U_{n^*}^*, \tag{1.12}$$

where we denote

$$U_k^* = \sum_{1 \le i_1 < \cdots < i_k \le n^*} (-1)^k g_k(X'_{i_1}, \ldots, X'_{i_k}), \qquad X'_j = X_{n+j}.$$

That is, T effectively depends on $n^* = N - n$ observations $X'_1, \ldots, X'_{n^*}$ only.

Remark 4. The bounds of Theorems 1.1 and 1.2 are optimal in the sense that it is impossible to approximate F by a continuous differentiable function, like G, with the remainder $o(\tau^{-2})$, if no additional smoothness condition apart from (1.11) is imposed. Already for U-statistics of degree two, Cramér's condition (1.11) together with moment conditions of arbitrary order do not suffice to establish the approximation of order $o(\tau^{-2})$. This fact is demonstrated by means of a counter example in Bentkus, Götze and van Zwet (1997) in the i.i.d. situation, and it is inherited by finite population statistics. Indeed, in the case where $N \to \infty$ and n remains fixed the simple random sample model approaches the i.i.d. situation. We have $\tau \to \sqrt{n}$, $p \to 0$, $q \to 1$. Replacing τ, p and q by $\sqrt{n}$, 0 and 1 respectively we obtain from G the one term Edgeworth expansion for the distribution function of symmetric statistic based on i.i.d. observations, which was constructed in Bentkus, Götze and van Zwet (1997).

Remark 5. The bound (1.8) involves moments (of nonlinear parts) which are higher than those which are necessary to define expansions. Thus, in an optimal dependence on moments one would like to replace $\gamma_4+\zeta+\delta_4/\sigma^2$ by $\gamma_2+\delta_3/\sigma^2$ in the remainder. Let us mention also that for U-statistics of degree k, where k is fixed, the bound (1.10) is more precise than (1.8). Indeed, a straightforward calculation shows that for some absolute constant c we have $\delta_3 \le c\sigma^2\zeta + c\delta_4\tau^{-2}$. Our technique allows us to prove (1.10) for the U-statistics only and with $c(k) \to \infty$, for $k \to \infty$.

In order to apply our results to particular classes of statistics one has to estimate moments δ_3 or δ_4 of differences $\mathbb{D}_3 T$ or $\mathbb{D}_4 T$. For U statistics and smooth functions of sample means this problem is easy and routine, see Bloznelis and Götze (2001). Some applications of our results to resampling procedures are considered ibidem.

The remaining part of the paper is organized as follows. In Section 2 we prove Theorem 1.2. Theorem 1.1 is a consequence of Theorem 1.2. In the proof we use the "data dependent smoothing technique", first introduced in Bentkus, Götze and van Zwet (1997), and expansions of characteristic functions. Expansions of characteristic functions are presented separately in Section 3. Section 4 collects auxiliary combinatorial lemmas. Lemma 4.2 of this section may be of independent interest.

2. Proofs

The section consists of two parts. In the first part, for reader's convenience, we collect several important facts about Hoeffding's decomposition of finite population statistics which are used in proofs below. These facts are shown in Bloznelis and Götze (2001). The second part contains proofs of Theorems 1.1 and 1.2.

Throughout this section and the next we shall assume without loss of generality that $\mathbf{E}T = 0$ and $\sigma^2 = \tau^{-2}$. For $k = 1, 2, \dots$, we write $\Omega_k = \{1, \dots, k\}$. By $C, c, c_0, c_1, \dots$ we denote positive absolute constants. Given two numbers a and $b > 0$, we write $a \ll b$ if $|a| \le c\, b$.

2.1. Hoeffding's decomposition of a finite population statistic

$$T = \mathbf{E}T + U_1 + \cdots + U_n \tag{2.1}$$

was studied by Zhao and Chen (1990) in the case of a U-statistic and by Bloznelis and Götze (2001) in the case of a general symmetric statistic. It was shown in the latter paper that if $n > N/2$ then $U_j \equiv 0$ for $j > N - n$. Note that if T is a U–statistic of degree k (that is, a statistic of the form (1.9)) then always $U_j \equiv 0$ for $j > k$. Recall that U_j is defined in (1.2).

Given $A = \{i_1, \dots, i_r\} \subset \Omega_n$ and $B = \{j_1, \dots, j_s\} \subset \Omega_N$ with $1 \le r \le n$ write

$$T_A = g_r(X_{i_1}, \dots, X_{i_r}) \qquad \text{and} \qquad \mathbf{E}(T_A|B) = \mathbf{E}(T_A|X_{j_1}, \dots, X_{j_s}),$$

and denote $T_\emptyset = \mathbf{E}T$. By symmetry, the random variables T_A and $T_{A'}$ are identically distributed for any $A, A' \subset \Omega_N$ such that $|A| = |A'| \le n$. A simple calculation shows that (1.3) extends to the following identity

$$\mathbf{E}(T_A|B) = 0, \qquad \text{for any} \quad A, B \subset \Omega_N, \quad \text{such that} \quad |B| < |A| \le n. \tag{2.2}$$

For $A, B \subset \Omega_N$ with $1 \le j = |A| = |B| \le n$ and $k = |A \cap B|$ denote

$$\sigma_j^2 = \mathbf{E}T_A^2, \qquad s_{j,k} = \mathbf{E}T_A T_B.$$

Using (2.2) it is easy to show, see e.g., Bloznelis and Götze (2001), that

$$s_{j,k} = (-1)^{j-k} \binom{N-j}{j-k}^{-1} \sigma_j^2, \qquad 0 \le k \le j \le n^*. \tag{2.3}$$

Since $\mathbf{E}T = 0$ we can write (2.1) in the form $T = \sum_{A \subset \Omega_n} T_A$. Similarly, for U_k from (1.2) we have $U_k = \sum_{A \subset \Omega_n,\, |A| = k} T_A$.

2.2 Proofs of Theorems 1.1 and 1.2. The expression $\exp\{ix\}$ is abbreviated by $\mathrm{e}\{x\}$. Given a complex function H defined on $\mathbb{R}$, we write $\|H(x)\| = \sup_{x \in \mathbb{R}} |H(x)|$. Write

$$\delta = 1 - \sup\{\mathbf{E}\cos(tg_1(X_1) + s) :\ s \in \mathbb{R},\ b_1\tau/\beta_3 \le |t| \le \tau^2\}.$$

It is easy to show that $\rho \le \delta$, see Bloznelis and Götze (2000). This inequality will be used in the proof below. We also use the inequalities $1 = \beta_2 \le \beta_3 \le \beta_4^{1/2}$ and $\gamma_2 \le \gamma_3^{2/3} \le \gamma_4^{1/2}$ which are simple consequences of Hölder's inequality.

We can assume that for sufficiently small c_0,

$$\beta_4 \le c_0\tau^2, \qquad \gamma_2 \le c_0\tau^2\delta^2, \qquad \delta^{-2/3}\ln\tau \le c_0\tau. \tag{2.4}$$

Indeed, if (2.4) fails, the bounds (1.8) and (1.10) follow from the inequalities

$$F(x) \le 1 \qquad \text{and} \qquad |G(x)| \ll 1 + \tau^{-1}(\beta_4^{1/2} + \gamma_2^{1/2}).$$

Note that $\beta_4 \ge 1$ and the first inequality in (2.4) imply that $\tau^2 \ge c_0^{-1}$ is sufficiently large.

Proof of Theorem 1.1. The theorem is a consequence of Theorem 1.2. Write $T = U_1 + U_2 + U_3 + R_3$, see (1.6). A Slutzky type argument gives

$$\Delta \le \Delta' + \tau^{-2}\|G^{(1)}(x)\| + \mathbf{P}\{|R_3| \ge \tau^{-2}\}, \tag{2.5}$$

where, $\Delta' := \|\mathbf{P}\{U_1 + U_2 + U_3 \le \sigma\tau x\} - G(x)\|$ satisfies, by (1.10),

$$\Delta' \ll \tau^{-2}\rho^{-2}(\beta_4 + \gamma_4 + \zeta).$$

Furthermore, by (2.4), $\|G^{(1)}(x)\| \ll 1$. We bound this quantity by $c\beta_4$, since $\beta_4 \ge 1$. Finally, by (1.6), we have $\mathbf{P}\{|R_3| \ge \tau^{-2}\} \le \delta_4$ and using the identity $\sigma^2 = \tau^{-2}$ we replace δ_4 by $\tau^{-2}\delta_4/\sigma^2$. Collecting these bounds in (2.5) we obtain (1.8).

Proof of Theorem 1.2. In view of (1.12) it suffices to prove the theorem in the case where $n \le N/2$. Therefore, we assume without loss of generality that $n \le N/2$ and $n^* = n$. The proof of (1.10) is complex and we first outline the main steps.

In the first step we replace T by $\tilde{T} = V_1 + \cdots + V_m + W$, where W is a function of the observations $(X_{m+1}, \ldots, X_n)$ (=:$\mathbb{X}_m$ for short), and where $V_j = V(X_j, \mathbb{X}_m)$ is a function of X_j and $\mathbb{X}_m$. Random variables $W, V_1, \ldots, V_m$ and the integer $m < n$ are specified below.

In the second step we apply Prawitz's (1972) smoothing lemma to construct upper and lower bounds for the conditional distribution function $\tilde{F}_1(x) = \mathbf{P}\{\tilde{T} \le \sigma\tau x | \mathbb{X}_m\}$:

$$\tilde{F}_1(x+) \le 1/2 + VP\int_{\mathbb{R}} \exp\{-itx\}H^{-1}K(t/H)f_1(t)dt, \tag{2.6}$$

$$\tilde{F}_1(x-) \ge 1/2 - VP\int_{\mathbb{R}} \exp\{-itx\}H^{-1}K(-t/H)f_1(t)dt. \tag{2.7}$$

Here $f_1(t) = \mathbf{E}\big(\exp\{it\tilde{T}\}|\mathbb{X}_m\big)$ denotes the conditional characteristic function of $\tilde{T}$; $F(x+) = \lim_{z\downarrow x} F(z)$, $F(x-) = \lim_{z\uparrow x} F(z)$, and VP denotes Cauchy's principal value. The kernel $K(s) = K_1(s)/2 + iK_2(s)/(2\pi s)$, where

$$K_1(s) = \mathbb{I}\{|s| \le 1\}\,(1-|s|), \qquad K_2(s) = \mathbb{I}\{|s| \le 1\}\,((1-|s|)\,\pi\, s\, \cot(\pi\, s) + |s|).$$

The positive random variable $H = O_P(n)$ (a function of $\mathbb{X}_m$) is specified below. Taking the expectations of (2.6) and (2.7) we obtain upper and lower bounds for $F_1(x+)$ and $F_1(x-)$, where $F_1(x) = \mathbf{P}\{\tilde{T} \le \sigma\tau x\}$. Combining these bounds with the inversion formula

$$G(x) = 1/2 + i/(2\pi) \lim_{M\to\infty} VP \int_{|t|\le M} \exp\{-itx\}\hat{G}(t)t^{-1}dt,$$

see, e.g., Bentkus, Götze and van Zwet (1997), we obtain upper bounds for $F_1(x+) - G(x) =: d_1$ and $G(x) - F_1(x-) =: d_2$ respectively. Here $\hat{G}$ denotes the Fourier-Stieltjes transform of G. Thus, for d_1, we have

$$F_1(x+) - G(x) \le \mathbf{E}I_1 + \mathbf{E}I_2 + \mathbf{E}I_3, \tag{2.8}$$

$$I_1 = \frac{1}{2}\, H^{-1} \int_R \exp\{-ix\,t\}\, K_1\big(\frac{t}{H}\big) f_1(t)\, dt,$$

$$I_2 = \frac{i}{2\,\pi}\ \text{V.P.} \int_R \exp\{-ix\,t\} K_2\big(\frac{t}{H}\big)\big(f_1(t) - \hat{G}(t)\big)\,\frac{dt}{t}\,,$$

$$I_3 = \frac{i}{2\,\pi}\ \text{V.P.} \int_R \exp\{-ix\,t\}\Big(K_2\big(\frac{t}{H}\big) - 1\Big)\hat{G}(t)\,\frac{dt}{t}\,.$$

A similar inequality holds for d_2. This step of the proof is called "data depending smoothing", see Bentkus, Götze and van Zwet (1997).

The final step of the proof provides upper bounds for d_1 and d_2. For this purpose we construct an upper bound for $|f_1(t)|$, for $|t| \ge cn^{1/2}/\beta_3$. Using Cramér's condition and the multiplicative structure of f_1 (note that $\tilde{T}$ is conditionally linear given $\mathbb{X}_m$) we show that $|f_1(t)|$ decay exponentially in $|t|$. Furthermore, for $|t| \le cn^{1/2}/\beta_3$ we replace the conditional characteristic function f_1 by the unconditional one $\hat{F}(t) = \mathbf{E}\exp\{itT\}$ and construct bounds for $|\hat{F}(t) - \hat{G}(t)|$ by means of expansions of $\hat{F}(t)$ in powers observations $X_1, \dots, X_n$.

Note that, usually, the validity of an Edgeworth expansion is proved using the conventional Berry–Eseen smoothing lemma, see e.g., Petrov (1975), Callaert, Janssen and Veraverbeke (1980), Bickel, Götze and van Zwet (1986). In the present paper (in the second step of the proof) we use Prawitz's smoothing lemma instead. This lemma is more precise in the sense that the right-hand sides of (2.6) and (2.7) do not involve *absolute values* of characteristic functions. Therefore, after taking the expected values of (2.6) and (2.7) we can interchange the order of integration in the right-hand sides and obtain the unconditional characteristic functions in the integrands. At the same time, the appropriate choice of the random cut-off H allows to control the nonlinear part of $\tilde{T}$ so that the exponential decay of $|f_1(t)|$ is established using the minimal smoothness condition (Cramér's condition on the linear part of the statistic). More restrictive smoothness conditions which involve nonlinear parts of the statistic are considered in Callaert, Janssen and Veraverbeke (1980), Bickel, Götze and van Zwet (1986).

Step 1. Let m denote the integer closest to the number $8\delta^{-1}\ln\tau$. Since, by (2.4), $\tau^2 \geq c_0^{-1}$ we can choose c_0 small enough so that $10 \leq m \leq n/2$. Split $T = V + W$, where

$$V = \sum_{B\subset\Omega_n,\, B\cap\Omega_m\neq\emptyset} T_B, \qquad W = \sum_{B\subset\Omega_n,\, B\cap\Omega_m=\emptyset} T_B,$$

$$V = \sum_{i=1}^{m} V_i + \Lambda_m + Y_m + Z_m, \qquad V_i = T_{\{i\}} + \xi_i + \eta_{m,i}, \qquad \xi_i = \sum_{j=m+1}^{n} T_{\{i,j\}}.$$

Here we denote

$$\Lambda_m = \sum_{B\subset\Omega_m,\, |B|=2} T_B, \qquad Z_m = \sum_{B\subset\Omega_n,\, |B\cap\Omega_m|\geq 3} T_B, \tag{2.9}$$
$$Y_m = \sum_{B\subset\Omega_n,\, |B\cap\Omega_m|=2, |B|\geq 3} T_B, \qquad \eta_{m,i} = \sum_{B\subset\Omega_n,\, B\cap\Omega_m=\{i\},\, |B|\geq 3} T_B.$$

We are going to replace T by $\tilde{T} = \sum_{i=1}^{m} V_i + W$. Write $T = \tilde{T} + R$, where $R = \Lambda_m + Y_m + Z_m$. Given $\varepsilon = \delta^{-2}\tau^{-2}$, a Slutzky type argument gives

$$\Delta \leq \Delta_1 + \varepsilon\, \|G^{(1)}(x)\| + \mathbf{P}\{|R| \geq \varepsilon\}, \qquad \Delta_1 = \|F_1(x) - G(x)\|,$$

where $F_1(x) = \mathbf{P}\{\tilde{T} \leq \sigma\tau x\}$. Invoking the bounds of Lemma 4.1 and the simple bound $\mathbf{E}|\Lambda_m|^3 \ll m^6 \mathbf{E}|T_{\{1,2\}}|^3$, we obtain, by Chebyshev's inequality, that

$$P\{|R| \geq \varepsilon\} \leq \mathbf{P}\{|\Lambda_m| \geq \tfrac{\varepsilon}{3}\} + \mathbf{P}\{|Y_m| \geq \tfrac{\varepsilon}{3}\} + \mathbf{P}\{|Z_m| \geq \tfrac{\varepsilon}{3}\} \ll \frac{\delta_3/\sigma^2 + \gamma_3}{\tau^2\delta^2}.$$

Finally, we bound $\|G^{(1)}(x)\|$ as in proof of Theorem 1.1 above. We obtain

$$\Delta \ll \Delta_1 + \tau^{-2}\delta^{-2}(1 + \delta_3/\sigma^2 + \gamma_3).$$

Therefore, in order to prove (1.10), it remains to bound Δ_1. Clearly, $\Delta_1 \leq \max\{d_1; d_2\}$, where $d_1 = F_1(x+) - G(x)$ and $d_2 = G(x) - F_1(x-)$. The remaining part of the proof provides bounds for d_1 and d_2.

Step 2. Let $\overline{A} = (A_1, \ldots, A_N)$ be a random permutation of $(x_1, \ldots, x_N)$ which is uniformly distributed over the class of permutations. Let $r = [(n+m)/2]$ denote the integer part of $(n+m)/2$. Introduce the sets $\mathcal{I}_0 = \{m+1, \ldots, n\}$, $\mathcal{J}_0 = \{1, \ldots, N\} \setminus \mathcal{I}_0$, $\mathcal{J}_1 = \mathcal{J}_0 \cup \{m+1, \ldots, r\}$ and $\mathcal{J}_2 = \mathcal{J}_0 \cup \{r+1, \ldots, n\}$. Define (random) sub-populations $\mathcal{A}_i = \{A_k,\, k \in \mathcal{J}_i\}$, $i = 0, 1, 2$, and given $\mathcal{A}_i$ let A_i^* be a random variable uniformly distributed in $\mathcal{A}_i$.

We assume that $X_j = A_j$, for $j \in \mathcal{I}_0$ and, given A_j, $j \in \mathcal{I}_0$, the observations $X_1, \ldots, X_m$, are drawn without replacement from $\mathcal{A}_0$. Write

$$H = n\,\delta / \big(32\, q^{-1}\, n\, (\Theta_1 + \Theta_2) + 1\big), \qquad \Theta_i = \mathbf{E}^* |v_i(A_i^*)|, \quad i = 1, 2,$$
$$v_1(a) = \sum_{r+1 \le j \le n} g_2(a, A_j), \qquad v_2(a) = \sum_{m+1 \le j \le r} g_2(a, A_j).$$

Here, given $\mathcal{A}_i$, we denote $\mathbf{E}^* f(A_i^*) = |\mathcal{J}_i|^{-1} \sum_{j \in \mathcal{J}_i} f(A_j)$, for $f : \mathcal{A}_i \to \mathbb{R}$.

In order to prove an upper bound for $F_1(x+) - G(x)$ we apply (2.8) and show that

$$|\mathbf{E} I_1| + |\mathbf{E}(I_2 + I_3)| \ll \tau^{-2} \delta^{-2} (\beta_4 + \gamma_4 + c(k) \delta_3 / \sigma^2). \tag{2.10}$$

An upper bound for $G(x) - F_1(x-)$ is obtained in a similar same way.

In the remaining part of the proof we verify (2.10). Write

$$\hat{F}_1(t) = \mathbf{E}\, \mathrm{e}\{t \tilde{T}\}, \qquad H_1 = b_1 \tau / \beta_3,$$
$$\mathcal{Z}_1 = \{|t| \le H_1\} \subset \mathbb{R}, \qquad \mathcal{Z}_2 = \{H_1 \le |t| \le H\} \subset \mathbb{R},$$
$$J_1 = \int_{\mathcal{Z}_1} \mathrm{e}\{-tx\} \frac{\hat{F}_1(t) - \hat{G}(t)}{t}\, dt, \qquad J_2 = \int_{\mathcal{Z}_2} \frac{|f_1(t)|}{|t|}\, dt,$$
$$J_3 = \int_{|t| > H_1} \frac{|\hat{G}(t)|}{|t|}\, dt, \qquad J_4 = \int_{\mathcal{Z}_1} \mathrm{e}\{-tx\} \frac{f_1(t)}{H}\, dt, \qquad J_5 = \int_{\mathcal{Z}_2} \frac{|f_1(t)|}{H}\, dt,$$

Here and below we write $\mathrm{e}\{x\}$ for $\exp\{ix\}$. Using the inequality $|K_1(s) - 1| \le |s|$ we replace $K_1(s)$ by 1 in $\mathbf{E} I_1$ and obtain

$$|\mathbf{E} I_1| \ll |\mathbf{E} J_4| + \mathbf{E} J_5 + R, \qquad R = \mathbf{E} H_1^2 H^{-2}.$$

Similarly, using the inequality $|K_2(s) - 1| \le 5 s^2$ and the fact that $K_2(s) = 0$ for $|s| > 1$, we obtain

$$|\mathbf{E}(I_2 + I_3)| \ll |\mathbf{E} \tilde{J}_1| + \mathbf{E} J_2 + J_3 + R, \qquad R = \mathbf{E} H_1^2 H^{-2},$$

where $\tilde{J}_1$ is defined as J_1 above, but with $\hat{F}_1$ replaced by f_1. Note that the change of the order of integration yields $\mathbf{E} \tilde{J}_1 = J_1$.

The bounds $J_3 \ll \tau^{-2}(\beta_4+\gamma_2)$ and $R \ll \delta^{-2}\tau^{-2}(1+\gamma_2)$ are proved in Bloznelis and Götze (2000). Furthermore, note that $J_5 \le J_2$. Therefore, in order to prove (2.10) it suffices to show that

(2.11)

$$\mathbf{E}J_2 \ll \frac{\beta_4+\delta_3/\sigma^2}{\tau^2}, \quad |\mathbf{E}J_4| \ll \frac{1+\gamma_2+\delta_3/\sigma^2}{\tau^2\delta^2}, \quad |J_1| \ll \frac{\beta_4+\gamma_4+c(k)\delta_3/\sigma^2}{\tau^2\delta^2}.$$

Step 3. Here we bound $|J_1|$, $\mathbf{E}J_2$, and $|\mathbf{E}J_4|$.

The bound for $\mathbf{E}J_2$. Given t, denote $\mathbb{I}_t = \mathbb{I}\{|t|\mathbf{E}^*|\eta_m(A_0^*)| < \delta/16\}$, where $\eta_m(x) = \mathbf{E}(\eta_{m,1}|X_1 = x, X_{m+1}, \dots, X_n)$. The identity $f_1(t) = \mathbb{I}_t f_1(t) + (1-\mathbb{I}_t) f_1(t)$ combined with the inequalities

$$1-\mathbb{I}_t \le 16^2\delta^{-2}t^2\big(\mathbf{E}^*|\eta_m(A_0^*)|\big)^2 \le 16^2\delta^{-2}t^2\mathbf{E}^*\eta_m^2(A_0^*)$$

yields $J_2 \le J_{2.1} + J_{2.2}$, where

$$J_{2.1} = \int_{\mathcal{Z}_2} \mathbb{I}_t \frac{|f_1(t)|}{|t|}\,dt \qquad \text{and} \qquad J_{2.2} = 16^2\delta^{-2}\mathbf{E}^*\eta_m^2(A_0^*)H^2.$$

In order to prove the bound (2.11) for $\mathbf{E}J_2$ we show that

$$\mathbf{E}J_{2.2} \ll \tau^{-2}\delta_3/\sigma^2 \qquad \text{and} \qquad \mathbf{E}J_{2.1} \ll \beta_3/n^2.$$

The first bound is a consequence of the inequalities $H \le n\delta$ and $\mathbf{E}\mathbf{E}^*\eta_m^2(A_0^*) \le n^{-3}\delta_3/\sigma^2$, where the latter inequality follows from (4.1) by symmetry.

The proof of the bound $\mathbf{E}J_{2.1} \ll \beta_3/n^2$ is almost the same as that of the corresponding inequality (3.12) in Bloznelis and Götze (2000). The only and minor difference is that now we add one more nonlinear term η_m. Namely, in the proof of (3.12) ibidem one should replace $v(a) = v_1(a) + v_2(a)$ by $\tilde{v}(a) = v(a) + \eta_m(a)$ and use the bound $E^*|t\,\eta_m(A_0^*)| \le \delta/16$ when estimating $\mathbf{E}^*(1+2u_2(A_0^*))$ in (3.17), ibidem. Indeed, the bound $E^*|t\,\eta_m(A_0^*)| \le \delta/16$ holds on the event $\{\mathbb{I}_t \ne 0\}$.

The bound for $\mathbf{E}J_4$. Define J_4' in the same way as J_4, but with $f_1(t)$ replaced by $\mathbf{E}(\mathrm{e}\{t\tilde{U}\}\,|\,\mathbb{X}_m)$, where $\tilde{U} = U - \Lambda_m$. We shall apply the bound $|\mathbf{E}J_4'| \ll n^{-1}(1+\gamma_2)/\delta^2$ which is proved in (3.20) of Bloznelis and Götze (2000). This bound and the inequality

$$|\mathbf{E}J_4 - \mathbf{E}J_4'| \ll \tau^{-2}\delta^{-2}(1+\delta_3/\sigma^2+\gamma_2) \tag{2.12}$$

yield the bound (2.11) for $\mathbf{E}J_4$.

In order to prove (2.12) we write $\tilde{T} - \tilde{U} = R_2 - Y_m - Z_m$, see (1.6) and (2.9). This identity in combination with the inequality $|\,\mathrm{e}\{x\} - \mathrm{e}\{y\}| \le |x-y|$ yields $|\mathbf{E}J_4 - \mathbf{E}J_4'| \ll R$, where

$$\begin{aligned} R &= H_1^2\mathbf{E}H^{-1}\mathbf{E}\big(|R_2 - Y_m - Z_m|\,\big|\,\mathbb{X}_m\big) \\ &\le H_1^2(\mathbf{E}H^{-2})^{1/2}\Big[(\mathbf{E}R_2^2)^{1/2} + (\mathbf{E}Y_m^2)^{1/2} + (\mathbf{E}Z_m^2)^{1/2}\Big]. \end{aligned}$$

In the last step we applied Cauchy–Schwarz. It follows from (1.6), (4.1) and the last inequality of (2.4) that the quantity in the brackets $[\dots] \ll \tau^{-2}\delta^{-1}\delta_3^{1/2}/\sigma$. Finally, invoking the bound $\mathbf{E}H^{-2} \ll n^{-2}\delta^{-2}(1+\gamma_2)$, which is proved in (5.1) of Bloznelis and Götze (2000), we complete the proof of (2.12).

The bound for $|J_1|$. The bound (2.11) is a consequence of the following two bounds

$$|\mathcal{I}_{\mathcal{Z}_1}(\hat{F}_1 - \hat{F})| \ll \tau^{-2}\delta^{-2}(1+\gamma_2+\delta_3/\sigma^2), \tag{2.13}$$

$$|\mathcal{I}_{\mathcal{Z}_1}(\hat{F} - \hat{G})| \ll \tau^{-2}\delta^{-2}(\beta_4+\gamma_4+c(k)\delta_3/\sigma^2). \tag{2.14}$$

Recall that $\hat{F}(t) = \mathbf{E}\,\mathrm{e}\{tT\}$. Here and below for a Borel set $\mathcal{B} \subset \mathbb{R}$ and an integrable complex function f, we write (for short) $\mathcal{I}_{\mathcal{B}}(f) = \int_{\mathcal{B}} t^{-1} f(t)dt$.

The proof of (2.14) is rather complex. We place it in a separate Section 3 below. Note that the bound (2.14) is the only step of the proof where we essentially use the assumption that T is a U–statistic.

Here we show (2.13). In the proof we replace $\hat{F}_1(t) - \hat{F}(t)$ by $f(t) = \mathbf{E}\,\mathrm{e}\{tT\}it\Lambda_m$ and then replace $f(t)$ by $g(t) = \mathbf{E}\,\mathrm{e}\{tU\}it\Lambda_m$. Finally, we invoke the bound

$$|\mathcal{I}_{\mathcal{Z}_1}(g)| \ll \binom{m}{2} n^{-3/2}(1+\gamma_2) \ll n^{-1}\delta^{-2}(1+\gamma_2), \tag{2.15}$$

which is proved in (3.38) of Bloznelis and Götze (2000). In order to show how (2.13) follows from (2.15) write $\tilde{T} = T - \Lambda_m - Y_m - Z_m$, see (2.9). We have $\hat{F}_1(t) = \mathbf{E}\,\mathrm{e}\{t(T - \Lambda_m - Y_m - Z_m)\}$. Expanding the exponent in powers of $it(Y_m + Z_m)$ and then in powers of $it\Lambda_m$ we get

$$|\hat{F}_1(t) - \hat{F}(t) - f(t)| \ll \mathbf{E}|tY_m| + \mathbf{E}|tZ_m| + t^2\mathbf{E}\Lambda_m^2.$$

Furthermore, the identity $T - U = R_2$ combined with the mean value theorem yields $|f(t) - g(t)| \ll \mathbf{E}t^2|\Lambda_m R_2|$. Combining these two inequalities we obtain

$$|\mathcal{I}_{\mathcal{Z}_1}(\hat{F}_1 - \hat{F})| \ll |\mathcal{I}_{\mathcal{Z}_1}(g)| + R, \quad R = 2H_1\mathbf{E}|Y_m| + 2H_1|Z_m| + H_1^2\mathbf{E}\Lambda_m^2 + H_1^2\mathbf{E}|\Lambda_m R_2|.$$

Now, the bound (2.11) follows from (2.15) and the bound $R \ll \tau^{-2}\delta^{-2}(1+\gamma_2+\delta_3/\sigma^2)$. The latter bound follows from the inequalities (4.1), (1.6), and the simple inequality $\mathbf{E}\Lambda_m^2 \le m^2 n^{-3}\gamma_2$, via Hölder's inequality.

The proof of Theorem 1.2 is complete.

3. Expansions

Here we show (2.14). Split $\mathcal{Z}_1 = \mathcal{B}_1 \cup \mathcal{B}_2$ where $\mathcal{B}_1 = \{|t| \le c_1\}$ and $\mathcal{B}_2 = \{c_1 \le |t| \le H_1\}$, and where c_1 is an absolute constant. Clearly, (2.14) follows from the obvious inequalities

$$\mathcal{I}_{\mathcal{Z}_1}(\hat{F} - \hat{G}) \le \mathcal{I}_{\mathcal{B}_1}(\hat{F} - \hat{G}) + \mathcal{I}_{\mathcal{B}_2}(\hat{F} - \hat{G}),$$

$$\mathcal{I}_{\mathcal{B}_1}(\hat{F} - \hat{G}) \le \mathcal{I}_{\mathcal{B}_1}(\hat{F} - \hat{F}_U) + \mathcal{I}_{\mathcal{B}_1}(\hat{F}_U - \hat{G}), \qquad \hat{F}_U(t) = \mathbf{E}\,\mathrm{e}\{tU\},$$

and the bounds

$$\mathcal{I}_{\mathcal{B}_2}(\hat{F}-\hat{G}) \ll \mathcal{R}, \qquad \mathcal{R} = \frac{\beta_4+\gamma_4+c(k)\delta_3/\sigma^2}{\tau^2\delta^2}, \tag{3.1}$$

$$\mathcal{I}_{\mathcal{B}_1}(\hat{F}_U-\hat{G}) \ll \frac{\beta_4+\gamma_4}{\tau^2}, \qquad \mathcal{I}_{\mathcal{B}_1}(\hat{F}-\hat{F}_U) \ll \frac{1+\delta_3/\sigma^2}{\tau^2}. \tag{3.2}$$

The first bound of (3.2) is proved in (4.1) of Bloznelis and Götze (2000). To prove the second bound we decompose $T = U + R_2$, see (1.6), and apply the mean value theorem to get $|\hat{F}(t) - \hat{F}_U(t)| \le \mathbf{E}|tR_2|$. Finally, an application of (1.6) gives $\mathbf{E}|R_2| \le (\mathbf{E}R_2^2)^{1/2} \le \tau^{-2}\delta_3^{1/2}/\sigma$, and, obviously, $\delta_3^{1/2}/\sigma \le 1+\delta^3/\sigma^2$.

In order to prove (3.1) we write the characteristic function $\hat{F}(t)$ in the Erdős-Rényi (1959) form, see (3.4) below. Let $\nu = (\nu_1,\dots,\nu_N)$ be i.i.d. Bernoulli random variables independent of $(X_1,\dots,X_N)$ and having probabilities $\mathbf{P}\{\nu_1 = 1\} = p$ and $\mathbf{P}\{\nu_1 = 0\} = q$. Observe, that the conditional distribution of

$$T^* = \sum_{A\subset\Omega_N, |A|\le n} T_A\nu_A^*, \qquad \text{where} \qquad \nu_A^* = \prod_{i\in A}\nu_i,$$

given the event $\mathcal{E} = \{S_\nu = n\}$, coincides with the distribution of T. Here $S_\nu = \nu_1+\dots+\nu_N$. Therefore, $\hat{F}$ can be written as follows

$$\hat{F}(t) = \lambda\int_{-\pi\tau}^{\pi\tau}\mathbf{E}\,\mathrm{e}\{tT^*+\tau^{-1}s(S_\nu-n)\}ds, \qquad \lambda^{-1} = 2\pi\mathbf{P}\{\mathcal{E}\}\tau. \tag{3.3}$$

Using (2.2) it is easy to show that, for $1 \le k \le n$, almost surely

$$\sum_{A\subset\Omega_N,\,|A|=k} T_A\nu_A^* = \sum_{A\subset\Omega_N,\,|A|=k} Q_A, \qquad Q_A = T_A\tilde{\nu}_A, \qquad \tilde{\nu}_A = \prod_{i\in A}(\nu_i-p).$$

Therefore, almost surely, $tT^*+\tau^{-1}s(S_\nu-n) = S+tQ$, where

$$S = \sum_{i=1}^N S_i, \quad S_i = (tT_{\{i\}}+\tau^{-1}s)(\nu_i-p), \qquad Q = \sum_{A\subset\Omega_N,\,2\le|A|\le n} Q_A.$$

Substitution of this identity in (3.3) gives

$$\hat{F}(t) = \lambda\int_{-\pi\tau}^{\pi\tau}\mathbf{E}\,\mathrm{e}\{S+tQ\}ds. \tag{3.4}$$

In view of (3.4), the bound (3.1) follows from the inequalities

$$\int_{t\in\mathcal{B}_2}\lambda\int_{|s|\le\pi\,\tau}\bigl|\mathbf{E}\ \mathrm{e}\{S+tQ\}-(h_1+h_2)\bigr|ds\,\frac{dt}{|t|} \ll \mathcal{R}, \tag{3.5}$$

$$\int_{t\in\mathcal{B}_2}\bigl|\lambda\int_{|s|\le\pi\,\tau}(h_1+h_2)ds-\hat{G}(t)\bigr|\,\frac{dt}{|t|} \ll \mathcal{R}. \tag{3.6}$$

Here

$$h_1 = \mathbf{E}\,\mathrm{e}\{S\}, \qquad h_2 = i^3\binom{n}{2}\mathbf{E}\,\mathrm{e}\{S_3+\dots+S_N\}V, \qquad V = tQ_{\{1,2\}}S_1S_2.$$

The inequality (3.6) is proved in Bloznelis and Götze (2000) (formula (4.2)).

We are going to prove (3.5). Before the proof we introduce some notation. Given a complex valued function $f(s,t)$ we write $f \prec \mathcal{R}$ if

$$\int_{\mathcal{B}_2}\frac{dt}{|t|}\int_{-\pi\tau}^{\pi\tau}|f(s,t)|ds \ll \mathcal{R}.$$

Furthermore, we write $f \sim g$ for $f-g \prec \mathcal{R}$. In view of the inequality $\lambda \le \sqrt{2}\pi$, see Höglund (1978), the bound (3.5) can be written as follows: $\mathbf{E}\,\mathrm{e}\{S+tQ\} \sim h_1+h_2$.

Let us prove (3.5). In what follows we assume that $t \in \mathcal{B}_2$ and $|s| \le \pi\tau$. Given s,t write $u = s^2+t^2$ and let (here and throughout the section) m denote the integer closest to the number $c_2N\,u^{-1}\ln u$, where c_2 is an absolute constant. We choose c_1 and c_2 so that $10 < m < N/2$. Split

$$(3.7) \quad Q = K+L+W+Y+Z, \qquad K = \zeta+\mu, \quad \zeta = \sum_{j=1}^m \zeta_j, \quad \mu = \sum_{j=1}^m \mu_j,$$

$$\zeta_j = \sum_{A\cap\Omega_m=\{j\},\,|A|=2} Q_A, \qquad \mu_j = \sum_{A\cap\Omega_m=\{j\},\,|A|\ge 3} Q_A, \qquad 1 \le j \le m,$$

$$L = \sum_{A\subset\Omega_m,\,|A|=2} Q_A, \qquad Y = \sum_{|A\cap\Omega_m|=2,\,|A|\ge 3} Q_A,$$

$$Z = \sum_{|A\cap\Omega_m|\ge 3} Q_A, \qquad W = \sum_{A\cap\Omega_m=\emptyset,\,|A|\ge 2} Q_A,$$

and denote $f_1 = \mathbf{E}\,\mathrm{e}\{S+t(K+W)\}$ and $f_2 = \mathbf{E}\,\mathrm{e}\{S+t(K+W)\}itL$.

In order to prove $\mathbf{E}\exp\{S+tQ\} \sim h_1+h_2$ we shall show that

$$(3.8) \qquad \mathbf{E}\,\mathrm{e}\{S+tQ\} \sim f_1+f_2,$$

$$(3.9) \qquad f_2 \sim h_3, \qquad h_3 = i^3\binom{m}{2}\mathbf{E}\,\mathrm{e}\{S_3+\dots+S_N\}V,$$

$$(3.10) \qquad f_1 \sim h_1+h_2-h_3.$$

Let us introduce some more notation. Given a sum $v = v_1+\dots+v_k$ we write $v_B = \sum_{j\in B} v_j$, for $B \subset \Omega_k$. Given $B \subset \Omega_N$, by $\mathbf{E}_{(B)}$ we denote the conditional expectation given all the random variables, but ν_j, $j \in B$. For $D \subset \Omega_m$, denote

$$(3.11) \qquad \mathbb{Y}_D = |\mathbf{E}_{(D)}\,\mathrm{e}\{S_D\}|, \quad \mathbb{Z}_D = |\mathbf{E}_{(D)}\,\mathrm{e}\{S_D+t\zeta_D\}|, \quad \mathbb{I}_D = \mathbb{I}\{\varkappa_D > c_3^{-1}\},$$

$$\varkappa_D = \tau^2|D|^{-1}\sum_{j\in D}\zeta^2(X_j), \qquad \zeta(a) = \sum_{j=m+1}^N g_2(a,X_j)(\nu_j-p).$$

Using the multiplicative structure of $\mathbb{Y}_D$ one can prove a sufficiently fast decay of its expected value as $u \to \infty$. More precisely, one can construct random variables $\mathbb{F}_D$, $D \subset \Omega_N$, of the form $\mathbb{F}_D = \prod_{i \in D} \tilde{u}(tg_1(X_i) + s/\tau)$, where $\tilde{u} \geq 0$ is a real function, such that for $|D| \geq m/4$ we have

$$\text{(3.12)} \quad \mathbb{Y}_D \leq \mathbb{F}_D, \quad \mathbb{Z}_D \leq \mathbb{I}_D + \mathbb{F}_D, \quad \mathbf{E}(\mathbb{F}_D^2 \,|\, X_i, X_j) \ll u^{-20}, \quad \forall\, i, j \in \Omega_N \setminus D.$$

Clearly, the latter inequality holds for the unconditional expectation as well. the formulas (4.9) and (4.10) ibidem. The proof of (3.12) and the construction of random variables $\mathbb{F}_D$ are provided in formulas (4.7-10) of Bloznelis and Götze (2000). Let us mention that in order to establish (3.12), one chooses the constants c_1, c_2 and c_3 in an appropriate way.

Split $\Omega_m = \Omega_m^1 \cup \Omega_m^2 \cup \Omega_m^3$, where Ω_m^i, $i = 1, 2, 3$ are disjoint consecutive intervals with cardinalities $|\Omega_m^i| \approx m/3$. For $i \leq j$, let $\Omega_{i,j}$ denote the set of all pairs $\{l, r\}$ such that $l \in \Omega_m^i$, $r \in \Omega_m^j$ and $l < r$.

Proof of (3.8). Expanding the exponent in powers of itZ and invoking (3.17) we get $\mathbf{E}\,\mathrm{e}\{S + tQ\} = f_3 + R$, where $f_3 = \mathbf{E}\,\mathrm{e}\{S + t(K + L + W + Y)\}$ and where

$$|R| \leq \mathbf{E}|tZ| \leq |t|(\mathbf{E}Z^2)^{1/2} \ll |t| c_1^{1/2}(k) u^{-3/2} \ln^{3/2} u\, \delta_3^{1/2} \sigma^{-1} \tau^{-2} \prec \mathcal{R}.$$

Furthermore, expanding f_3 in powers of $it(L + Y)$ we obtain

$$f_3 = f_1 + f_2 + f_4 + R, \qquad f_4 = \mathbf{E}\,\mathrm{e}\{S + t(K + W)\} itY,$$
$$|R| \ll t^2 \mathbf{E}(L + Y)^2 \ll t^2 u^{-2} \ln^2 u\, \tau^{-2}(\gamma_2 + c_1(k)\delta_3/\sigma^2) \prec \mathcal{R}.$$

In the last step we invoked (3.18) and used the identity $\mathbf{E}L^2 = \binom{m}{2} p^2 q^2 \tau^{-6} \gamma_2$. We obtain $\mathbf{E}\,\mathrm{e}\{S + tQ\} \sim f_1 + f_2 + f_4$.

It remains to prove that $f_4 \prec \mathcal{R}$. To this aim we show that $f_4 \sim f_5$ and $f_5 \prec \mathcal{R}$, where $f_5 = \mathbf{E}\,\mathrm{e}\{S + t(\zeta + W)\} itY$. By the mean value theorem $|f_4 - f_5| \leq \mathbf{E}t^2|Y\mu|$. Furthermore, by Cauchy–Schwarz and (3.18), (3.19),

$$\mathbf{E}t^2|Y\mu| \leq t^2(\mathbf{E}Y^2)^{1/2}(\mathbf{E}\mu^2)^{1/2} \ll t^2 c_1(k) u^{-3/2} \ln^{3/2} u\, \tau^{-4} \delta_3/\sigma^2 \prec \mathcal{R}.$$

Therefore, $f_4 \sim f_5$. In order to prove $f_5 \prec \mathcal{R}$ we split $f_5 = \sum_{1 \leq i \leq j \leq 3} f_{i,j}$ and show that $f_{i,j} \prec \mathcal{R}$, for every $i \leq j$. Here $f_{i,j}$ is defined in the same way as f_5, but with Y replaced by $Y_{i,j}$, where $Y_{i,j}$ denotes the sum of all Q_A such that $A \cap \Omega_m \in \Omega_{i,j}$. Given i, j choose r from $\{1, 2, 3\} \setminus \{i, j\}$. Note that the random variable $Y_{i,j}$ and the sequence $\{\nu_l,\, l \in \Omega_m^r\}$ are independent. Therefore, by (3.12),

$$|f_{i,j}| \leq |t| \mathbf{E}\mathbb{Z}_{\Omega_m^r} |Y_{i,j}| \leq |t|(\mathbf{E}\mathbb{Z}_{\Omega_m^r}^2)^{1/2}(\mathbf{E}Y_{i,j}^2)^{1/2}$$
$$\leq |t|(\mathbf{E}\mathbb{F}_{\Omega_m^r}^2 + c_3 \mathbf{E}\varkappa_{\Omega_m^r})^{1/2}(\mathbf{E}Y_{i,j}^2)^{1/2}.$$

In the last step we used the simple inequality $\mathbb{I}_D \le c_3 \varkappa_D$, for $D = \Omega_m^r$. Note that the bound (3.18) applies to $\mathbf{E}Y_{i,j}^2$ as well. This bound in combination with (3.12) and (3.24) implies $f_{i,j} \prec \mathcal{R}$ thus completing the proof of (3.8).

Proof of (3.9). Split $W = W_0 + W_1$, where

$$W_0 = \sum_{A \subset \Omega_N :\, A \cap \Omega_m = \emptyset, |A|=2} Q_A, \qquad W_1 = \sum_{A \subset \Omega_N :\, A \cap \Omega_m = \emptyset, \, |A| \ge 3} Q_A. \tag{3.13}$$

In order to prove (3.9) we replace f_2 by $f_6 = \mathbf{E}\,\mathrm{e}\{S + t(\zeta + W)\} itL$ and then replace f_6 by $f_7 = \mathbf{E}\,\mathrm{e}\{S + t(\zeta + W_0)\} itL$. Finally, we invoke the relation $f_7 \sim h_3$ which is proved in Bloznelis and Götze (2000) (formula (4.15)).

Let us prove $f_2 \sim f_6$ and $f_6 \sim f_7$. By the mean value theorem, $|f_2 - f_6| \le t^2 \mathbf{E}|L\mu|$. Invoking (3.19) and the bound $\mathbf{E}L^2 \le m^2 p^2 q^2 \tau^{-6} \gamma_2$ we obtain, by Cauchy–Schwarz, $t^2 \mathbf{E}|L\mu| \prec \mathcal{R}$. Hence, $f_2 \sim f_6$.

Let us prove $f_6 \sim f_7$. Split $L = \sum_{1 \le i \le j \le 3} L_{i,j}$, where $L_{i,j} = \sum_{A \in \Omega_{i,j}} Q_A$. Write $f_{6|i,j} = \mathbf{E}\,\mathrm{e}\{S + t(\zeta + W)\} itL_{i,j}$ and $f_{7|i,j} = \mathbf{E}\,\mathrm{e}\{S + t(\zeta + W_0)\} itL_{i,j}$. In order to prove $f_6 \sim f_7$ we show that $f_{6|i,j} \sim f_{7|i,j}$, for every $1 \le i \le j \le 3$. Given $i \le j$, choose $r \in \{1, 2, 3\} \setminus \{i, j\}$ and denote $D = \Omega_m^r$. Expanding the exponent in $f_{6|i,j}$ in powers of itW_1 we obtain $f_{6|i,j} = f_{7|i,j} + t^2 R$, where

$$|R| \le \mathbf{E}\mathbb{Z}_D |L_{i,j} W_1| \le R_1 + R_2, \quad R_1 = \mathbf{E}\mathbb{F}_D |L_{i,j} W_1|, \quad R_2 = \mathbf{E}\mathbb{I}_D |L_{i,j} W_1|,$$

by (3.12). By Cauchy–Schwarz, we have

$$R_1^2 \le \mathbf{E}W_1^2 \mathbf{E}L_{i,j}^2 \mathbb{F}_D^2, \qquad R_2^2 \le \mathbf{E}W_1^2 \mathbf{E}L_{i,j}^2 \mathbb{I}_D \le \mathbf{E}W_1^2 \mathbf{E}L_{i,j}^2 c_3 \varkappa_D \tag{3.14}$$

Fix $\{i_1, i_2\} \in \Omega_{i,j}$. By symmetry, (3.12) and (3.24),

$$\mathbf{E}L_{i,j}^2 \mathbb{F}_D^2 = |\Omega_{i,j}| p^2 q^2 \mathbf{E} g_2^2(X_{i_1}, X_{i_2}) \mathbf{E}(\mathbb{F}_D^2 \,|i_1, i_2) \ll u^{-20} \tau^{-2} \gamma_2, \tag{3.15}$$

$$\mathbf{E}L_{i,j}^2 \varkappa_D = |\Omega_{i,j}| p^2 q^2 \mathbf{E} g_2^2(X_{i_1}, X_{i_2}) \mathbf{E}(\varkappa_D \,|i_1, i_2) \ll u^{-2} \ln^2 u \, \tau^{-4} \gamma_2^2. \tag{3.16}$$

Here we estimated $|\Omega_{i,j}| < m^2$ and $m^2 p^2 q^2 \ll \tau^4 u^{-2} \ln^2 u \ll \tau^4$. It follows from (3.14), (3.15), (3.16) and (3.20) that $t^2 R \prec \mathcal{R}$. We obtain $f_{6|i,j} \sim f_{7|i,j}$ thus completing the proof of (3.9).

Proof of (3.10). In the proof we replace f_1 by $f_8 = \mathbf{E}\,\mathrm{e}\{S + t(\zeta + W)\}$ and then replace f_8 by the sum $f_{10} + f_{11}$, where $f_{10} = \mathbf{E}\,\mathrm{e}\{S + tW\}$ and $f_{11} = \mathbf{E}\,\mathrm{e}\{S + tW\} it\zeta$. Finally, we replace f_{10} by $f_{12} = \mathbf{E}\,\mathrm{e}\{S + tW_0\}$ and f_{11} by $f_{13} = \mathbf{E}\,\mathrm{e}\{S + tW_0\} it\zeta$ (recall that W_0 is defined in (3.13)), and invoke the relation $f_{12} + f_{13} \sim h_1 + h_2 - h_3$, which is proved in Bloznelis and Götze (2000) (formulas (4.36-37)).

Let us prove $f_1 \sim f_8$. Expanding in f_1 powers of $it\mu$ we obtain

$$f_1 = f_8 + f_9 + R, \qquad f_9 = \mathbf{E}\,\mathrm{e}\{S + t(\zeta + W)\} it\mu,$$

where $|R| \le t^2\mu^2 \prec \mathcal{R}$, by (3.19). Therefore, $f_1 \sim f_8 + f_9$. In order to show $f_9 \prec \mathcal{R}$, split $\mu = \mu_1^* + \mu_2^* + \mu_3^*$, where $\mu_j^* = \sum_{k \in \Omega_m^j} \mu_k$, and write $f_9 = f_1^* + f_2^* + f_3^*$, where $f_j^* = \mathbf{E}\,\mathrm{e}\{S + t(\zeta + W)\}it\mu_j^*$. Given j, we show $f_j^* \prec \mathcal{R}$. Fix $r \in \{1,2,3\} \setminus \{j\}$ and denote $D = \Omega_m^r$. By (3.12),

$$|f_j^*| \le |t|\mathbf{E}\mathbb{Z}_D|\mu_j^*| \le |t|(\mathbf{E}(\mu_j^*)^2)^{1/2}\big(\mathbf{E}2(\mathbb{F}_D^2 + \mathbb{I}_D)\big)^{1/2}$$
$$\le 2|t|(\mathbf{E}(\mu_j^*)^2)^{1/2}(\mathbf{E}\mathbb{F}_D^2 + c_3^{3/2}\mathbf{E}\varkappa_D^{3/2})^{1/2}.$$

Here we applied Cauchy–Schwarz and the inequality $\mathbb{I}_D \le c_3^{3/2}\varkappa_D^{3/2}$. Note that the bound (3.19) holds for μ_j^* as well. This bound in combination with (3.24) and (3.12) gives $f_j^* \prec \mathcal{R}$. We obtain $f_9 \prec \mathcal{R}$, thus, completing the proof of $f_1 \sim f_8$.

In order to replace f_8 by $f_{10} + f_{11}$ we use the relation $f_8 \sim f_{10} + f_{11}$. The proof of this relation is almost the same as that of the corresponding relation (4.35) in Bloznelis and Götze (2000).

It remains to show that $f_{10} \sim f_{12}$ and $f_{11} \sim f_{13}$. Expanding f_{10} in powers of itW_1 we get $f_{10} = f_{12} + R$, where $|R| \le \mathbf{E}\mathbb{Y}_{\Omega_m}|tW_1|$. It follows from (3.12) and (3.20), by Cauchy–Schwarz, that $|R| \prec \mathcal{R}$. Therefore, $f_{10} \sim f_{12}$. In order to prove $f_{11} \sim f_{13}$ split $\Omega_m = V_1 \cup V_2$, where $V_1 \cap V_2 = \emptyset$ and the cardinality $|V_i| \approx m/2$, for $i = 1,2$, and write $\zeta = \zeta_{V_1} + \zeta_{V_2}$. Expanding f_{11} in powers of itW_1, we get $f_{11} = f_{13} + R_{(1)} + R_{(2)}$, where $|R_{(i)}| \le \mathbf{E}t^2|W_1\zeta_{V_i}|\mathbb{Y}_{\Omega_m \setminus V_i}$. Fix $r \in V_1$. By Cauchy–Schwarz and symmetry,

$$|R_{(1)}|^2 \le t^2\mathbf{E}W_1^2\mathbf{E}\zeta_{V_1}^2\mathbb{Y}_{V_2}^2 = t^2\mathbf{E}W_1^2|V_1|(N-m)p^2q^2\mathbf{E}g_2^2(X_r, X_N)\mathbf{E}(\mathbb{Y}_{V_2}^2|X_r, X_N)$$
$$\ll t^2\tau^{-6}u^{-20}\gamma_2 c(k)\delta_3/\sigma^2.$$

Here we applied (3.20) and (3.12). It follows that $R_{(1)} \prec \mathcal{R}$. The same bound holds for $R_{(2)}$ as well. Therefore $f_{11} \sim f_{13}$. The proof of (3.10) is complete.

Lemma 3.1. *Assume that $n \le N/2$ and that T is a U–statistic of degree k. Then*

$$\mathbf{E}Z^2 \le c_1(k)m^3p^3q^3\tau^{-10}\delta_3\sigma^{-2}, \tag{3.17}$$
$$\mathbf{E}Y^2 \le c_1(k)m^2p^2q^2\tau^{-8}\delta_3\sigma^{-2}, \tag{3.18}$$
$$\mathbf{E}\mu^2 \le c_1(k)mpq\tau^{-6}\delta_3\sigma^{-2}, \tag{3.19}$$
$$\mathbf{E}W_1^2 \le c_1(k)\tau^{-4}\delta_3\sigma^{-2}, \tag{3.20}$$

where Y, Z and μ are defined in (3.7), and W_1 is defined in (3.13). Here $c_1(k)$ denotes a constant which depends on k only.

Proof of Lemma 3.1. Note that our assumption that T is a U-statistic of degree k implies $\sigma_j^2 = 0$ for $j > k$. In the proof we use the bound which follows from (4.3),

(4.4)

$$\sum_{j=3}^{k} a_j \sigma_j^2 \ll \tau^{-8}\delta_3, \qquad a_j = \binom{n-3}{j-3}\binom{N-n}{j-3}\binom{N-j}{j-3}^{-1}. \tag{3.21}$$

A simple calculation shows that $a_3 = 1$ and $a_j = b_1 \cdots b_{j-3} b_1^* \dots b_{j-3}^*$, where $b_i = (n-2-i)/(N-j-i+1)$ and $b_i^* = N-n-i+1$. For $i = 1, \dots, j-3$ and $j = 4, \dots, k$ we apply the following simple bounds $p \le c(k) b_i$ and $q \le c(k) b_i^*/N$, for some constant $c(k)$ depending on k only. These bounds imply for $j = 3, \dots, k$, that $(pqN)^{j-3} \le c_1(k) a_j$, where $c_1(k)$ is a constant depending on k only. Combining this bound and (3.21) we obtain

$$\sum_{j=3}^{k} p^{j-3} q^{j-3} N^{j-3} \sigma_j^2 \ll c_1(k) \tau^{-8} \delta_3. \tag{3.22}$$

We introduce some more notation. Consider a random variable $M = \sum_{A \in \mathcal{M}} Q_A$, where $\mathcal{M}$ denotes a class of subsets of Ω_N, i.e., $\mathcal{M} \subset \{A \subset \Omega_N : 3 \le |A| \le n\}$. Note that $\mathbf{E} Q_A = 0$, for every A and $\mathbf{E} Q_A^2 = \sigma_j^2 p^j q^j$, for $|A| = j$. Since Q_A and Q_B are uncorrelated for $A \ne B$, we have

$$\mathbf{E} M^2 = \sum_{A \in \mathcal{M}} \mathbf{E} Q_A^2 = \sum_{j=3}^{n} e_j[M] \sigma_j^2 p^j q^j, \tag{3.23}$$

where $e_j[M]$ denotes the number of different subsets $A \in \mathcal{M}$ of size $|A| = j$.

Let us prove (3.17). A simple calculation shows that

$$e_j[Z] = \sum_{r=0}^{j-3} \binom{m}{j-r}\binom{N-m}{r}, \qquad 3 \le j \le k.$$

Invoking the inequality $\binom{m}{j-r} \le m^3 \binom{m}{j-r-3}$ we obtain

$$e_j[Z] \le m^3 \sum_{r=0}^{j-3} \binom{m}{j-r-3}\binom{N-m}{r} = m^3 \binom{N}{j-3}.$$

by Vandermonde's convolution formula. Therefore, $e_j[Z] \le m^3 N^{j-3}$. In view of (3.23) (applied to Z) we obtain (3.17) from this inequality and (3.22).

The proof of the remaining bounds (3.18-20) is similar. We find

$$e_j[Y] = \binom{m}{2}\binom{N-m}{j-2}, \qquad e_j[\mu] = m\binom{N-m}{j-1}, \qquad e_j[W_1] = \binom{N-m}{j}.$$

It follows from these identities that

$$e_j[Y] \le m^2 N^{j-2}, \qquad e_j[\mu] \le m N^{j-1}, \qquad e_j[W_1] \le N^j.$$

Combining these bounds and (3.22) we obtain (3.18-20), thus, completing the proof of the lemma.

Lemma 3.2. *For every $D \subset \Omega_m$, with cardinality $|D| \geq m/4$, and for any $i, j \in \Omega_m \setminus D$ we have*

$$\mathbf{E}(\varkappa_D \,|X_i, X_j) \ll \tau^{-2}\gamma_2, \qquad \mathbf{E}(\varkappa_D^{3/2} \,|X_i, X_j) \ll \tau^{-3}\gamma_3, \tag{3.24}$$
$$\mathbf{E}\varkappa_D \ll \tau^{-2}\gamma_2, \qquad \mathbf{E}\varkappa_D^{3/2} \ll \tau^{-3}\gamma_3,$$

where $\varkappa_D$ is defined in (3.11) and where $m \approx c_2 N u^{-1} \ln u$.

Proof of Lemma 3.2. Clearly, the first two inequalities imply the rest ones. The first inequality is proved in Bloznelis and Götze (2000) (formula (4.22)).

Let us prove the second inequality. An application of the simple inequality $(|D|^{-1} \sum_{r \in D} x_r^2)^{3/2} \leq |D|^{-1} \sum_{r \in D} |x_r|^3$ to $x_r = \zeta(X_r)$ gives

$$\varkappa_D^{3/2} \leq \tau^3 |D|^{-1} \sum_{r \in D} |\zeta(X_r)|^3.$$

Invoking the bound $\mathbf{E}(|\zeta(X_r)|^3 | X_i, X_j) \ll (Npq)^{3/2} \mathbf{E}|g_2(X_r, X_N)|^3$ we complete the proof. In order to prove this latter bound we first apply Rosenthal's inequality to the conditional expectation of $|\zeta(X_r)|^3$ given all the random variables but $\nu_{m+1}, \dots, \nu_N$ and then take the expected value given X_i, X_j and apply the inequalities (4.5) of Bloznelis and Götze (2000), see also (5.4) ibidem. Lemma is proved.

4. Combinatorial lemmas

Here we prove two lemmas. Lemma 4.1 establishes bounds for the second moments of random variables Y_m, Z_m and $\eta_{m,i}$ defined in (2.9) above. Lemma 4.2 provides an auxiliary combinatorial identity.

We first introduce some more notation. For $k \leq n$ write $\Omega_k^c = \Omega_n \setminus \Omega_k$, where $\Omega_k = \{1, \dots, k\}$. Let H be a random variable of the form $H = \sum_{A \subset \mathcal{H}} T_A$, where $\mathcal{H}$ is a class of subsets of Ω_N, $\mathcal{H} = \{A \subset \Omega_N : 1 \leq |A| \leq n\}$. Denote $U_j(H) = \sum_{A \in \mathcal{H}, |A| = j} T_A$ and write $e_j(H) = \mathbf{E} U_j^2(H) \sigma_j^{-2}$, for $\sigma_j^2 > 0$. For $\sigma_j^2 = 0$ put $e_j(H) = 0$. We have

$$H = \sum_{1 \leq j \leq n} U_j(H) \qquad \text{and} \qquad \mathbf{E}H^2 = \sum_{1 \leq j \leq n} \mathbf{E}U_j^2(H) = \sum_{1 \leq j \leq n} e_j(H) \sigma_j^2,$$

where the second identity follows from the fact that T_A and T_B are uncorrelated for $|A| \neq |B|$, see (2.2) above.

For non–negative integers k, s, t, u such that $u \geq \min\{s, t\} + k$ denote

$$r_k(s, t, u) = \sum_{v=0}^{s \wedge t} (-1)^{v+k} \binom{s}{v} \binom{t}{v} \binom{u}{v+k}^{-1}.$$

Here $s \wedge t = \min\{s, t\}$. Recall that for $x \in \mathbb{R}$, $\binom{x}{r} = [x]_r / r!$, if the integer $r \geq 0$, and $\binom{x}{r} = 0$, for $r < 0$. Here $[x]_r = x(x-1) \cdots (x - r + 1)$, for $r > 0$, and $[x]_0 = 1$. In particular, for non-negative integers $s < v$, we have $\binom{s}{v} = 0$

Lemma 4.1. *Assume that* $100 \le n \le N/2$. *Given* $3 \le m \le n$ *and* $i \in \Omega_m$, *we have*

$$\mathbf{E}Y_m^2 \ll n^{-3}m^4\delta_3, \qquad \mathbf{E}Z_m^2 \ll n^{-4}m^6\delta_3, \qquad \mathbf{E}\eta_{m,i}^2 \ll n^{-2}\delta_3. \tag{4.1}$$

Proof of Lemma 4.1. In order to prove (4.1) we show that

$$\mathbf{E}Z_m^2 \ll m^6\mathbf{E}Z_3^2, \qquad \mathbf{E}Y_m^2 \ll n\,m^4\mathbf{E}Z_3^2, \qquad \mathbf{E}\eta_{m,i}^2 \ll n^2\mathbf{E}Z_3^2, \tag{4.2}$$
$$\mathbf{E}Z_3^2 \ll n^{-4}\delta_3. \tag{4.3}$$

Let us show that, for $3 \le j \le n$,

$$e_j(Z_3) = \binom{n-3}{j-3} r_0(j-3, n-j, N-j) = \binom{n-3}{j-3}\binom{N-n}{j-3}\binom{N-j}{j-3}^{-1}. \tag{4.4}$$

By symmetry,

$$\mathbf{E}U_j^2(Z_3) = \binom{n-3}{j-3}\mathbf{E}T_{\Omega_j}U_j(Z_3), \quad \mathbf{E}T_{\Omega_j}U_j(Z_3) = \sum_{v=0}^{(j-3)\wedge(n-j)} M_v s_{j,j-v},$$

where $M_v = \binom{n-j}{v}\binom{j-3}{v}$ counts the summands T_A of the sum $U_j(Z_3)$ which satisfy $|A \cap \Omega_j| = j - v$. Invoking (2.3), we obtain $\mathbf{E}T_{\Omega_j}U_j(Z_3) = r_0(j-3, n-j, N-j)\sigma_j^2$, thus proving the first part of (4.4). The second part follows from (4.19).

Let us prove the first inequality of (4.2). Write $Z_m = Z_3 + D_4 + \cdots + D_m$, where $D_s = Z_s - Z_{s-1}$. By the inequality $(a_1 + \cdots + a_s)^2 \le s(a_1^2 + \cdots + a_s^2)$, we have

$$\mathbf{E}Z_m^2 \le (m-2)(\mathbf{E}Z_3^2 + \mathbf{E}D_4^2 + \cdots + \mathbf{E}D_m^2).$$

Now (4.2) follows from the inequalities $\mathbf{E}D_s^2 \ll s^4\mathbf{E}Z_3^2$, for $4 \le s \le m$. To prove these inequalities we show that, for $3 \le j \le n$,

$$e_j(D_s) \ll s^4 e_j(Z_3), \tag{4.5}$$

Observe, that $e_j(D_s) = 0$, for $n - s + 3 < j \le n$, by the definition of D_s. In the case where $3 \le j \le n-s+3$, the inequalities (4.5) follow from the identities (4.6), (4.7) and the bound (4.9), see below.

We have

$$D_s = \sum_{A \subset \Omega_{s-1},\, |A|=2} \; \sum_{B \subset \Omega_s^c} T_{A \cup \{s\} \cup B}.$$

Given j, this sum has $\binom{s-1}{2}\binom{n-s}{j-3}$ different summands $T_{A\cup\{s\}\cup B}$, such that $|A\cup\{s\}\cup B|=j$. Fix $B_0\subset\Omega_s^c$ with $|B_0|=j-3$ and denote $A_0=\Omega_2\cup\{s\}\cup B_0$. By symmetry

$$\mathbf{E}U_j^2(D_s)=\binom{s-1}{2}\binom{n-s}{j-3}\mathbf{E}T_{A_0}U_j(D_s). \tag{4.6}$$

In the next step we show that

$$\mathbf{E}T_{A_0}U_j(D_s)=\Big(L_0(j)+2(s-3)L_1(j)+\binom{s-3}{2}L_2(j)\Big)\sigma_j^2, \tag{4.7}$$

where we denote $L_i(j)=r_i(j-3,\kappa,N-j)$, where $\kappa=n-s-j+3$. To this aim split $U_j(D_s)=W_0+W_1+W_2$, where

$$W_i=\sum_{A\in\mathcal{A}_i}\ \sum_{B\subset\Omega_s^c,\,|B|=j-3}T_{A\cup\{s\}\cup B},\quad \mathcal{A}_i=\{A\subset\Omega_{s-1}:|A|=2,\,|A\cap\Omega_2|=2-i\},$$

and write $\mathbf{E}T_{A_0}U_j(D_s)=\mathbf{E}T_{A_0}W_0+\mathbf{E}T_{A_0}W_1+\mathbf{E}T_{A_0}W_2$. Note that (4.7) follows from the identities

$$\mathbf{E}T_{A_0}W_i=|\mathcal{A}_i|L_i(j)\sigma_j^2,\quad i=0,1,2, \tag{4.8}$$

which we are going to prove now. Denote $H_0=\{1,2\}$, $H_1=\{1,3\}$, $H_2=\{3,4\}$. By symmetry, for $i=0,1,2$,

$$\mathbf{E}T_{A_0}W_i=|\mathcal{A}_i|\mathbf{E}T_{A_0}\sum_{B\subset\Omega_s^c,\,|B|=j-3}T_{H_i\cup\{s\}\cup B}=|\mathcal{A}_i|\sum_{v=0}^{(j-3)\wedge\kappa}M_v\,s_{j,j-v-i}.$$

Here $M_v=\binom{\kappa}{v}\binom{j-3}{v}$ counts the subsets $B\subset\Omega_s^c$ such that $|B\cap B_0|=j-3-v$. Invoking (2.3) we obtain (4.8).

We complete the proof of (4.5), by showing that, for $3\le j\le n-s+3$,

$$L_i'(j)\ll 1,\qquad \text{where}\qquad L_i'(j):=\binom{n-s}{j-3}\frac{|L_i(j)|}{e_j(Z_3)},\qquad i=0,1,2. \tag{4.9}$$

To evaluate $L_i'(j)$ we use the expression (4.4) for $e_j(Z_3)$ and invoke the formulas (4.21) for $L_i(j)$. For $i=0$ a simple calculation shows that

$$L_0'(j)=\frac{[n-s]_{j-3}}{[N-n]_{j-3}}\ \frac{[N-n+s-3]_{j-3}}{[n-3]_{j-3}}=\prod_{r=0}^{j-4}\frac{x_r}{y_r}\ \frac{y_r+s-3}{x_r+s-3},$$

where we denote $x_r = n - s - r$ and $y_r = N - n - r$. Now the inequality $L_0'(j) \le 1$ follows from the inequalities $x_r \le y_r$, which are consequences of the inequality $n \le N/2$.

For $i = 1, 2$ the proof of (4.9) is similar.

Let us prove the second inequality of (4.2). To this aim we shall show that

$$e_j(Y_m) \ll n\, m^4 e_j(Z_3), \qquad 3 \le j \le n. \tag{4.10}$$

Note that (by the definition of Y_m) $e_j(Y_m) = 0$, for $j > n - m + 2$. Let us prove (4.10), for $3 \le j \le n - m + 2$. Given j, fix $B_0 \subset \Omega_m^c$, with $|B_0| = j - 2$. By symmetry,

$$\mathbf{E}U_j^2(Y_m) = \binom{m}{2}\binom{n-m}{j-2}\mathbf{E}T_{\Omega_2 \cup B_0} U_j(Y_m). \tag{4.11}$$

Proceeding as in the proof of (4.7) above, we obtain

$$\mathbf{E}T_{\Omega_2 \cup B_0} U_j(Y_m) = \Bigl(L_0(j) + 2(m-2)L_1(j) + \binom{m-2}{2} L_2(j)\Bigr)\sigma_j^2, \tag{4.12}$$

where $L_i(j) = r_i(j-2, n-m-j+2, N-j)$, for $i = 0, 1, 2$. Furthermore, arguing as in proof of (4.9) we obtain

$$\binom{n-m}{j-2} L_i(j) \ll n\, e_j(Z_3), \quad \text{for} \quad i = 0, 1, 2. \tag{4.13}$$

Finally, combining (4.11-13) we get $\mathbf{E}U_j^2(Y_m) \ll n\, m^4 e_j(Z_3)\sigma_j^2$. This bound implies (4.10) thus completing the proof of the second inequality of (4.2)

In order to prove the last inequality of (4.2) we shall show that, for $3 \le j \le n$,

$$e_j(\eta_{m,i}) \le n^2 e_j(Z_3). \tag{4.14}$$

Note that (by the definition of $\eta_{m,i}$) $e_j(\eta_{m,i}) = 0$, for $j > n - m + 1$. Let us prove (4.14), for $3 \le j \le n - m + 1$. Given j, fix $B_0 \subset \Omega_m^c$, with $|B_0| = j - 1$. By symmetry,

$$\mathbf{E}U_j^2(\eta_{m,i}) = \binom{n-m}{j-1}\mathbf{E}T_{\{i\} \cup B_0} U_j(\eta_{m,i}). \tag{4.15}$$

Denote $\kappa = n - m - j + 1$. A direct calculation shows that

$$\mathbf{E}T_{\{i\} \cup B_0} U_j(\eta_{m,i}) = \sum_{v=0}^{(j-1)\wedge\kappa} \binom{j-1}{v}\binom{\kappa}{v} s_{j,j-v} = r_0(j-1, \kappa, N-j)\sigma_j^2,$$

where in the last step we invoke (2.3). Furthermore, proceeding as in the proof of (4.9) we obtain

$$\binom{n-m}{j-1} r_0(j-1,\kappa,N-j) \le n^2 e_j(Z_3). \tag{4.16}$$

Now (4.14) follows from (4.15) and (4.16).

Proof of (4.3). Note that the inequality $n \le N/2$ implies $\tau^2 \ge n/2$. Therefore, in order to prove (4.3) it suffices to show that $\mathbf{E} Z_3^2 \ll \mathbf{E}(\mathbb{D}_3 T)^2$. For this purpose we show that

$$e_j(Z_3) \ll e_j(\mathbb{D}_3 T), \qquad 3 \le j \le n. \tag{4.17}$$

We invoke the formula for $\mathbf{E} U_j^2(\mathbb{D}_i T)$ provided in Lemma 2 of Bloznelis and Götze (2001):

$$\mathbf{E} U_j^2(\mathbb{D}_i T) = \frac{\binom{n-i}{j-i}\binom{N-n-i}{j-i}}{\binom{N-i-j}{j-i}} \, \frac{N-j+1}{N-i-j+1}\, 2^i \sigma_j^2.$$

Combining this formula and (4.4) we find $\dfrac{e_j(Z_3)}{e_j(\mathbb{D}_3 T)} = A\,B\,C$ where

$$A = \frac{[N-n]_3}{[N-j]_3}, \quad B = \frac{[N-2j+3]_3}{[N-n-j+3]_3}, \quad C = 2^{-3}\,\frac{N-j-2}{N-j+1}.$$

The inequality (4.17) follows from the inequalities $A \le 1$, $C \le 2^{-3}$ and $B \le 2^3$. The first two inequalities are obvious. In order to show the last one we write $j = n - \varepsilon$ and use the fact that $\varepsilon \ge 0$. The lemma is proved.

In the remaining part of the section we evaluate the coefficients $r_k(s,t,u)$. Using the identity, see Feller (1968), Chapter II,

$$\sum_v (-1)^v \binom{a}{v}\binom{u-v}{t} = \binom{u-a}{u-t}, \qquad a,t,u \in \{0,1,2,\dots\}, \tag{4.18}$$

Zhao and Chen (1990) showed that for $u \ge s \wedge t$ and $s \ge t$

$$r_0(s,t,u) = \binom{u-s}{t}\binom{u}{t}^{-1}. \tag{4.19}$$

Given integers $0 \le s \le t$, let $l(t,s)$ denote the coefficients of the expansion

$$[v+t]_t = l(t,t)[v]_t + l(t,t-1)[v]_{t-1} + \dots + l(t,0)[v]_0. \tag{4.20}$$

Lemma 4.2. *Let $k, s, t, u \in \{0, 1, 2, \dots\}$. For $u \geq s \wedge t + k$, we have*

$$r_k(s,t,u) = \sum_{r=0}^{k} l(k,r)(-1)^{r+k} A_{k,r}, \tag{4.21}$$

where

$$\begin{aligned} A_{k,r} &= 0 && \text{for} \quad r > s \wedge t; \\ A_{k,r} &= 0 && \text{for} \quad u < s + t + k - r; \\ A_{k,r} &= \frac{(u-s-k)!(u-t-k)![s]_r[t]_r}{(u-s-t-k+r)!u!} && \text{otherwise.} \end{aligned}$$

Clearly, the numbers $l(i,j)$ can be expressed by Stirling numbers. A direct calculation shows that

$$l(0,0) = 1; \quad l(1,0) = 1, \quad l(1,1) = 1; \qquad l(2,0) = 2, \quad l(2,1) = 4, \quad l(2,2) = 1;$$
$$l(3,0) = 6, \quad l(3,1) = 18, \quad l(3,2) = 9, \quad l(3,3) = 1.$$

Proof of Lemma 4.2. Write $a = \min\{s,\, t\}$ and $b = \max\{s,\, t\}$. We have

$$r_k(s,t,u) = \sum_{v=0}^{a} (-1)^{v+k} \binom{b}{v} M_v, \quad \text{where} \quad M_v = \binom{a}{v} \binom{u}{v+k}^{-1}. \tag{4.22}$$

A simple calculation shows that

$$M_v = [v+k]_k \binom{u-k-v}{u-k-a} w_k(a,u), \qquad w_k(a,u) = \binom{u-k}{a}^{-1} [u]_k^{-1}.$$

Invoking the expansion (4.20), for the function $v \to [v+k]_k$, we obtain an expression for M_v. Substituting this expression in (4.22) we get

$$r_k(s,t,u) = w_k(a,u) \sum_{r=0}^{k} l(k,r) S_r, \qquad S_r = \sum_{v=0}^{a} (-1)^{v+k} \binom{b}{v} \binom{u-k-v}{u-k-a} [v]_r.$$

We complete the proof of (4.21) by showing that, for $0 \leq r \leq k$,

$$S_r = (-1)^{r+k} [b]_r \binom{u-b-k}{a-r} \quad \text{and} \quad [b]_r \binom{u-b-k}{a-r} w_k(a,u) = A_{k,r}. \tag{4.23}$$

Note that $[v]_r = 0$, for $v < r$. For $r \leq v \leq b$, we have $[v]_r \binom{b}{v} = [b]_r \binom{b-r}{v-r}$. Therefore, denoting $v' = v - r$, we can write

$$S_r = \sum_{v'=0}^{a-r} (-1)^{v'+r+k} [b]_r \binom{b-r}{v'} \binom{u-k-r-v'}{u-k-a}.$$

Finally, invoking (4.18) we obtain the first identity of (4.23). The second identity is trivial.

References

Albers, W., Bickel, P.J. and van Zwet, W.R., *Asymptotic expansions for the power of distribution free tests in the one-sample problem*, Ann. Statist. **4** (1976), 108–156.

Babu, Gutti Jogesh, Bai, Z. D., *Mixtures of global and local Edgeworth expansions and their applications*, J. Multivariate Anal. **59** (1996), 282–307.

Babu, G.J. and Singh, K., *Edgeworth expansions for sampling without replacement from finite populations*, J. Multivar. Analysis. **17** (1985), 261–278.

Bentkus, V., Götze, F. and van Zwet, W. R., *An Edgeworth expansion for symmetric statistics,*, Ann. Statist. **25** (1997), 851–896.

Bertail, P., *Second-order properties of an extrapolated bootstrap without replacement under weak assumptions*, Bernoulli **3** (1997), 149–179.

Bhattacharya, R.N. and Ghosh, J.K., *On the validity of the formal Edgeworth expansion*, Ann. Statist. **6** (1978), 434–451.

Bickel, P.J., *Edgeworth expansions in non parametric statistics*, Ann. Statist. **2** (1974), 1–20.

Bickel, P.J., Götze, F. and van Zwet, W. R., *The Edgeworth expansion for U-statistics of degree two*, Ann. Statist. **14** (1986), 1463–1484.

Bickel, P.J., Götze, F. and van Zwet, W. R., *Resampling fewer than n observations: gains, losses, and remedies for losses*, Statist. Sinica. **7** (1997), 1–31.

Bickel, P.J. and Robinson, J., *Edgeworth expansions and smoothness*, Ann. Probab. **10** (1982), 500–503.

Bickel, P.J. and van Zwet, W. R., *Asymptotic expansions for the power of distribution free tests in the two-sample problem*, Ann. Statist. **6** (1978), 937–1004.

Bikelis, A., *On the estimation of the remainder term in the central limit theorem for samples from finite populations.*, Studia Sci. Math. Hungar. **4** (1972), 345–354. (In Russian).

Bloznelis, M., *A Berry–Esseen bound for finite population Student's statistic*, Ann. Probab. **27** (1999), 2089–2108.

Bloznelis, M. and Götze, F., *An Edgeworth expansion for finite population U-statistics*, Bernoulli **6** (2000), 729–760.

Bloznelis, M. and Götze, F., *One term Edgeworth expansion for finite population U–statistics of degree two*, Acta Applicandae Mathematicae **58** (1999), 75–90.

Bloznelis, M. and Götze, F., *Orthogonal decomposition of finite population statistic and its applications to distributional asymptotics*, Ann. Statist. **29** (2001).

Bolthausen, E. and Götze, F., *The rate of convergence for multivariate sampling statistics*, Ann. Statist. **21** (1993), 1692–1710.

Booth, J. G. and Hall, P., *An improvement of the jackknife distribution function estimator*, Ann. Statist. **21** (1993), 1476–1485.

Callaert, H., Janssen, P., *A Berry–Esseen theorm for U–statistics*, Ann. Statist. **6** (1978), 417–421.

Callaert, H., Janssen, P. and Veraverbeke, N., *An Edgeworth expansion for U-statistics*, Ann. Statist. **8** (1980), 299–312.

Chibisov, D.M., *Asymptotic expansion for the distribution of a statistic admitting a stochastic expansion*, Theory Probab. Appl. **17** (1972), 620–630.

Erdős, P. and Rényi, A., *On the central limit theorem for samples from a finite population*, Publ. Math. Inst. Hungar. Acad. Sci. **4** (1959), 49–61.

Feller, W., *An introduction to probability theory and its applications. I*, Wiley, New York, 1968.

Götze, F., *Asymptotic expansions for bivariate von Mises functionals*, Z. Wahrsch. verw. Gebiete **50** (1979), 333–355.

Götze, F. and van Zwet, W.R., *Edgeworth expansions for asymptotically linear statistics* (1991), Preprint 91 - 034, SFB 343, Universität Bielefeld.

Helmers, R., *Edgeworth expansions for linear combinations of order statistics* **105** (1982), Math. Centre Tracts, Math. Centrum, Amsterdam.

Helmers, R., and van Zwet, W. R., *The Berry–Esseen bound for U-statistics.* **1** (1982), Statistical Decision Theory and Related Topics,III. Vol. 1. (S.S. Gupta and J.O. Berger, eds.), Academic Press, New York, 497–512.

Höglund, T., *Sampling from a finite population: a remainder term estimate*, Scand. J. Stat. **5** (1978), 69–71.

Kokic, P.N. and Weber, N.C., *An Edgeworth expansion for U-statistics based on samples from finite populations*, Ann. Probab. **18** (1990), 390–404.

Lai, T. L. and Wang, J.Q., *Edgeworth expansions for symmetric statistics with applications to bootstrap methods*, Statist. Sinica **3** (1993), 517–542.

Nandi, H. K. and Sen, P.K., *On the properties of U-statistics when the observations are not independent.II. Unbiased estimation of the parameters of a finite population.*, Calcutta Statist. Assoc. Bull. **12** (1963), 124–148.

Petrov, V.V., *Sums of independent random variables*, Springer, New York (1975).

Pfanzagl, J., *Asymptotic expansions related to minimum contrast estimators*, Ann. Statist. **1** (1973), 993–1026.

Prawitz, H., *Limits for a distribution, if the characteristic function is given in a finite domain*, Scand. AktuarTidskr. (1972), 138–154.

Politis, D. N. and Romano, J. P., *Large sample confidence regions based on subsamples under minimal assumptions*, Ann. Statist. **22** (1994), 2031–2050.

Rao, C. R., Zhao, L. C., *Berry-Esseen bounds for finite-population t-statistics*, Statist. Probab. Lett. **21** (1994), 409–416.

Robinson, J., *An asymptotic expansion for samples from a finite population*, Ann. Statist. **6** (1978), 1005–1011.

Schneller, W., *Edgeworth expansions for linear rank statistics*, Ann. Statist. **17** (1989), 1103–1123.

Serfling, R. J., *Approximation Theorems of Mathematical Statistics*, Wiley, New York, 1980.

Shao, J., *The efficiency and consistency of approximations to the jackknife variance estimators*, J. Amer. Statist. Assoc. **84** (1989), 114–119.

Wu, C.F.J., *On the asymptotic properties of the jackknife histogram*, Ann. Statist. **18** (1990), 1438–1452.

Zhao, L. C. and Chen, X. R., *Berry–Esseen bounds for finite-population U-statistics*, Sci.Sinica Ser. A **30** (1987), 113–127.

Zhao, L. C. and Chen, X. R., *Normal approximation for finite-population U-statistics*, Acta Math. Appl. Sinica **6** (1990), 263–272.

van Zwet, W.R., *A Berry-Esseen bound for symmetric statistics*, Z. Wahrsch. verw. Gebiete **66** (1984), 425–440.

Mindaugas Bloznelis, Faculty of Mathematics and Informatics,
Vilnius University, Naugarduko 24, Vilnius 2006, Lithuania
E-mail address: Mindaugas.Bloznelis@maf.vu.lt

Friedrich Gotze, Fakultat fur Mathematik, Universitat Bielefeld,
Postfach 100131, Bielefeld 33501, Germany
E-mail address: goetze@mathematik.uni-bielefeld.de