AN EDGEWORTH EXPANSION FOR FINITE POPULATION U-STATISTICS

M. Bloznelis[1] F. Götze[1]

Vilnius University and IMI
University of Bielefeld

January 1997

Abstract. Suppose that U is a U-statistic of degree two based on N random observations drawn without replacement from a finite population. For the distribution of a standardized version of U we construct an Edgeworth expansion with remainder $O(N^{-1})$ provided that the linear part of the statistic satisfies a Cramér type condition.

1. Introduction and results

Let $\mathcal{A}$ denote a population of n elements $a_1, \dots, a_n$ and let $\mathcal{H} : \mathcal{A} \times \mathcal{A} \to R$ denote symmetric function of its two arguments. By $X_1, \dots, X_N$, $N \le n$ we denote random variables with values in $\mathcal{A}$ such that $X = \{X_1, \dots, X_N\}$ represents a random sample from $\mathcal{A}$ of size N drawn without replacement, i.e., $\mathbf{P}\{X \in B\} = \binom{n}{N}^{-1}$ for any subset B of $\mathcal{A}$ of size N. Assume that $\mathbf{E}\,\mathcal{H}(X_1, X_2) = 0$.

We shall investigate the asymptotic distribution of the statistic

$$U = \sum_{1 \le i < j \le N} \mathcal{H}(X_i, X_j).$$

We shall decompose the statistic as

$$U = L + Q, \tag{1.1}$$

---

[1]Research supported by the SFB 343 in Bielefeld.

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

Key words and phrases. Edgeworth expansion, finite population, U-statistic, Central Limit Theorem, sampling without replacement.

where

$$L = \sum_{i=1}^{N} g_1(X_i) \qquad \text{respectively} \qquad Q = \sum_{1 \le i < j \le N} g_2(X_i, X_j)$$

is the linear respectively the quadratic part of the statistic. Here

$$g_1(x) = (N-1)\, t(x), \qquad t(x) = \frac{n-1}{n-2}\, \mathbf{E}\big(\mathcal{H}(X_1,\, X_2) \,|\, X_1 = x\big)$$

and

$$g_2(x_1,\ x_2) = \mathcal{H}(x_1,\, x_2) - t(x_1) - t(x_2).$$

Since

$$\mathbf{E}\big(g_2(X_1,\, X_2) \,|\, X_1 = x\big) = 0, \qquad \forall x \in \mathcal{A}, \tag{1.2}$$

the random variables $g_1(X_i)$ and $g_2(X_j,\, X_k)$, $1 \le i \le N$, $1 \le j < k \le N$, (and thus the parts L and Q) are uncorrelated.

If the linear part dominates the statistic for large N, the distribution of U can be approximated by a Gaussian distribution using the Central Limit Theorem (CLT).

The asymptotic normality of linear statistics based on samples drawn without replacement from finite populations has been studied by a number of authors. Erdös and Rényi (1959) proved the CLT under very mild conditions. The rate of convergence in the CLT was first studied by Bikelis (1969). Berry–Esseen bounds of order $O(N^{-1/2})$ were obtained by Höglund (1978). Robinson (1978) proved the validity of an Edgeworth expansion with a remainder of order $O(N^{-3/2})$, see also Bickel and van Zwet (1978).

Nandi and Sen (1963) studied the behavior of finite population U-statistics and showed that under proper regularity conditions the sequence of distributions of normalized U-statistics converges to the standard normal distribution. The rate of this convergence was investigated by Zhao and Chen (1987, 1990), Kokic and Weber (1990, 1991) and, as a particular case of the rate of convergence of general multivariate sampling statistics, by Bolthausen and Götze (1993). In contrast to the case of independent and identically distributed observations where the second order asymptotic theory has been developed for U-statistics, see Bickel (1974), Götze (1979), Callaert, Janssen and Veraverbeke (1980), Bickel, Götze and van Zwet (1986), and for more general asymptotically normal symmetric statistics, see Bentkus, Götze and van Zwet (1994), there are only a few results concerned with higher order asymptotics of nonlinear finite population statistics. Babu and Singh (1985) proved the validity of an Edgeworth expansion with a remainder $o(N^{-1/2})$ for second order polynomials of variables $X_1, \dots, X_N$. Kokic and Weber (1990) established one term Edgeworth expansion with the remainder $o(N^{-1/2})$ for finite population U-statistics of degree 2.

Assume that

$$\sigma^2 = N\,\mathbf{E} g_1^2(X_1) > 0.$$

The distribution function of the standardized statistic, $F(x) = \mathbf{P}\{U \le x\,\sigma\}$, will be approximated by the one term Edgeworth expansion,

$$G(x) = \Phi\Big(\frac{x}{\sqrt{q}}\Big) - \frac{(q-p)\,q^{-1/2}\,\alpha + 3\,q^{1/2}\,\kappa}{6\,\sigma^3\,N^{1/2}}\,\Phi'''\Big(\frac{x}{\sqrt{q}}\Big). \tag{1.3}$$

Here $\Phi(x)$ is the standard normal distribution function,

$$p = N/n, \qquad q = 1 - p$$

and

$$\alpha = N^{3/2}\,\mathbf{E} g_1^3(X_1), \qquad \kappa = N^{5/2}\,\mathbf{E} g_2(X_1, X_2)\,g_1(X_1)\,g_1(X_2). \tag{1.4}$$

We shall derive bounds for the remainder,

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

In the classical case of standardized sums $S = (Y_1 + \cdots + Y_N)/\sqrt{N}$ of independent and identically distributed (i.i.d.) random variables $Y_1, \dots, Y_N$, such that $\mathbf{E}\,Y_1 = 0$, $\mathbf{E}\,Y_1^2 = 1$ and $\mathbf{E}\,Y_1^4 < \infty$ asymptotic expansions for the distribution F_S of S with the remainder $O(N^{-1})$ are obtained assuming Cramer's condition (C),

$$\sup_{|t|>a} \big|\,\mathbf{E}\,\exp\{i\,t\,Y_1\}\,\big| < 1. \tag{C}$$

Bentkus, Götze and van Zwet (1994), [BGZ] for short, introduced a local version of Cramer's condition (C), namely,

$$\rho_{Y_1}(a, b) = 1 - \sup_{a \le |t| \le b} \big|\mathbf{E}\exp\{i\,t\,Y_1\}\big| > 0. \tag{C'}$$

Condition (C') (with $a = 1/\mathbf{E}\,|Y_1|^3$ and $b = N^{1/2}$) is somewhat weaker than (C) but still sufficient to prove the validity of Edgeworth expansions for F_S up to an order $O(N^{-1})$. This modification is useful in more general situations, where Y_1 depends on N in an implicit way, see [BGZ].

We shall assume that the random variable $Z = \sqrt{N}\,g_1(X_1)/\sigma$ satisfies condition (C') with $a = b_1/\mathbf{E}\,|Z|^3$ and $b = N^{1/2}$, i.e.,

$$\rho = \rho_Z(a,\, b) > 0. \tag{1.5}$$

Here b_1 is a sufficiently small absolute constant. One may take, e.g., $b_1 = 0.00144$. Write

$$\beta_r = \mathbf{E}\,\big|N^{1/2}\,g_1(X_1)\big|^r \quad \text{and} \quad \gamma_r = \mathbf{E}\,\big|N^{3/2}\,g_2(X_1, X_2)\big|^r, \quad r = 1, 2, \dots\,. \tag{1.6}$$

The following estimate on the remainder Δ holds.

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

$$\sup_{x\in R}\big|F(x) - G(x)\big| \le \frac{A}{N}\,\frac{\beta_4+\gamma_4}{\rho^2\,q^2\,\sigma^4}\,.$$

For linear statistics we obtain the following result.

Theorem 1.2. *There exists an absolute constant $B > 0$ such that*

$$\Big|\mathbf{P}\{L \le x\} - \Phi\big(\frac{x}{\sqrt{q}}\big) + \frac{(q-p)\,q^{-1/2}\,\alpha}{6\,\sigma^3\,N^{1/2}}\,\Phi'''\big(\frac{x}{\sqrt{q}}\big)\Big| \le \frac{B}{N}\,\frac{\beta_4}{\rho^2\,q\,\sigma^4}\,.$$

The estimates in Theorems 1.1 and 1.2 hold for any fixed sample size N, population size n and functions $\mathcal{H}$. If β_4/σ^4 and γ_4/σ^4 are bounded and q and ρ are bounded away from 0 as $N \to \infty$ and $n \to \infty$, then these results establish Edgeworth expansions with remainders $O(N^{-1})$.

The case where $n \to \infty$ and N is fixed corresponds to the i.i.d. situation. By the law of large numbers we obtain a corollary for independent observations:

Let $\mathcal{E}$ denote a measurable space and let $\mathcal{X}_1, \mathcal{X}_2, \dots$ be i.i.d. random variables with values in $\mathcal{E}$. Write

$$\widehat{U} = \sum_{1\le i<j\le N} \mathcal{H}(\mathcal{X}_i,\, \mathcal{X}_j).$$

Here $\mathcal{H} : \mathcal{E}\times\mathcal{E} \to R$ denotes a measurable function symmetric in its two arguments such that $\mathbf{E}\,\mathcal{H}^2(\mathcal{X}_1, \mathcal{X}_2) < \infty$. Again we assume $\mathbf{E}\,\widehat{U} = 0$ and decompose

$$\widehat{U} = \sum_{i=1}^{N} \widehat{g}_1(\mathcal{X}_i) + \sum_{1\le i<j\le N} \widehat{g}_2(\mathcal{X}_i, \mathcal{X}_j).$$

Here $\widehat{g}_1$ and $\widehat{g}_2$ are defined in the same way as g_1 and g_2, but by means of $\widehat{t}(x) = \mathbf{E}\big(\mathcal{H}(\mathcal{X}_1,\, \mathcal{X}_2)\,|\,\mathcal{X}_1 = x\big)$ instead of $t(x)$. Let $\widehat{\sigma}$, $\widehat{\alpha}$, $\widehat{\beta}_k$, $\widehat{\gamma}_k$, $k = 2, 3, 4$, and $\widehat{\kappa}$ denote the moments of $\widehat{g}_1(\mathcal{X}_1)$ and $\widehat{g}_2(\mathcal{X}_1, \mathcal{X}_2)$ corresponding to σ, α, β_k, γ_k and κ. We shall assume that

$$\widehat{\rho} = \rho_{\mathcal{Z}}\big(\widehat{a}, \widehat{b}\big) > 0, \qquad \mathcal{Z} = \sqrt{N}\,\widehat{g}_1(\mathcal{X}_1)/\widehat{\sigma},$$

where $\widehat{a} = b_1/\mathbf{E}\,|\mathcal{Z}|^3$ and $\widehat{b} = \sqrt{N}$. Then we have

Corollary 1.3. *There exists an absolute constant $A > 0$ such that*

$$\Big|\mathbf{P}\{\widehat{U} \le \hat{\sigma}x\} - \Phi(x) + \frac{\hat{\alpha} - 3\hat{\kappa}}{6\,\hat{\sigma}^3\,N^{1/2}}\,\Phi'''(x)\Big| \le \frac{A}{N}\,\frac{\hat{\beta}_4+\hat{\gamma}_4}{\hat{\rho}^2\,\hat{\sigma}^4}\,.$$

Hence, Theorem 1.1 which yields this result as a special case may be regarded as a partial extension of the result of Bentkus, Götze and van Zwet (1994) to sampling models. These authors proved the validity of an Edgeworth expansion with the remainder $O(N^{-1})$ for general symmetric asymptotically normal statistics based on i.i.d. observations. In the case of U-statistics of degree two their result yields the estimate as in Corollary 1.3 but with the lower moment $\hat{\gamma}_3/\hat{\sigma}^3$ instead of $\hat{\gamma}_4/\hat{\sigma}^4$ in the remainder.

An example given in Theorem 1.4 in [BGZ] shows that a Cramér type condition on the linear part and the existence of moments of arbitrarily high order of the linear and quadratic parts of the statistic are not sufficient to obtain an estimate similar to that in Corollary 1.3 for higher order Edgeworth expansions (up to an order $o(N^{-1})$). Hence, in this sense Corollary 1.3 and thus Theorem 1.1 are best possible.

Let us compare our results with those of Robinson (1978) and Kokic and Weber (1990). Robinson (1978) proved the validity of a two term Edgeworth expansion with the remainder $O(N^{-3/2})$ for linear statistics like L in (1.1) assuming the following Cramér type condition. This condition, first used in Albers, Bickel and van Zwet (1976), requires for a random variable Z that there exists an $\varepsilon > 0$ such that

$$\tau_Z(\varepsilon, a, b) = 1 - \sup_{s \in R,\, a \le |t| \le b} \mathbf{P}\big\{t\, Z \in L^{\varepsilon} + s\big\} > 0. \tag{c}$$

Here $L = \{2\,\pi\, r,\ r = 0, \pm 1, \pm 2, \dots\}$ and $\mathcal{B}^{\varepsilon}$ denotes the ε neighborhood of a set $\mathcal{B} \subset R$. Notice that $\varepsilon_1 \le \varepsilon_2$ implies $\tau_Z(\varepsilon_1, a, b) \ge \tau_Z(\varepsilon_2, a, b)$. Robinson assumed that given $C' > 0$

$$\text{there exist} \quad \varepsilon,\, \delta > 0 \quad \text{and} \quad C > 0 \quad \text{such that} \quad \tau_Z(\varepsilon, a, b) > \delta, \tag{1.7}$$

for

$$Z = \sqrt{N}\, g_1(X_1)/\sigma, \quad a^{-1} = \max_{1 \le i \le n} |z_i|/C' \quad \text{and} \quad b^{-1} = p\, \mathbf{E}\, |Z|^5/(C\, N).$$

Here $\{z_1, \dots, z_n\}$ denotes the set of values of the r.v. Z. For a sequence of finite population linear statistics, say (L_n), Robinson's (1978) Theorem establish an Edgeworth expansion with the remainder $O(N^{-3/2})$ provided that β_5/σ^5 is bounded, p and q are bounded away from 0 and (1.7) holds with ε, δ and C not depending on n as $n \to \infty$. Robinson's (1978) result was used by Kokic and Weber (1990) to show $\Delta = o(N^{-1/2})$. Bounds for remainders in these papers involve constants (correspondent to A and B of the present paper) which implicitly depend on p.

In Section 2 we compare conditions (c) and (1.5). It is shown there that these conditions are, in fact, equivalent. Proofs of the Theorems 1.1, 1.2 and the Corollary 1.3 are given in Sections 3 and 4. Our approach can be regarded as an extension of that given in [BGZ] to the case where the sample is drawn without replacement from a finite population. Basic devices used in proofs are: the data depending smoothing technique, developed in [BGZ], and the Erdös and Rényi formula for characteristic functions (Erdös and Rényi (1959)). Auxiliary results are collected in Section 5.

Acknowledgment. The authors would like to thank V. Bentkus for useful discussions and comments. I particular, the smoothness condition (c^*) in Section 2 below was suggested by V. Bentkus.

2. Smoothness conditions

Let Z_n, $n = 1, 2, \ldots$, be a sequence of random variables. Let (a_n) and (b_n) denote sequences of positive numbers. We shall show that the condition

$$\text{(2.1)} \qquad \text{there exists an} \quad \varepsilon > 0 \quad \text{such that} \quad \liminf_n \tau_{Z_n}(\varepsilon, a_n, b_n) > 0$$

implies

$$\text{(2.2)} \qquad \liminf_n \rho_{Z_n}(a_n, b_n) > 0.$$

Furthermore, (2.2) implies (2.1). More specifically, we show in Lemma 2.1 that given a random variable Z and numbers $0 < a < b$, (1.5) implies $\tau_Z(\rho, a, b) > \rho/4$ and if (c) holds for some $\varepsilon > 0$, then $\rho_Z(a, b) > \varepsilon^2 \tau_Z(\varepsilon, a, b)/\pi^2$. Hence, conditions (2.1) and (2.2) as well as (c) and (1.5) are equivalent.

In order to check condition (c) one needs to minimize a bivariate function over the set $(s, t) \in [-\pi, \pi] \times \{a \le |t| \le b\}$. Such a minimization problem can be difficult to solve numerically. A symmetrization argument suggests a version of condition (c) which is easier to check. Let Z' denote an independent copy of Z. Let $Z^* = Z - Z'$ denote a symmetrization of Z. The condition

$$(c^*) \quad \text{there exists} \quad \varepsilon > 0 \quad \text{such that} \quad \tau^*_Z(\varepsilon, a, b) = 1 - \sup_{a \le |t| \le b} \mathbf{P}\big\{ t\, Z^* \in L^\varepsilon \big\} > 0$$

requires to estimate the minimum of an univariate function only. Notice that $\varepsilon_1 < \varepsilon_2$ implies $\tau^*_Z(\varepsilon_2, a, b) \le \tau^*_Z(\varepsilon_1, a, b)$. We show in Lemma 1 that conditions (c^*) and (c) are, in fact, equivalent.

Modifications of Cramer's condition (C) that ensure the validity of Edgeworth expansions for sums of random variables assuming a finite number of values only were considered by Albers, Bickel and van Zwet (1976), van Zwet (1982), Does (1983), Schneller (1989), see also Bickel and Robinson (1982).

Lemma 2.1. *Let Z be a random variable. For $0<a<b$ and $0<\varepsilon<\pi$ write*

$$\rho=\rho_Z(a,b),\quad \tau_\varepsilon=\tau_Z(\varepsilon,a,b),\quad \tau^*_\varepsilon=\tau^*_Z(\varepsilon,a,b),\quad u=\pi^{-1}\varepsilon\,\tau^*_\varepsilon \quad and \quad v=\pi^{-1}\varepsilon\,\tau_\varepsilon.$$

We have
i) if $\tau_\varepsilon>0$ then $\rho>\pi^{-2}\varepsilon^2\tau_\varepsilon$ and if $\rho>0$ then $\tau_\rho>\rho/4$;
ii) if $\tau^*_\varepsilon>0$ then $\rho>\pi^{-2}\,\varepsilon^2\tau^*_\varepsilon$ and if $\rho>0$ then $\tau^*_\rho>\rho/4$;
iii) if $\tau_\varepsilon>0$ then $\tau^*_v>2^{-1}\,\pi^{-2}\,\varepsilon^2\,\tau_\varepsilon$ and if $\tau^*_\varepsilon>0$ then $\tau_u>4^{-1}\,\pi^{-2}\,\varepsilon^2\,\tau^*_\varepsilon$.

Proof of Lemma 1. For a random variable Y write $\tau_\delta(Y)=\mathbf{P}\{\cos Y<\cos\delta\}$, $\delta>0$. Clearly, $\tau_{\delta_1}(Y)\le\tau_{\delta_2}(Y)$, for $0<\delta_2<\delta_1<\pi$. We shall show that for any $0<\delta<\pi$ and $0<d<1$,

$$\mathbf{E}\,(1-\cos\,Y)\ge 2\,\pi^{-2}\delta^2\,\tau_\delta(Y), \tag{2.3}$$

$$\tau_{d'}(Y)>\frac{1}{2}\Big(\mathbf{E}\,(1-\cos\,Y)-d\Big),\quad d'=\sqrt{2\,d}. \tag{2.4}$$

The first inequality follows from Chebyshev's inequality,

$$\mathbf{E}\,(1-\cos\,Y)\ge(1-\cos\,\delta)\,\mathbf{P}\{1-\cos\,Y>1-\cos\,\delta\},$$

the identity $\tau_\delta(Y)=\mathbf{P}\{1-\cos\,Y>1-\cos\,\delta\}$ and the inequality

$$1-\cos\,u\ge 2\,\pi^{-2}\,u^2,\qquad 0\le u\le\pi. \tag{2.5}$$

To get (2.4) combine the inequality $\arccos(1-d)>\sqrt{2\,d}$ and

$$\begin{aligned}\mathbf{E}\,(1-\cos\,Y)&\le d\,\mathbf{P}\{1-\cos\,Y\le d\}+2\,\mathbf{P}\{1-\cos\,Y>d\}\\&=d+(2-d)\,\mathbf{P}\{\cos\,Y<1-d\}.\end{aligned}$$

Write $Y_{s,t}=t\,Z+s$. We have

$$\tau_\varepsilon=\inf\{\tau_\varepsilon(Y_{s,t}):\ s\in R,\ a<|t|<b\}\quad\text{and}\quad \tau^*_\varepsilon=\inf\{\tau_\varepsilon(t\,Z^*):\ a<|t|<b\}.$$

Recall that $Z^*=Z-Z'$, where Z' denotes a random variable independent of Z and having the same distributions as Z. Put

$$\varkappa=\sup\Big\{\,\mathbf{E}\,\cos Y_{s,t}:\ s\in R,\ a\le|t|\le b\Big\},\quad \varkappa^*=\sup\Big\{\,\mathbf{E}\,\cos\big(t\,Z^*\big):\ a\le|t|\le b\Big\}$$

and $\tilde\varkappa=\sup\Big\{\big|\mathbf{E}\,\exp\{i\,t\,Z\}\big|\ :\ a\le|t|\le b\Big\}$. Notice, that $\varkappa^*=\tilde\varkappa^2$ and $0\le\varkappa,\varkappa^*,\tilde\varkappa\le 1$. Furthermore,

$$\varkappa^*\le\varkappa,\qquad \varkappa^2\le\varkappa^*. \tag{2.6}$$

The first inequality is obvious. The second one follows from the inequality

$$\Big(\mathbf{E}\,\cos(t\,Z+s)\Big)^2 \le \mathbf{E}\,\cos\Big((t\,Z+s)-(t\,Z'+s)\Big)$$

which is a consequence of $\cos(x-y)=\cos x\,\cos y+\sin x\,\sin y$ with $x=t\,Z+s$, $y=t\,Z'+s$ and the independence of x and y.

Let us show i). Assume that $\tau_\varepsilon>0$. It follows from (2.3) that

$$\mathbf{E}\,\cos Y_{s,t}\le 1-2\,\pi^{-2}\varepsilon^2\,\tau_\varepsilon(Y_{s,t})\le 1-2\,\pi^{-2}\varepsilon^2\,\tau_\varepsilon.$$

Hence,

$$1-\varkappa\ge 2\,\pi^{-2}\varepsilon^2\,\tau_\varepsilon. \tag{2.7}$$

By (2.6), $1-\tilde\varkappa^2\ge 1-\varkappa$. Therefore, $\rho=1-\tilde\varkappa=(1-\tilde\varkappa^2)/(1+\tilde\varkappa)\ge(1-\varkappa)/2$ and we obtain $\rho>\pi^{-2}\varepsilon^2\,\tau_\varepsilon$. Let us prove the second implication of i). Assume that $\rho>0$. By (2.6), $\varkappa\le\tilde\varkappa$. Hence, $\rho=1-\tilde\varkappa\le 1-\varkappa$. An application of (2.4) gives

$$\tau_{d'}(Y_{s,t})\ge\frac12\Big(1-\mathbf{E}\,\cos Y_{s,t}-d\Big)\ge\frac12\,(1-\varkappa-d).$$

Hence,

$$\tau_{d'}\ge\frac12\,(1-\varkappa-d),\qquad d'=\sqrt{2\,d}. \tag{2.8}$$

Taking $d=\rho/2$ we obtain $\tau_{d'}\ge\rho/4$ and using $d'=\rho^{1/2}\ge\rho$ we get $\tau_\rho\ge\rho/4$.

The proof of ii) is similar. Assume that $\tau^*_\varepsilon>0$. Using (2.3) with $Y=t\,Z^*$ we get

$$1-\varkappa^*\ge 2\,\pi^{-2}\,\varepsilon^2\,\tau^*_\varepsilon. \tag{2.9}$$

Since $\varkappa^*=\tilde\varkappa^2$, we have $1-\varkappa^*\le 2\,\rho$. Therefore, $\rho\ge\pi^{-2}\,\varepsilon^2\,\tau^*_\varepsilon$. Now assume that $\rho>0$. We have $\rho\le 1-\tilde\varkappa^2=1-\varkappa^*$. An application of (2.4) with $Y=t\,Z^*$ and $d=\rho/2$ gives $\tau^*_{d'}\ge\rho/4$, where $d'=\rho^{1/2}$. It follows that $\tau^*_\rho\ge\rho/4$.

Let us prove iii). Assume that $\tau_\varepsilon>0$. Put $d=\pi^{-2}\,\varepsilon^2\,\tau_\varepsilon$. It follows from (2.7) and the inequality $\varkappa^*\le\varkappa$, see (2.6), that

$$1-\varkappa^*\ge 2\,\pi^{-2}\,\varepsilon^2\,\tau_\varepsilon=2\,d.$$

Invoking (2.4) with $Y=t\,Z^*$ we get

$$2d\le 1-\varkappa^*\le 1-\mathbf{E}\,\cos Y\le d+2\tau_{d'}(Y),\quad d'=\sqrt{2\,d}.$$

It follows that $\tau^*_{d'}\ge d/2=2^{-1}\pi^{-2}\varepsilon^2\tau_\varepsilon$. Since $d'\ge d^{1/2}\ge v$ we obtain $\tau^*_v\ge 2^{-1}\,\pi^{-2}\varepsilon^2\tau_\varepsilon$. The proof of the second implication of iii) is similar. We use (2.9) and inequalities $1-\varkappa\ge(1-\varkappa^2)/2\ge(1-\varkappa^*)/2$, see (2.6), to get $1-\varkappa\ge 2d$, $d=2^{-1}\pi^{-2}\varepsilon^2\tau^*_\varepsilon$. Now an application of (2.8) gives $\tau_{d'}\ge d/2$, where $d'=\sqrt{2\,d}$. Since $d'\ge u$ we obtain $\tau_u\ge\tau_{d'}\ge 4^{-1}\pi^{-2}\varepsilon^2\,\tau^*_\varepsilon$ thus completing the proof.

3. Proofs

3.1. Notation. Throughout the rest of the paper we shall use the following notation.

By C and c we denote generic absolute constants. We shall write C respectively c to denote a "sufficiently large" respectively "sufficiently small" absolute constant. We shall write $A \ll B$ if $A < C\,B$. The expression $\exp\{i\,x\}$ will be abbreviated by $\mathrm{e}\{x\}$. Write

$$\tau = \sqrt{n\,p\,q}, \quad b_1 = 0.00144, \quad b_2 = 0.075, \quad H_1 = b_1\,N^{1/2}/\beta_3$$

and write $\mathcal{K} = \big\{k :\ H_1\,|g_1(a_k)| \ge b_2\big\}$. Let us mention in advance, that our choice of the constants b_1 and b_2 ensures the validity of inequality (4.26) in Section 4 below (in fact this inequality holds if b_2 and b_1/b_2 are sufficiently small positive numbers). Introduce the function, cf. Höglund (1978),

$$\Theta(t) = \Big(\frac{2}{\pi}\ \frac{\pi - t}{\pi + t}\Big)^2.$$

By $\overline{A} = \big(A_1, A_2, \ldots, A_n\big)$ we shall denote a random permutation which is uniformly distributed on the permutations of the population $\big\{\,a_1, a_2, \ldots, a_n\,\big\}$, $\mathbf{E}^*$ denotes the conditional expectation given $\overline{A}$. Furthermore, $\mathbf{E}^{i_1,\ldots,i_k}$ denotes the conditional expectation given $A_{i_1}, \ldots, A_{i_k}$.

3.2. Proofs of Theorems 1.1, 1.2, and Corollary 1.3. The proof of Theorem 1.1 is close to that of given in [BGZ] where a similar result was proved for more general symmetric statistics based on i.i.d. observations. In the proof we use an inversion formula (smoothing lemma) and expansions for characteristic functions. The proof consists of two parts. In the first part we apply the smoothing lemma due to Prawitz (1972) conditionally, cf. "the data dependent smoothing" introduced in [BGZ], and bound the characteristic function $\hat{F}(t)$ of F for large values of t, $N^{1/2}\,\beta_3^{-1} \ll |t| \ll N$. In the second part we estimate the difference $\hat{F}(t) - \hat{G}(t)$ in the region $|t| \ll N^{1/2}\,\beta_3^{-1}$, where $\hat{G}$ denotes the Fourier transform of the function G. Here we expand $\hat{F}(t)$ in powers of t and then use Lindeberg's method, cf. [BGZ] and Höglund (1978).

Proof of Theorem 1.1. We shall assume without loss of generality, that $\sigma^2 = 1$ (equivalently, $\beta_2 = 1$). Write

$$\delta = 1 - \sup\big\{\mathbf{E}\,\cos\big(t\,N^{1/2}\,g_1(X_1) + s\big) :\ s \in R, a \le |t| \le b\big\}, \quad a = b_1/\beta_3, \quad b = N^{1/2}.$$

It is easy to show that $\delta \ge \rho$, see, e.g., the proof of the statement i) of Lemma 2.1.

We may and shall assume, that for sufficiently small but absolute constant $c_0 > 0$,

$$\text{(3.1)} \qquad q^{-1}\beta_4/N < c_0, \quad \delta^{-2}\,q^{-2}\,\gamma_2/N \le c_0, \quad \delta^{-1}\,q^{-1}\,n^{-1}\,\ln N < c_0.$$

Indeed, if the (3.1) fails, then the theorem follows from the inequalities $F(x) \le 1$ and $|G(x)| \ll 1 + q^{-1/2}\beta_4^{1/2}/N^{1/2} + q^{1/2}\gamma_2/N^{1/2}$.

Let m be an integer approximately equal to $C\delta^{-1}\ln N$, where $C > 0$ is a sufficiently large absolute constant. Put $\Lambda = \sum_{1\le k<l\le m} g_2(X_k, X_l)$, $U_1 = U - \Lambda$ and let F_1 denote the probability distribution of U_1. Then

$$\Delta \le \Delta_1 + \mathbf{P}\bigl\{|\Lambda| \ge \delta^{-3/2} N^{-1}\bigr\} + \delta^{-3/2} N^{-1} \max_x \bigl|G'(x)\bigr|, \quad \Delta_1 = \sup_x \bigl|F_1(x) - G(x)\bigr|.$$

By Chebyshev's inequality and the estimate $\mathbf{E}|\Lambda|^3 \le m^6\,\mathbf{E}\,|g_2(X_1, X_2)|^3$,

$$\mathbf{P}\{|\Lambda| \ge \delta^{-3/2}N^{-1}\} \le \delta^{9/2}N^3\mathbf{E}\,|\Lambda|^3 \le \delta^{-3/2}\,\gamma_3\,N^{-3/2}\,\ln^6 N.$$

Furthermore, using $|G'(x)| \ll \beta_4/q + \gamma_2$ we obtain

$$\Delta \ll \Delta_1 + N^{-1}\,\delta^{-3/2}\,\bigl(\beta_4/q + \gamma_2 + \gamma_3\bigr). \tag{3.2}$$

Hence, it remains to bound Δ_1.

Data dependent smoothing. Let k be an integer approximately equal to $(N + m)/2$. Put $\mathcal{I}_0 = \{m + 1, \dots, N\}$, $\mathcal{J}_0 = \{1, \dots, n\} \setminus \mathcal{I}_0$, $\mathcal{J}_1 = \mathcal{J}_0 \cup \bigl\{m + 1, \dots, k\bigr\}$ and $\mathcal{J}_2 = \mathcal{J}_0 \cup \bigl\{k + 1, \dots, N\bigr\}$. Define (random) subpopulations $\mathcal{A}_i = \{A_k,\ k \in \mathcal{J}_i\}$, $i = 0, 1, 2$ and let $\mathcal{A}_i^*$ be random variables uniformly distributed in $\mathcal{A}_i$, $i = 0, 1, 2$. Furthermore, write

$$v_1(a) = \sum_{j=k+1}^{N} g_2(a, A_j), \quad v_2(a) = \sum_{j=m+1}^{k} g_2(a, A_j) \quad \text{and} \quad \Theta_i = \mathbf{E}^*|v_i(A_i^*)|, \tag{3.3}$$

$i = 1, 2$. Notice, that Θ_1 is a function of the r.v. $A_{k+1}, \dots, A_N$ and Θ_2 is a function of r.v. $A_{m+1}, \dots, A_N$.

The sampling without replacement sample is splited as follows. Put $X_j = A_j$, for $m < j \le N$. The rest of the sample, $X_1, \dots, X_m$, is obtained by simple random sampling without replacement from the (random) subpopulation $\mathcal{A}_0$.

Put $H = N\,\delta/\bigl(32\,q^{-1}\,N\,(\Theta_1 + \Theta_2) + 1\bigr)$. Notice, that H is a function of the random variables $A_{m+1}, \dots, A_N$ only. Let $f_1 = \mathbf{E}\,(\mathrm{e}\{t\,U_1\}\,\bigl|A_{m+1}, \dots, A_N)$ denote the conditional characteristic function of the probability distribution F_1, given $A_{m+1}, \dots, A_N$. An application of Prawitz's (1972) smoothing lemma yields conditionally, given $A_{m+1}, \dots, A_N$,

$$F_1(x+) \le \frac{1}{2} + \mathbf{E}\,\mathrm{V.P.}\int_R \mathrm{e}\{-x\,t\}\,\frac{1}{H}\,K\Bigl(\frac{t}{H}\Bigr)\,f_1(t)\,dt, \tag{3.4}$$

$$F_1(x-) \ge \frac{1}{2} - \mathbf{E}\,\mathrm{V.P.}\int_R \mathrm{e}\{-x\,t\}\,\frac{1}{H}\,K\Bigl(\frac{-t}{H}\Bigr)\,f_1(t)\,dt, \tag{3.5}$$

where $F(x+) = \lim_{z\downarrow x} F(z)$, $F(x-) = \lim_{z\uparrow x} F(z)$ and V.P. denotes Cauchy's Principal Value,

$$\text{V.P.} \int_R = \lim_{h\downarrow 0} \Big(\int_{-\infty}^{-h} + \int_h^{+\infty} \Big)$$

and $2\,K(s) = K_1(s) + i\,K_2(s)/(\pi\,s)$. Here

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

Combining (3.4) and the inversion formula, see e.g. [BGZ],

$$G(x) = \frac{1}{2} + \frac{i}{2\,\pi} \lim_{M\to\infty} \text{V.P.} \int_{|t|\le M} \mathrm{e}\{-t\,x\}\,\hat{G}(t)\,\frac{dt}{t}\,, \tag{3.6}$$

we get

$$F_1(x+) - G(x) \le \mathbf{E}I_1 + \mathbf{E}I_2 + \mathbf{E}I_3, \tag{3.7}$$

$$\begin{aligned} I_1 &= \frac{1}{2}\,H^{-1} \int_R \mathrm{e}\{-x\,t\}\,K_1\big(\frac{t}{H}\big)\,f_1(t)\,dt, \\ I_2 &= \frac{i}{2\,\pi}\ \text{V.P.} \int_R \mathrm{e}\{-x\,t\}K_2\big(\frac{t}{H}\big)\big(f_1(t) - \hat{G}(t)\big)\,\frac{dt}{t}\,, \\ I_3 &= \frac{i}{2\,\pi}\ \text{V.P.} \int_R \mathrm{e}\{-x\,t\}\Big(K_2\big(\frac{t}{H}\big) - 1\Big)\hat{G}(t)\,\frac{dt}{t}\,, \end{aligned}$$

where V.P. means that one should take $\lim_{M\to\infty}$, if it is necessary.

Combining (3.5) and (3.6) we obtain a bound for $G(x) - F_1(-x)$ similar to that given in (3.7). We shall bound $F_1(x+) - G(x)$ only:

$$|\mathbf{E}I_1| + |\mathbf{E}(I_2+I_3)| \ll N^{-1}\big(\beta_4/q + \delta^{-1}(\delta^{-1}+q^{-1}) + \delta^{-2}q^{-2}(\gamma_2^{1/2}+\gamma_2) + \gamma_4\big). \tag{3.8}$$

The analogous bound for $G(x) - F_1(x-)$ can be derived in the same way. Using these bounds and (3.2) we obtain the estimate of the theorem.

Estimate for $|\mathbf{E}I_1|$. Let $\nu = \big\{\nu_1, \dots, \nu_n\big\}$ be a sequence of i.i.d. Bernoulli random variables independent of $\overline{A}$ such that $\mathbf{P}\{\nu_1 = 1\} = 1 - \mathbf{P}\{\nu_1 = 0\} = p_1$, where $p_1 = m/(n - N + m)$. Write $q_1 = 1 - p_1$. By Erdös and Rényi (1959), the (conditional) characteristic function f_1 can be represented as

$$f_1(t) = \mathbf{P}^{-1}\{\mathcal{B}\}\,\frac{1}{2\,\pi} \int_{-\pi}^{\pi} \mathbf{E}\big(\mathrm{e}\{\tilde{U}_1\}\,\big|\,A_{m+1}, \dots, A_N\big)ds, \qquad \mathcal{B} = \{\sum_{k\in\mathcal{J}_0} \nu_k = m\},$$

where $\tilde U_1=\tilde V_1+\tilde T_1+\tilde S_1+\tilde\xi_1+\tilde Q_1$. Here $\tilde V_1=\sum_{k\in\mathcal J_0}s(\nu_k-p_1)$,

$$\tilde T_1=\sum_{k\in\mathcal J_0}tg_1(A_k)\,\nu_k,\qquad \tilde S_1=\sum_{k=m+1}^N tg_1(A_k),$$
$$\tilde Q_1=\sum_{m<k<l\le N}t\,g_2(A_k,A_l),\qquad \tilde\xi_1=\sum_{k\in J_0}\sum_{j=m+1}^N t\,g_2(A_k,A_j)\nu_k.$$

We shall replace the random bound H in the integral I_1 by a non–random one and $K_1(t/H)$ by 1. We have $|\mathbf E I_1|\le|\mathbf E I_4|+\mathbf E I_5$ where,

$$I_4=H^{-1}\int_{|t|\le H_1}\mathrm e\{-t\,x\}K_1\big(\frac{t}{H}\big)f_1(t)dt,$$
$$I_5=H^{-1}\int_{H_1\le|t|}K_1\big(\frac{t}{H}\big)\big|f_1(t)\big|dt\le H^{-1}\,\mathbb I\big\{H\ge H_1\big\}\int_{H_1\le|t|\le H}\big|f_1(t)\big|dt.$$

Let us show

$$\mathbf E I_5\ll N^{-2}\beta_3. \tag{3.9}$$

We have
$$I_5\le H_1^{-1}\int_{H_1\le|t|\le H}\big|f_1(t)\big|dt.$$

Put $W=\big|\mathbf E^*\,\mathrm e\{\tilde U_1\}\big|$ and

$$\eta_k=(\nu_k-p_1)\,\big(\beta(A_k)+t\,v(A_k)\big),\quad \beta(a)=t\,g_1(a)+s,v(a)=v_1(a)+v_2(a),$$
$$u(a)=\big|\mathbf E\,\mathrm e\{(\nu_1-p_1)\,(\beta(a)+t\,v(a))\}\big|^2=1-2\,p_1\,q_1\,\big(1-\cos(\beta(a)+t\,v(a)\,)\big).$$

Notice that $W=\big|\mathbf E^*\,\mathrm e\big\{\sum_{k\in\mathcal J_0}\eta_k\big\}\big|\le\prod_{k\in\mathcal J_0}u^{1/2}(A_k)$. Expanding the cosine function in powers of $t\,v(a)$ we get $u(a)\le u_1(a)+u_2(a)$, where

$$u_1(a)=1-2p_1q_1\big(1-\cos\beta(a)\big)\qquad\text{and}\qquad u_2(a)=2\,p_1\,q_1\,|t\,v(a)|.$$

Notice that by (3.1), p_1 is less than a sufficiently small absolute constant. Hence, we may and shall assume that $p_1<1/8$. Thus, $u_1(a)\ge 1-4\,p_1q_1\ge1/2$ and we can write $u_1(a)+u_2(a)\le u_1(a)\,(1+2\,u_2(a))$. We obtain

$$W\le\prod_{k\in\mathcal J_0}u_1^{1/2}(A_k)\prod_{k\in\mathcal J_0}\big(1+2\,u_2(A_k)\big)^{1/2}. \tag{3.10}$$

By Hölder's inequality, $\mathbf{E}(W\,|A_{m+1},\dots,A_N) \le W_1^{1/2}\,W_2^{1/2}$, where

$$W_1 = \mathbf{E}\,\Big(\prod_{k\in\mathcal{J}_0} u_1(A_k)\,\Big|\,A_{m+1},\dots,A_N\Big)$$
$$W_2 = \mathbf{E}\,\Big(\prod_{k\in\mathcal{J}_0} \big(1+2\,u_2(A_k)\big)\,\Big|\,A_{m+1},\dots,A_N\Big).$$

We shall show that

$$\mathbf{E}W_1^{1/2} \le \exp\{-m\,q_1\,\delta\} \qquad \text{and} \qquad W_2 \le \exp\{m\,q_1\,\delta/4\}. \tag{3.11}$$

An application of the well known arithmetic-geometric mean inequality,

$$s_1\,s_2\,\cdots\,s_r \le \Big(\frac{s_1+s_2+\cdots+s_r}{r}\Big)^r, \quad r=1,\,2,\,\dots, \tag{3.12}$$

yields

$$W_2 \le \big(\mathbf{E}^*(1+2\,u_2(A_0^*))\,\big)^{n-N+m}. \tag{3.13}$$

Recall that A_0^* denotes a random variable uniformly distributed over the (random) subpopulation $\mathcal{A}_0$. Furthermore, it follows from Lemma 5.1 that, for $|t|\le H$,

$$\mathbf{E}^*(1+2\,u_2(A_0^*)) \le 1+p_1\,q_1\,\delta/16 \le \exp\{p_1\,q_1\,\delta/16\}.$$

Indeed, by (5.1),

$$\mathbf{E}^*u_2(A_0^*) \le 2\,p_1q_1H\,E^*\big(|v_1(A_0^*)|+|v_2(A_0^*)|\big) \le 2\,p_1q_1Hq^{-1}(\Theta_1+\Theta_2) \le p_1q_1\delta/16.$$

Now the second inequality of (3.11) follows from (3.13).

Let us prove the first inequality of (3.11). By Hölder's inequality, $\mathbf{E}W_1^{1/2} \le \big(\mathbf{E}W_3\big)^{1/2}$, where $W_3 = \prod_{k\in\mathcal{J}_0} u_1(A_k) = \exp\{\sum_{k\in\mathcal{J}_0}\ln u_1(A_k)\}$. An application of Theorem 4, Hoeffding (1963), yields $\mathbf{E}\,W_3 \le \big(\mathbf{E}\,u_1(X_1)\big)^{n-N+m}$. Notice that

$$\mathbf{E}u_1(X_1) \le 1-2\,p_1\,q_1\,\delta \le \exp\big\{-2\,p_1\,q_1\,\delta\big\}.$$

We obtain, $\mathbf{E}W_3 \le \exp\{-2\,p_1\,q_1\,\delta\,(n-N+m)\}$ thus completing the proof of (3.11).

Combining the inequality $1/\mathbf{P}\{\mathcal{B}\} \ll m^{1/2}$, see (5.25), with (3.11) we get

$$\mathbf{E}I_5 \le H_1^{-1}\int_{H_1\le|t|\le N\delta}\mathbf{E}\,W\,dt \ll \frac{\beta_3}{N^{1/2}}\;m^{1/2}\,N\,\delta\exp\{-m\,q_1\,\delta/2\}. \tag{3.14}$$

Notice that (3.1) implies that $q_1 = 1 - p_1 \ge 1/2$. Hence, to get (3.9) it suffices to choose the constant in the definition of m big enough so that the exponent in (3.14) is less that N^{-3}.

Let us show

$$\mathbf{E} I_4 \ll N^{-1}\,\delta^{-2}\,(1 + q^{-2}\,\gamma_2) + N^{-1}\,\delta^{-1}\,q^{-1}(1 + q^{-1}\,\gamma_2^{1/2}). \tag{3.15}$$

By the inequality $\bigl|K_1(u) - 1\bigr| \le |u|$, we get $|\mathbf{E} I_4| \le |\mathbf{E} I_6| + \mathbf{E} I_7$, where

$$I_6 = H^{-1} \int_{|t| \le H_1} \mathrm{e}\{-t\,x\} f_1(t)dt \qquad \text{and} \qquad I_7 \le 2\,H^{-1} \int_0^{H_1} \frac{t}{H}\,dt. \tag{3.16}$$

Clearly, $I_7 \le H_1^2\,H^{-2}$. We have $H^{-2} \ll \delta^{-2} N^{-2} + \delta^{-2} q^{-2}(\Theta_1^2 + \Theta_2^2)$. An application of Lemma 5.1 gives

$$\mathbf{E} I_7 \ll \delta^{-2}\,N^{-1}\,\bigl(1 + q^{-2}\,\gamma_2\bigr).$$

Let us now estimate $\mathbf{E} I_6$. Write $f_2(t) = \mathbf{E}(\mathrm{e}\{t\,U\}\,\bigl|A_{m+1}, \dots, A_N)$. By the mean value theorem, $|f_2(t) - f_1(t)| \le \mathbf{E}(|t\,\Lambda|\,\bigl|A_{m+1}, \dots, A_N)$. Therefore, $\mathbf{E} f_6 = \mathbf{E} f_8 + R$, where

$$f_8 = H^{-1} \int_{|t| \le H_1} \mathrm{e}\{-t\,x\}\,f_2(t)\,dt \quad \text{and} \quad |R| \ll H_1^2\,\mathbf{E}\,H^{-1}\,\mathbf{E}\bigl(|\Lambda|\,\bigl|A_{m+1}, \dots, A_N\bigr).$$

We have $\mathbf{E}\,H^{-1}\mathbf{E}(|\Lambda|\,\bigl|A_{m+1}, \dots, A_N) = \mathbf{E} H^{-1}|\Lambda| \le \mathbf{E} H^{-2} + \mathbf{E}\Lambda^2$. By Lemma 5.1, $|R| \ll \delta^{-2}\,N^{-1}\,\bigl(1 + q^{-2}\,\gamma_2\bigr)$.

Let us estimate $\mathbf{E} I_8$. Clearly, $\mathbf{E} I_8 = \mathbf{E} I_9 + R$, where

$$I_9 = H^{-1} \int_{q^{-1}C' \le |t| \le H_1} \mathrm{e}\{-t\,x\}\,f_2(t)dt \quad \text{and} \quad |R| \ll q^{-1}\,\bigl(\mathbf{E} H^{-2}\bigr)^{1/2}$$

and where $C' > 2$ is a sufficiently large absolute constant. By Lemma 5.1,

$$|R| \ll q^{-1}\,\delta^{-1}\,N^{-1}\,\bigl(1 + q^{-1}\,\gamma_2^{1/2}\bigr).$$

Furthermore, $I_9 = \delta^{-1}\,I_{10} + 32\,q^{-1}\,\delta^{-1}\,(I_{11} + I_{12})$, where

$$I_{10} = N^{-1} \int_{q^{-1}C' \le |t| \le H_1} \mathrm{e}\{-t\,x\} f_2(t)dt,$$
$$I_{10+i} = \int_{q^{-1}C' \le |t| \le H_1} \mathrm{e}\{-t\,x\}\Theta_i f_2(t)dt, \quad \text{for} \quad i = 1, 2.$$

We shall show

$$\mathbf{E} I_{11} \ll N^{-1}\,(\gamma_2^{1/2} + \gamma_2). \tag{3.17}$$

Since Θ_1 is a function of $A_{k+1}, \dots, A_N$ only and $m < k$, we have $\mathbf{E}\Theta_1\, f_2(t) = \mathbf{E}\Theta_1\, f_3(t)$, where $f_3 = \mathbf{E}\,(\mathrm{e}\{t\,U\}\big|X_{k+1}, \dots, X_N)$.

Now we modify the sampling scheme. Let m_1 be an integer approximately equal to $C''\,N\,t^{-2}\,\ln|t|$. Here $C'' > 0$ is a large absolute constant to be specified latter. Notice, that $k \ge N/2$. Hence, we can choose the constant C' so big that $m_1 \le k$ and $m_1 \le q\,N/10$, for $q^{-1}\,C' \le |t| \le H_1$. Put $\mathcal{J}_3 = \{1, \dots, m_1\} \cup \{N+1, \dots, n\}$ and let $\mathcal{A}_3 = \{A_k,\, k \in \mathcal{J}_3\}$. Let A_3^* denote a random variable which is uniformly distributed in the (random) subpopulation $\mathcal{A}_3$,

$$v_3(a) = \sum_{k=m_1+1}^{N} g_2(a, A_j) \qquad \text{and} \qquad \Theta_3 = \mathbf{E}^*|v_3(A_3^*)|. \tag{3.18}$$

Notice, that Θ_3 is a function of $A_{m_1+1}, \dots, A_N$. We put $X_k = A_k$, for $k = m_1+1, \dots, N$. The remaining part of the sample, $X_1, \dots, X_{m_1}$, is obtained by a simple random sampling without replacement from $\mathcal{A}_3$.

Put $\Lambda_1 = \sum_{1 \le i < j \le m_1} g_2(A_i, A_j)$ and let $U_2 = U - \Lambda_1$. By the mean value theorem,

$$|f_3(t) - f_4(t)| \le \mathbf{E}\,\big(|t\,\Lambda_1|\,\big|A_{k+1}, \dots, A_N\big), \quad f_4(t) = \mathbf{E}\,\big(\mathrm{e}\{t\,U_2\}\big|A_{k+1}, \dots, A_N\big).$$

Therefore $\mathbf{E} I_{11} = \mathbf{E} I_{13} + R$, where

$$I_{13} = \int_{q^{-1}C' \le |t| \le H_1} \mathrm{e}\{-t\,x\}\,\Theta_1\, f_4(t)\,dt \quad \text{and} \quad |R| \le \int_{q^{-1}C' \le |t| \le H_1} |t|\,\mathbf{E}\,\Theta_1\,|\Lambda_1|\,dt.$$

By Hölder's inequality and the inequalities of Lemma 5.1,

$$\mathbf{E}\,\Theta_1\,|\Lambda_1| \le \big(\mathbf{E}\Theta_1^2\big)^{1/2}\,\big(\mathbf{E}\Lambda_1^2\big)^{1/2} \ll \frac{\ln|t|}{|t|}\,N^{-3/2}\,\gamma_2.$$

Therefore, $|R| \ll N^{-1}\,\gamma_2$. Furthermore, $|\mathbf{E}\,I_{13}| \le \mathbf{E}\,I_{14} + \mathbf{E}\,I_{15}$, where

$$I_{14} = \int_{q^{-1}C' \le |t| \le H_1} \Theta_1\,|f_4(t)|\,\mathbb{I}\big\{\Theta_3 \le \tfrac{c_1|t|}{N}\big\}dt,$$
$$I_{15} = \int_{\mathbf{R}} \Theta_1\,|f_4(t)|\,\mathbb{I}\big\{\Theta_3 > \tfrac{c_1|t|}{N}\big\}dt,$$

and where $c_1 > 0$ is small absolute constant to be specified later. Clearly,

$$\mathbf{E} I_{15} \le \mathbf{E}\,\Theta_1 \int_{|t| \le c_1^{-1} N\,\Theta_3} dt \ll N\,\mathbf{E}\,\Theta_1\,\Theta_3 \ll N\,\big(\mathbf{E}\,\Theta_1^2\big)^{1/2}\,\big(\mathbf{E}\,\Theta_3^2\big)^{1/2}.$$

By Lemma 5.1, $\mathbf{E}\, I_{15} \ll N^{-1}\, \gamma_2$.

Let us estimate $\mathbf{E} I_{14}$. Let $\nu_1', \dots, \nu_n'$ be a sequence of i.i.d. Bernoulli random variables independent of $\overline{A}$ and with probabilities $\mathbf{P}\{\nu_1' = 1\} = 1 - \mathbf{P}\{\nu_1' = 0\} = p_2$, where $p_2 = m_1/(n - N + m_1)$. Denote $q_2 = 1 - p_2$. By Erdös and Rényi (1959),

$$f_4(t) = \lambda \int_{-\pi\, \tau_1}^{\pi\, \tau_1} \mathbf{E}\big(\mathrm{e}\{\tilde U_2\}\,\big|A_{m_1+1}, \dots, A_N\big)\, ds.$$

Here $\tau_1 = (m_1\, q_2)^{1/2}$ and $\lambda^{-1} = 2\,\pi\,\tau_1\, \mathbf{P}\{\sum_{k \in \mathcal{J}_3} \nu_k' = m_1\}$. Notice, that $\lambda \ll 1$, see (5.25). Furthermore, $\tilde U_2 = \tilde V_2 + \tilde T_2 + \tilde S_2 + \tilde\xi_2 + \tilde Q_2$, with

$$\tilde V_2 = \sum_{k \in \mathcal{J}_3} (s/\tau)\, (\nu_k' - p_2), \quad \tilde T_2 = \sum_{k \in \mathcal{J}_3} t\, g_1(A_k)\, \nu_k', \quad \tilde S_2 = \sum_{k = m_1+1}^{N} t\, g_1(A_k),$$

$$\tilde\xi_2 = \sum_{k \in \mathcal{J}_3} \sum_{j = m_1+1}^{N} t\, g_2(A_k, A_j)\, \nu_k', \quad \tilde Q_2 = \sum_{m_1 < k < j \le N} t\, g_2(A_k, A_j).$$

Notice that the inequality $m_1 \le q\, N/10$ implies $p_2 \le 1/10$. Hence, the argument that leads to (3.10) applies to $\tilde W = |\mathbf{E}^*\, \mathrm{e}\{\tilde U_2\}|$ as well and yields

$$\tilde W \le \tilde W_1\, \tilde W_2, \quad \tilde W_1 = \prod_{k \in \mathcal{J}_3} \tilde u_1^{1/2}(A_k), \quad \tilde W_2 = \prod_{k \in \mathcal{J}_3} \big(1 + 2\, \tilde u_2(A_k)\big)^{1/2},$$

with $\tilde u_1(a) = 1 - 2\, p_2 q_2 \big(1 - \cos \tilde\beta(a)\big)$, $\tilde\beta(a) = t\, g_1(a) + s/\tau$ and $\tilde u_2 = 2\, p_2 q_2 |t\, v_3(a)|$. We obtain

$$\mathbf{E} I_{14} \ll \int_{q^{-1} C' \le |t| \le H_1} \int_{-\pi\tau}^{\pi\tau} \mathbf{E}\, M dt, \quad M = \Theta_1 \tilde W_1 \tilde W_2\, \mathbb{I}\big\{\Theta_3 \le \tfrac{c_1 |t|}{N}\big\}. \tag{3.19}$$

By Hölder's inequality,

$$\mathbf{E}\, M \le M_1^{1/2}\, M_2^{1/2}, \qquad M_1 = \mathbf{E}\, \Theta_1^2 \tilde W_2^2\, \mathbb{I}\big\{\Theta_3 \le \tfrac{c_1 |t|}{N}\big\}, \qquad M_2 = \mathbf{E}\, \tilde W_1^2.$$

An application of (3.12) yields

$$\tilde W_2^2 \le \big(1 + 2\, \mathbf{E}^* \tilde u_2(A_3^*)\big)^{|\mathcal{J}_3|} = \big(1 + 4\, p_2 q_2 |t|\, \Theta_3\big)^{|\mathcal{J}_3|}.$$

By the inequality $1 + x \le \exp\{x\}$, we have $\tilde W_2^2 \le \exp\{4\, m_1 q_2 |t|\, \Theta_3\}$. Thus, we obtain

$$M_1 \le \mathbf{E}\, \Theta_1^2\, \exp\{4\, C''\, c_1 \ln |t|\}. \tag{3.20}$$

Now consider M_2. An application of Theorem 4, Hoeffding (1963), yields $\mathbf{E}\,\tilde W_1^2 \le \big(\mathbf{E}\,\tilde u(X_1)\big)^{|\mathcal{J}_3|}$. By the inequality $1+x \le \exp\{x\}$, we get

$$\mathbf{E}\,\tilde W_1^2 \le \exp\{-2\,m_1\,q_2\,L\}, \qquad L = \mathbf{E}\big(1-\cos\tilde\beta(X_1)\big).$$

Furthermore,

$$L \ge n^{-1} \sum_{1\le k\le n,\, k\notin\mathcal{K}} \big(1-\cos\tilde\beta(a_k)\big).$$

By (5.26), $1-\cos\tilde\beta(a_k) \ge \frac{1}{2}\,\tilde\beta^2(a_k)\,\Theta(b_2)$, for $k\notin\mathcal{K}$. We obtain

$$L \ge \frac{1}{2}\,\Theta(b_2)\,\mathbf{E}\,(t\,g_1(X_1)+s/\tau)^2\,\mathbb{I}\big\{H_1\,|g_1(X_1)| < b_2\big\}.$$

An application of Lemma 5.6 yields $L \ge (t^2/N + (s/\tau)^2)\,b'$, where $b' = \frac{1}{2}\,\Theta(b_2)\,(1-2\,b_1/b_2)$. It is easy to check that $b'>0$. We obtain

$$m_1\,q_2\,L \ge q_2\,b'\,C''\,\ln|t| + b'\,s^2.$$

Since, $q_2 = 1-p_2 \ge 7/8$, we have

$$\mathbf{E}\tilde W_2^2 \le \exp\big\{-b'\,C''\,\ln|t| - b'\,s^2\big\}. \tag{3.21}$$

Collecting the estimates (3.20) and (3.21) in (3.19) one obtains

$$\mathbf{E}I_{14} \ll \big(\mathbf{E}\,\Theta_1^2\big)^{1/2}\,I, \quad I = \int_{-\pi\tau}^{\pi\tau} ds \int_{2\le|t|\le H_1} \exp\big\{-C''\,(b'-4\,c_1)\,\ln|t| - b'\,s^2\big\}dt.$$

Choose the constant C'' large enough such that $I \ll 1$. Then, by (5.1), $I_{14} \ll N^{-1}\,\gamma_2^{1/2}$ thus proving (3.17).

Clearly, the bound (3.17) extends to $\mathbf{E}\,I_{12}$ as well, i.e., $|\mathbf{E}\,I_{12}| \ll N^{-1}\,(\gamma_2^{1/2}+\gamma_2)$. Furthermore, $\mathbf{E}\,I_{10}$ can be estimated in the same way as $\mathbf{E}I_{11}$, just writing N^{-1} instead of Θ_1 everywhere in the proof of (3.17). Therefore, $|\mathbf{E}I_{10}| \ll N^{-1}\,(1+\gamma_2)$ and we obtain

$$|\mathbf{E}\,I_9| \ll N^{-1}\,\delta^{-1}\,q^{-1}\,(1+\gamma_2)$$

thus concluding the proof of (3.15).

It follows from (3.9) and (3.15) that

$$|\mathbf{E}I_1| \ll N^{-1}\beta_4 + N^{-1}\,\delta^{-2}\,(1+q^{-2}\,\gamma_2) + N^{-1}\,\delta^{-1}\,q^{-1}(1+q^{-1}\,\gamma_2^{1/2}). \tag{3.22}$$

Estimation of $|\mathbf{E}(I_2+I_3)|$. We have $I_2+I_3=(2\,\pi\,i)^{-1}\,(I_4-I_5+I_6+I_7)$, where

$$I_4=\int_{|t|\le H_1}\mathrm{e}\{-t\,x\}\big(f_1(t)-\hat G(t)\big)\,\frac{dt}{t}\,,\quad I_5=\int_{H_1\le|t|\le H}\mathrm{e}\{-t\,x\}K_2\big(\frac{t}{H}\big)f_1(t)\,\frac{dt}{t}\,,$$
$$I_6=\int_{|t|>H_1}\mathrm{e}\{-t\,x\}\,\hat G(t)\,\frac{dt}{t}\,,\qquad I_7=\int_{|t|<H_1}\mathrm{e}\{-t\,x\}\,\big(K_2\big(\frac{t}{H}\big)-1\big)f_1(t)\,\frac{dt}{t}\,.$$

Let us estimate $\mathbf{E}I_5$, $\mathbf{E}I_6$ and $\mathbf{E}I_7$. We have

$$\hat G(t)=\hat G_1(t)+\hat G_2(t), \tag{3.23}$$
$$\hat G_1(t)=\exp\{-t^2q/2\}\,\Big(1+i^3t^3\,\frac{q(q-p)\alpha}{6\,N^{1/2}}\Big),\quad \hat G_2(t)=\exp\{-t^2q/2\}\,i^3t^3\,\frac{q^2\,\kappa}{2\,N^{1/2}}\,.$$

Using (3.1) it is easy to show that $|\mathbf{E}I_6|\ll q^{-1}\beta_4/N+\gamma_2/N$. To estimate $\mathbf{E}I_7$ we use the inequality $|K_2(s)-1|\le 5\,s^2$. We have $|\mathbf{E}I_7|\le\mathbf{E}H^{-2}\,H_1^2$. Therefore, $|\mathbf{E}I_7|\ll\delta^{-2}\,N^{-1}\,(1+q^{-2}\gamma_2)$, cf (3.16). Furthermore, $|\mathbf{E}I_5|\le\mathbf{E}I_8$, where

$$I_8=\int_{H_1\le|t|\le H}|f_1(t)|\,\frac{dt}{|t|}\,.$$

The argument that leads to (3.14) applies to I_8 as well and yields $\mathbf{E}I_8\le N^{-2}\beta_3$. We obtain

$$\big|\mathbf{E}\,(I_2+I_3)\big|\ll|\mathbf{E}I_4|+N^{-1}\,\big(q^{-1}\,\beta_4+\delta^{-2}\,q^{-2}\,\gamma_2\big).$$

It remains to estimate $\mathbf{E}I_4$. We have

$$\mathbf{E}\,I_4=\int_{|T|\le H_1}\mathrm{e}\{-t\,x\}\,\big(\mathbf{E}\;\mathrm{e}\{t\,U_1\}-\hat G(t)\big)\,\frac{dt}{t}\,.$$

Expanding $\mathrm{e}\{-t\,\Lambda\}=-i\,t\,\Lambda+O(t^2\Lambda^2)$, we get $\mathbf{E}I_4=I_9-i\,2^{-1}m\,(m-1)I_{10}+R$, where

$$I_9=\int_{-H_1}^{H_1}\mathrm{e}\{-t\,x\}(\hat F(t)-\hat G(t))\,\frac{dt}{t}\,,\quad I_{10}=\int_{-H_1}^{H_1}\mathrm{e}\{-t\,x\}\mathbf{E}\,\mathrm{e}\{t\,U\}\,g_2(X_{N-1},X_N)dt,$$

and where $|R|\le H_1^2\,\mathbf{E}\,\Lambda^2\ll\delta^{-2}\,N^{-1}\gamma_2$. The last inequality follows from (5.3).

Consider I_{10}. The argument that leads to (3.17) applies to I_{10} as well and yields $|I_{10}|\ll N^{-3/2}\,(\gamma_2^{1/2}+\gamma_2)$. Indeed, just take $\tilde f_3=\mathbf{E}\,(\mathrm{e}\{t\,U\}\big|X_{N-1},X_N)$ instead of f_3 and write $g_2(X_{N-1},X_N)$ instead of Θ_1 in the proof of (3.17). We obtain $|\mathbf{E}I_4-I_9|\ll\delta^{-2}\,N^{-1}\,(\gamma_2^{1/2}+\gamma_2)$.

In the next section we shall show $|I_9|\ll N^{-1}\,(\beta_4/q+\gamma_4)$ thus completing the proof of the theorem.

Proof of Theorem 1.2. We again assume (without loss of generality) that $\sigma^2 = 1$. For the linear statistic L, we have $g_2(x,y) = 0$, for any $x,\, y \in \mathcal{A}$. Therefore the estimate of the theorem follows from (3.8).

Proof of Corollary 1.3. The corollary follows from Theorem 1.1, by the law of large numbers (LLN) for U- statistics, see, e.g., Serfling (1980). Given N, the function $\mathcal{H}$ and a sequence of i.i.d. observations $\mathcal{X}_1, \mathcal{X}_2, \dots$, introduce the sequence of finite populations $\mathcal{A}_n = \{\mathcal{X}_1, \dots, \mathcal{X}_n\}$ and the corresponding sequence of U-statistics, (U_n). Given $x \in R$, use the estimate of Theorem 1.1 applied to the sequence of probabilities $P_n\{x\} = \mathbf{P}\{U_n \le x\}$, n=1,2, … . By LLN, we have $\lim_n P_n\{x\} = \mathbf{P}\{\widehat{U} \le x\}$. Furthermore, the moments of the linear and quadratic parts of U_n in the expansion and in the remainder (in the estimate of Theorem 1.1) converge to the corresponding moments of the statistics $\widehat{U}$ thus proving Corollary 1.3.

4. Expansions

In this section we shall prove the inequality

$$\int_{|t|\le H_1} |t|^{-1}\big|\hat{F}(t) - \hat{G}(t)\big| dt \ll N^{-1}\,\big(\beta_4/q + \gamma_4\big). \tag{4.1}$$

Let $\overline{\nu} = \big(\nu_1,\, \nu_2, \dots,\, \nu_n\big)$ denote a sequence of i.i.d. Bernoulli r.v. with probabilities $\mathbf{P}\big(\nu_k = 1\big) = 1 - \mathbf{P}\big(\nu_k = 0\big) = p$. We shall assume that $\overline{A}$ and $\overline{\nu}$ are independent. By Erdös and Rényi (1959) and (1.1),

$$\hat{F}(t) = \lambda \int_{-\pi\,\tau}^{\pi\,\tau} \mathbf{E}\, \mathrm{e}\{\tilde{U}\} ds,$$

where $\tau = \sqrt{n\,p\,q}$, $\lambda^{-1} = 2\,\pi\,\tau\, \mathbf{P}\big(\sum_1^n \nu_k = N\big)$ and where

$$\tilde{U} = t\sum_{k=1}^{n} g_1(A_k)\,\nu_k + t\sum_{1\le k<l\le n} g_2(A_k,\, A_l)\nu_k\,\nu_l + \frac{s}{\tau}\sum_{1\le k\le n}(\nu_k - p).$$

Notice, that $(2\,\pi)^{-1/2} \le \lambda \le \sqrt{2}/\pi$, see (5.25). Furthermore, it follows from (1.2) that $\tilde{U} = T + Q$, where $T = \sum_1^n T_k$ and $Q = \sum_{1\le k<l\le n} Q_{k,l}$,

$$T_k = (\nu_k - p)\,\big(t\, g_1(A_k) + s/\tau\big), \quad Q_{k,l} = t\,(\nu_k - p)\,(\nu_l - p)\, g_2(A_k,\, A_l).$$

Observe, that the r. v. T_k, $k = 1,\, 2, \cdots, n$ as well as $Q_{k,l}$, $1 \le k < l \le n$ are identically distributed.

In what follows θ_1, θ_2, ... denote i.i.d. random variables which are uniformly distributed in $[0,1]$. For a vector valued smooth function H, the Taylor expansion may be written as

$$H(x) = H(0) + H'(0)\,x + \cdots + H^{(n)}(0)\,\frac{x^n}{n!} + \mathbf{E} H^{(n+1)}(\theta_1\,x)\,(1-\theta_1)^n\,\frac{x^{n+1}}{n!}.$$

In particular, we have the mean value formula, $H(x) - H(0) = \mathbf{E} H'(\theta_1\,x)\,x$. For $S = s_1 + \cdots + s_k$ we shall write $S^{(i)} = S - s_i$. Similarly, write $S^{(i,j)} = S - s_i - s_j$.

For a set of indices $\mathcal{B} \subset \{1,\ldots,m\}$ let $Y_{\mathcal{B}}$, respectively $Z_{\mathcal{B}}$, denote the absolute value of the conditional expectation of $\mathrm{e}\{\sum_{k\in\mathcal{B}} T_k\}$ respectively $\mathrm{e}\{\sum_{k\in\mathcal{B}}(T_k + \xi_k)\}$, given $A_1,\ldots,A_n$ and $\{\nu_j,\ j\notin\mathcal{B}\}$. We write $Y_{\mathcal{B}} = Z_{\mathcal{B}} = 1$ if $\mathcal{B} = \emptyset$. Here

$$\xi_k = t\,(\nu_k - p)\,\zeta(A_k), \qquad \zeta(A_k) = \sum_{m+1\le l\le n} (\nu_l - p)\,g_2(A_k, A_l).$$

We shall prove (4.1) in two steps. In the first step, see Lemmas 4.1 and 4.2 below, we estimate the difference $\mathbf{E}\ \mathrm{e}\{\tilde{U}\} - (h_1 + h_2)$, where

$$h_1 = \mathbf{E}\ \mathrm{e}\{T\}, \qquad h_2 = i^3\,C_n^2\,\mathbf{E}\,Q_{1,2}\,T_1\,T_2\ \mathrm{e}\{T^{(1,2)}\}. \tag{4.2}$$

In the proofs of Lemmas 4.1 and 4.2 we expand $\exp\{i\,\tilde{U}\}$ in powers of T_k and $Q_{k,l}$. In order to ensure the integrability (with respect to the measure $ds\,dt/|t|$) of the remainders of the expansion we split $\mathbf{E}\ \exp\{i\,\tilde{U}\}$ into a product of two functions (different for different values of s and t) so that the first one is the characteristic function of a sum of conditionally independent r. v. which vanishes sufficiently fast as s and t tends to infinity. This type of approach has been used by Helmers and van Zwet (1982), van Zwet (1984) and was extended in Götze and van Zwet (1991), [BGZ] to the problem of asymptotics expansions of general symmetric statistics of i.i.d. random variables. In the second step, see Lemma 4.3, we estimate the difference

$$\lambda\int_{-\pi\tau}^{\pi\tau}(h_1 + h_2)ds - \hat{G}.$$

The proof of Lemma 4.3 uses the Lindeberg method, cf. Lemma 6.2 in [BGZ]. Höglund (1978) used the Lindeberg method to prove the Berry-Essen bound for finite population sample means. Our proof of Lemma 4.3 is close to the proof of Höglund (1978).

Lemma 4.1. *Assume that $\beta_2 = 1$ and that (3.1) holds. There exist absolute constants $C_1 > 0$ and $c > 0$ such that*

$$\int_{C_1\,q^{-1}\le|t|\le H_1} \lambda\int_{|s|\le\pi\,\tau} \big|\mathbf{E}\ \mathrm{e}\{\tilde{U}\} - (h_1 + h_2)\big|ds\,\frac{dt}{|t|} \ll N^{-1}\,\mathcal{R}, \tag{4.3}$$

where $\mathcal{R} = \beta_4/q + \gamma_4$.

Proof of Lemma 4.1. Denote $x_i = g_1(A_i)$, and $y_{i,j} = g_2(A_i, A_j)$, $1 \le i, j \le n$, $i \ne j$. Throughout the proof we shall bound conditional moments by unconditional ones, that is, $\mathbf{E}^{3,4,5}\,|y_{1,2}|^r \ll \mathbf{E}\,|y_{1,2}|^r$ and $\mathbf{E}^{2,3,4}\,|x_1|^r \ll \mathbf{E}\,|x_1|^r$, $r = 2, 3, 4$, and will reduce the powers of $|t|$ and $|s|/\tau$, using the inequalities $|s|/\tau \le c$ and $|t| \le N^{1/2}$ (the last inequality follows from $H_1 \le N^{1/2}$).

For a function $f(s,t)$ we write $f \prec L$ if

$$\int_{C_1\, q^{-1} \le |t| \le H_1} \lambda \int_{|s| \le \pi\,\tau} |f| ds\, \frac{dt}{|t|} \ll L.$$

Given two functions $f(s,t)$ and $g(s,t)$ we write $f \sim g$ if $f - g \prec N^{-1}\,\mathcal{R}$.

Let $m = m(t,s)$ be the integer closest to the number $C_2\,q^{-1}\,n\,\varphi$, where

$$\varphi = \varphi(t,s) = \ln\bigl(t^2 + s^2/q\bigr)/\bigl(t^2 + s^2/q\bigr), \tag{4.4}$$

and where $C_2 > 0$ is a large, but absolute constant. We may and shall choose C_1 and C_2 large enough so that $10 < m < n$. Indeed, if $C_1 > e$ we have

$$\varphi(H_1, \pi\,\tau) \le \varphi(t,s) \le \ln\bigl((C_1/q)^2\bigr)/(C_1/q)^2 \le q\,C_1^{-1}\,(2/e), \tag{4.5}$$

for $C_1/q \le |t| \le H_1$ and $|s| \le \pi\,\tau$. Since $H_1 \le N^{1/2}$ we have $\varphi(H_1, \pi\,\tau) \ge c/(n\,p)$. Therefore, by (4.5), $m \ge C_2 c$. We choose C_2 large enough so that $m(s,t) > 4$. Furthermore, by (4.5), $m \le n\,(2/e)\,C_2/C_1$. Now choosing C_1 sufficiently large we obtain $m(s,t) < n$.

Put $\Lambda = \sum_{1 \le k < l \le m} Q_{k,l}$, $\tilde T = \sum_1^m T_k$, $\tilde Q = \sum_{m+1 \le k < l \le n} Q_{k,l}$ and

$$\xi = \sum_{1 \le k \le m} \xi_k, \qquad \xi_k = t\,(\nu_k - p)\,\zeta(A_k), \quad \zeta(A_k) = \sum_{m+1 \le l \le n} (\nu_l - p)\,y_{k,l}.$$

Then $\tilde U = W + \tilde T + \xi + \Lambda$, where $W = \sum_{m+1}^n T_k + \tilde Q$. Expanding the exponent in powers of Λ one obtains $\mathbf{E}\,\mathrm{e}\{\tilde U\} = f_1 + f_2 + R$, where

$$f_1 = i\,C_m^2\,\mathbf{E}\,Q_{1,2}\,\mathrm{e}\bigl\{W + \tilde T + \xi\bigr\}, \qquad f_2 = \mathbf{E}\,\mathrm{e}\bigl\{W + \tilde T + \xi\bigr\},$$

and $|R| \le \mathbf{E}\Lambda^2 = t^2\,m^2 (p\,q)^2\,N^{-3}\,\gamma_2$. It is easy to show $R \prec q\,N^{-1}\,\gamma_2$. Therefore, $R \sim 0$ and $\mathbf{E}\,\mathrm{e}\{\tilde U\} \sim f_1 + f_2$. The exponents in f_1 and f_2 consist of sums $W + \tilde T + \xi$, where W and $\tilde T$ are conditionally independent, given $A_1, \dots, A_n$, and ξ is stochastically negligible compared to $\tilde T$. Furthermore, given $A_1, \dots, A_n$, $\tilde T$ is a sum of conditionally independent r.v. Therefore the conditional expectation of $\mathrm{e}\{\tilde T\}$ vanishes rapidly as s and t tends to infinity.

We shall prove (4.3) in two steps. In the first step we show that

$$f_1 \simeq h_3, \qquad h_3 = i^3\, C_m^2\, \mathbf{E}\, Q_{1,2}\, T_1\, T_2\ \mathrm{e}\{T^{(1,2)}\}. \tag{4.6}$$

In the second step we prove

$$f_2 \simeq h_1 + h_4, \qquad h_4 = i^3 \Big(C_n^2 - C_m^2\Big)\, \mathbf{E}\, Q_{1,2}\, T_1\, T_2\ \mathrm{e}\{T^{(1,2)}\}. \tag{4.7}$$

Step 1. Expanding $\mathrm{e}\{\xi\} = \mathrm{e}\big\{\xi^{(1,2)} + (\xi_1 + \xi_2)\big\}$ in powers of $\xi_1 + \xi_2$ one obtains $f_1 = f_3 + f_4 + f_5 + f_6$, where $f_3 = i\, \mathbf{E}\, K$, $f_4 = 2i^2\, \mathbf{E}\, K\, \xi_1$,

$$f_5 = 2i^3 \mathbf{E} K\, \mathrm{e}\{\theta_1(\xi_1 + \xi_2)\}\xi_1^2(1 - \theta_1), \quad f_6 = 2i^3 \mathbf{E} K\, \mathrm{e}\{\theta_1\,(\xi_1 + \xi_2)\}\xi_1\xi_2(1 - \theta_1)$$

and where $K = C_m^2\, Q_{1,2}\ \mathrm{e}\big\{W + \tilde{T} + \xi^{(1,2)}\big\}$.
Let us show $f_5 \sim 0$ and $f_6 \sim 0$. We have

$$|f_6| \le m^2\, \mathbf{E}\, |Q_{1,2}\xi_1\, \xi_2| = m^2\, |t|^3\, (p\, q)^2\, \mathbf{E}\, \big|\zeta(A_1)\, \zeta(A_2)\, y_{1,2}\big|.$$

It follows from the inequalities $2\, |\zeta(A_1)\, \zeta(A_2)| \le \zeta^2(A_1) + \zeta^2(A_2)$ and

$$\mathbf{E}\, \zeta^2(A_1)\, |y_{1,2}| \le (n - m)\, p\, q\, \mathbf{E}\, y_{1,n}^2\, |y_{1,2}| \ll q\, N^{-7/2}\, \gamma_3$$

that $|f_6| \ll q\, \varphi^2\, |t|^3\, N^{-3/2}\, \gamma_3$. Using $|t| \le N^{-1/2}$ one may easily show that $f_6 \prec q\, N^{-1}\, \gamma_3$. Hence, $f_6 \sim 0$.
Let us show that $f_5 \sim 0$. First expanding the exponent in powers of $\theta_1\, \xi_2$ and then in powers of T_2 we get $|f_5| \le m^2\, (R_1 + R_2)$, where $R_1 = \mathbf{E}\, |Q_{1,2}\, \xi_2|\, \xi_1^2$ and $R_2 = \mathbf{E}\, |Q_{1,2}\, T_2|\, \xi_1^2$. Now consider R_1. An application of $a\, b^2 \ll |a|^3 + |b|^3$ yields

$$R_1 = (p\, q)^2\, t^4\, \mathbf{E}|y_{1,2}\, \zeta(A_2)|\, \zeta^2(A_1) \ll (p\, q)^2\, t^4\, \mathbf{E}|y_{1,2}|\, |\zeta(A_1)|^3.$$

It follows from (5.9) and the inequality $a\, b^3 \ll a^4 + b^4$ that

$$\mathbf{E}|y_{1,2}|\, |\zeta(A_1)|^3 \ll q^{3/2}\, N^{3/2}\, \mathbf{E}\, \big|y_{1,2}\, y_{1,n}^3\big| \ll q^{3/2}\, N^{-9/2}\, \gamma_4.$$

This inequality in combination with $t^2 \le N$ gives $m^2\, R_1 \ll q^{3/2}\, t^2\, \varphi^2\, N^{-3/2}\, \gamma_4$ and it is easy to show that $m^2\, R_1 \prec q^{3/2}\, N^{-3/2}\, \gamma_4$. Hence, $m^2\, R_1 \sim 0$. Finally, consider R_2. We have

$$R_2 = (p\, q)^3\, |t|^3\, (n - m)\, L, \qquad L = \mathbf{E}\, \big|y_{1,2}\, y_{1,n}^2\, (t\, x_2 + s/\tau)\big|.$$

It follows from the inequalities $|y_{1,2}\, x_2| \le N\, y_{1,2}^2 + N^{-1}\, x_2^2$ and $|s/\tau| \le c$ that $L \ll |t|\, N^{-5}(\gamma_4 + \gamma_2) + N^{-9/2}\gamma_3$. This inequality combined with $|t| \le N^{1/2}$

implies $m^2\,R_2 \ll q\,\varphi^2\,t^2\,N^{-1}\,\big(\gamma_3+\gamma_4+\gamma_2\big)$. Therefore, $m^2\,R_2 \prec q\,N^{-1}\,(\gamma_2+\gamma_4)$ and thus $m^2\,R_2 \sim 0$. We obtain $f_5 \sim 0$. Therefore $f_1 \sim f_3+f_4$.

Let us show $f_4 \simeq 0$. Expanding $\mathrm{e}\{\tilde T\} = \mathrm{e}\big\{\tilde T^{(2)}+T_2\big\}$ in powers of T_2 we obtain

$$f_4 = 2\,i^3\,C_m^2\,\mathbf{E}\,Q_{1,2}\,\xi_1\,T_2\;\mathrm{e}\big\{W+\tilde T^{(2)}+\xi^{(1,2)}+\theta_1\,T_2\big\}.$$

Take $\mathcal{B} = \{3,\,4\,\dots,\,m\}$. Then $|f_4| \le 2\,C_m^2\;\mathbf{E}\big|Q_{1,2}\,\xi_1\,T_2\big|\,Z_{\mathcal{B}}$. An application of Lemma 5.4 gives

$$|f_4| \le m^2\,\mathbf{E}\,\big(R_1+R_2\big),\quad \text{where}\quad R_1 = \big|Q_{1,2}\,\xi_1\,T_2\big|\,\mathbb{I}_{\mathcal{B}}\quad\text{and}\quad R_2 = \big|Q_{1,2}\,\xi_1\,T_2\big|\,\Psi_{\mathcal{B}}$$

and where $\mathbb{I}_{\mathcal{B}}$ and $\Psi_{\mathcal{B}}$ are defined in (5.13).

Let us show $m^2\,\mathbf{E}\,R_1 \sim 0$. It follows from Hölder's inequality (for conditional expectations) and Chebyshev's inequality that

$$\mathbf{E}\,(|\xi_1|\,\mathbb{I}_{\mathcal{B}}\big|A_1,A_2,\nu_1) \le V_1^{1/2}V_2^{1/2},\quad V_1 = \mathbf{E}(\xi_1^2\big|A_1,A_2,\nu_1),\quad V_2 = \mathbf{E}^{1,2}\varkappa_{\mathcal{B}}/\delta.$$

The r.v. $\varkappa_{\mathcal{B}}$ and constant δ are defined in (5.13). Clearly,

$$V_1 \le t^2(\nu_1-p)^2 q\,N\,\tilde\gamma_2,\quad\text{and}\quad V_2 \ll N\,p\,q\,(n-m)\mathbf{E}^{1,2}\,y_{3,4}^2 \ll q\,N^{-1}\,\gamma_2, \tag{4.8}$$

where $\tilde\gamma_2 = \mathbf{E}^{1,2}\,y_{1,n}^2$. These inequalities combined with the estimate $|T_2| \le |\nu_2-p|(N^{1/2}|x_2|+c)$ and the inequality

$$\mathbf{E}\,\tilde\gamma_2^{1/2}\,|y_{1,2}|\,N^{1/2}\,|x_2| \ll N^{-3}\,\gamma_2. \tag{4.9}$$

(to get it apply $a\,b \le a^2+b^2$ with $a = y_{1,2}$ and $b = N^{1/2}\,x_2\,\tilde\gamma_2^{1/2}$) implies $\mathbf{E}\,R_1 \ll q\,(p\,q)^2\,t^2\,N^{-3}\,\gamma_2^{3/2}$. Now it is easy to show $m^2\mathbf{E}\,R_1 \prec q^2N^{-1}\gamma_2^{3/2}$. Therefore, $m^2\,\mathbf{E}\,R_1 \sim 0$.

Let us show $m^2\,R_2 \sim 0$. Combining Hölder's inequality and Lemma 5.4 we obtain $\mathbf{E}(|\xi_1|\,\Psi_{\mathcal{B}}\big|A_1,A_2,\nu_1)\le V_1^{1/2}\,F_{\mathcal{B}}$ almost surely. The (non random) function $F_{\mathcal{B}}$ is defined in Lemma 5.4. We shall use the inequality $F_{\mathcal{B}} < (t^2+s^2/q)^r$, $|t| \ge C_1/q$, see (5.16). Here the positive number r is chosen big enough so that all powers of s and t appearing in the remainder terms are integrable over the region $\{(s,t):\,|t|\ge C_1/q\}$ with respect to the measure $dsdt/|t|$. We have $m^2\mathbf{E}R_2\le n^2\,F_{\mathcal{B}}\,\mathbf{E}|Q_{1,2}T_2|\,V_1^{1/2}$. Using (4.8) and (4.9) it follows that

$$\mathbf{E}|Q_{1,2}T_2|\,V_1^{1/2} \ll (p\,q)^2t^2\,\big(|t|\sqrt{q}+|s|\big)\,N^{-3}\,\gamma_2.$$

Simple calculations give $m^2R_2 \prec q^2\,N^{-1}\gamma_2$ and we obtain $f_4 \sim 0$. Hence, $f_1 \sim f_3$.

Expanding $\mathrm{e}\{\tilde{T}\} = \mathrm{e}\big\{\tilde{T}^{(1,2)} + T_1 + T_2\big\}$ first in powers of T_2 and then in powers of T_1 one obtains $f_3 = f_7 + R$, where

$$f_7 = i^3\, C_m^2\, \mathbf{E}\ \mathrm{e}\big\{W + \tilde{T}^{(1,2)} + \xi^{(1,2)}\big\}\, Q_{1,2}\, T_1\, T_2$$

and $|R| \le 2\, C_m^2\, \mathbf{E}\, Z_{\mathcal{B}} V_3$. Here $\mathcal{B} = \{3, 4, \dots, m\}$ and $V_3 = |Q_{1,2} T_1|\, T_2^2$. Let us show $R \sim 0$. An application of Lemma 5.4 and Chebyshev's inequality gives $\mathbf{E}^{1,2}\, Z_{\mathcal{B}} \le V_2 + F_{\mathcal{B}}$. Invoking (4.8) we get $|R| \le m^2 (c\, q\, N^{-1} \gamma_2 + F_{\mathcal{B}})\, \mathbf{E} V_3$. Some calculations show that

$$\mathbf{E}\, V_3 \ll (p\,q)^2\, |t|\, \mathbf{E}\, |y_{1,2}|\, z, \qquad z = |t|^3\, |x_1|\, x_2^2 + t^2\, x_2^2\, |s/\tau| + |s/\tau|^3.$$

Estimating $\mathbf{E}\, x_2^2 \ll N^{-1}$, $\mathbf{E}\, x_1^2\, x_2^2 \ll N^{-2}$ and using (5.7) we obtain

$$\mathbf{E} V_3 \ll N^{-3}\, (p\,q)^2\, |t|\, \Big(t^2\, (|t| + q^{-1/2} |s|)\, (\gamma_4^{1/2} + \beta_4^{1/2}) + q^{-3/2}\, |s|^3\, \gamma_2^{1/2}\Big). \tag{4.10}$$

The inequalities $\gamma_2\, \gamma_4^{1/2} \le \gamma_4$, $\gamma_2\, \beta_4^{1/2} \le \gamma_4 + \beta_4$ together with the bounds $|t| \le N^{1/2}$ and $q^{-1/2}\, |s| \ll N^{1/2}$ imply that

$$m^2\, q\, N^{-1} \gamma_2\, \mathbf{E}\, V_3 \ll N^{-1}\, q\, \varphi^2\, |t|\, \big(|t| + q^{-1/2}\, |s|\big)\, (\gamma_4 + \beta_4).$$

Therefore, $m^2\, q\, N^{-1} \gamma_2\, \mathbf{E}\, V_3 \prec q^2 N^{-1}\, (\gamma_4 + \beta_4)$. Using (4.10) one may easily show that $n^2 F_{\mathcal{B}}\, \mathbf{E} V_3 \prec q^3 N^{-1} (\gamma_4 + \beta_4)$. Hence, we get $R \sim 0$ and therefore, $f_3 \sim f_7$.

Next we shall show that $f_7 \sim f_8$, where

$$f_8 = i^3\, C_m^2\, \mathbf{E}\, \mathrm{e}\big\{W + \tilde{T}^{(1,2)}\big\}\, V_4, \qquad V_4 = Q_{1,2}\, T_1\, T_2.$$

Expanding the exponent in powers of $\xi^{(1,2)}$ we obtain $f_7 = f_8 + R_1 + R_2$, where

$$R_1 = i^4\, C_m^2\, (m-2)\, \mathbf{E}\ \mathrm{e}\big\{W + \tilde{T}^{(1,2)}\big\}\, \xi_3\, V_4$$

and $|R_2| \ll m^2 \mathbf{E}\, \big(\xi^{(1,2)}\big)^2\, |V_4|$. Let us show $R_2 \sim 0$. We have

$$\mathbf{E}^{1,2}\, \big(\xi^{(1,2)}\big)^2 \ll t^2\, m\, p\, q\, \mathbf{E}^{1,2}\, \zeta^2(A_3) \ll N^{-1}\, q^2\, t^2\, \varphi\, \gamma_2.$$

Some calculations yield

$$\mathbf{E}\, |V_4| \ll N^{-5/2}\, (p\,q)^2\, |t|\, (t^2 + s^2/q)\, (\gamma_2 + 1). \tag{4.11}$$

Estimating $|t| \le N^{1/2}$ we obtain $|R_2| \ll \varphi^3\, t^2\, (t^2 + s^2/q)\, N^{-1}\, (\gamma_2 + 1)$. It follows that $R_2 \prec N^{-1}\, \gamma_2\, (\gamma_2 + 1)$. Hence, $R_2 \sim 0$. It remains to show $R_1 \sim 0$. Expanding $\mathbf{E}\ \mathrm{e}\{T_3\} = 1 + i\, T_3\, \mathbf{E}\ \mathrm{e}\{\theta_1\, T_3\}$ we obtain $|R_1| \ll m^3\, \mathbf{E} Y_{\mathcal{B}} |\xi_3\, T_3\, V_4|$ with $\mathcal{B} = \{4, \dots, m\}$. The estimate

$$\mathbf{E}^{1,2} |\xi_3\, T_3| \ll N^{1/2}\, \mathbf{E}^{1,2}\, \xi_3^2 + N^{-1/2}\, \mathbf{E}\, T_3^2 \ll N^{-3/2}\, p\, q\, (t^2 + s^2/q)\, (\gamma_2 + 1)$$

combined with (5.14), (5.15) and (4.11) yields

$$|R_1| \ll N^{-1} q^3 F_{\mathcal{B}} |t| (t^2 + s^2/q)^2 (\gamma_2 + 1)^2.$$

Therefore, $R_1 \prec N^{-1} q^3 (\gamma_2 + 1)^2$ and thus $R_1 \sim 0$ which implies $f_7 \sim f_8$.

Let us show that $f_8 \sim h_3$. Expanding the exponent in powers of $\tilde{Q}$ we obtain $f_8 = h_3 + f_9 + R$, where

$$f_9 = i^4 C_m^2 C_{n-m}^2 \mathbf{E}\, \mathrm{e}\{T^{(1,2)}\} V_4 Q_{n-1,n} \quad \text{and} \quad |R| \le C_m^2 \mathbf{E}|V_4| \tilde{Q}^2 Y_{\mathcal{D}},$$

with $\mathcal{D} = \{3, \dots, m\}$. We shall show $R \sim 0$. Notice that given $A_1, \dots, A_m$ the r. v. $y_{k,l}$, $m+1 \le k < l \le n$ are identically distributed. Therefore,

$$|R| \le t^2 (p\, q)^2 C_m^2 C_{n-m}^2 \mathbf{E} |V_4| y_{n-1,n}^2 Y_{\mathcal{D}}.$$

Combining the inequality

$$\mathbf{E} |V_4| \mathbf{E}^{1,2} \big(y_{n-1,n}^2 \mathbf{E}^{1,2,n-1,n} Y_{\mathcal{D}}\big) \le F_{\mathcal{D}} N^{-3} \gamma_2 \mathbf{E} |V_4|$$

with (4.11) one obtains

$$|R| \ll N^{-11/2} n^2 m^2 (p\, q)^4 |t|^3 (t^2 + s^2/q) F_{\mathcal{D}} (\gamma_2 + 1) \gamma_2.$$

Now it is easy to show that $R \prec q^4 N^{-3/2} (\gamma_2 + 1) \gamma_2$. Hence, $R \sim 0$.

Let us show $f_9 \sim 0$. Expanding the exponent in powers of T_{n-1} and then in powers of T_n, we obtain $|f_9| \le C_m^2 C_{n-m}^2 \mathbf{E} \big|V_4 T_{n-1} T_n Q_{n-1,n}\big| Y_{\mathcal{D}}$. By (5.14) and (5.15), $|f_9| \le F_{\mathcal{D}} m^2 n^2 \mathbf{E}\big|V_4 T_{n-1} T_n Q_{n-1,n}\big|$. Furthermore, the inequality $\mathbf{E}^{1,2} |T_{n-1} T_n Q_{n-1,n}| \ll \mathbf{E} |T_{n-1} T_n Q_{n-1,n}|$ combined with (4.11) gives

$$|f_9| \ll m^2 n^2 F_{\mathcal{D}} \big(\mathbf{E}|V_4|\big)^2 \le c\, q^4 N^{-1} t^2 (t^4 + s^4/q^2) F_{\mathcal{D}} (\gamma_2 + 1)^2.$$

It is easy to show that $f_9 \prec q^4 N^{-1} (\gamma_2 + 1)^2$. Hence, $f_9 \sim 0$. We obtain $f_8 \sim h_3$ and this completes the proof of (4.6).

Step 2. Let us show $f_2 \sim f_{10} + f_{11}$, where $f_{10} = \mathbf{E}\, K$, $f_{11} = i\, \mathbf{E}\, K\, \xi$ and where $K = \mathrm{e}\{T + \tilde{Q}\}$. Put $m_0 = 0$, $m_4 = m$ and let m_1, m_2, m_3 be integers approximately equal to $m/4$, $m/2$ and $(3/4)m$ respectively. Put $\delta_i = \sum_{m_{i-1} < k \le m_i} \xi_k$, $i = 1, 2, 3, 4$. Expanding the exponent in f_2 in powers of ξ we get $f_2 = f_{10} + f_{11} + f_{12}$, where

$$f_{12} = i^2\, \mathbf{E}\, K\; \mathrm{e}\{\theta_1 \xi\} (1 - \theta_1) \sum_{1 \le i \le j \le 4} \delta_i\, \delta_j.$$

Hence, we need to show $f_{12} \sim 0$. Clearly, it suffices to prove $f_{13} \sim 0$ and $f_{14} \sim 0$, where

$$f_{13} = i^2 \mathbf{E} K \; \mathrm{e}\{\theta_1 \xi\} (1-\theta_1)\delta_1^2 \quad \text{and} \quad f_{14} = i^2 \mathbf{E} K \; \mathrm{e}\{\theta_1 \xi\} (1-\theta_1)\delta_1 \delta_2.$$

Let us show $f_{13} \sim 0$. Expanding the exponent in powers δ_2 we get $f_{13} = f_{15} + R$, where

$$f_{15} = i^3 \mathbf{E} K \; \mathrm{e}\big\{\theta_1 (\delta_1 + \theta_2\delta_2 + \delta_3 + \delta_4)\big\} (1-\theta_1)\delta_1^2\delta_2, \qquad |R| \le \mathbf{E}\, \delta_1^2 Y_{\mathcal{D}}$$

and $\mathcal{D} = \{m_1+1, \dots, m_2\}$. Using the fact that given A_j, $j \in \mathcal{D}$ the r. v. $y_{k,l}$, $1 \le k \le m_1$, $m < l \le n$ are identically distributed we obtain

$$\mathbf{E}\, \delta_1^2 Y_{\mathcal{D}} \;=\; t^2 \,(p\,q)^2\, m_1\, (n-m)\, \mathbf{E}\, y_{1,n}^2\, Y_{\mathcal{D}} \;\le\; q^2\, t^2\, F_{\mathcal{D}}\, N^{-1}\, \gamma_2. \tag{4.12}$$

In the last step we estimated $\mathbf{E}\, y_{1,n}^2\, Y_{\mathcal{D}} \le N^{-3}\, F_{\mathcal{D}}\, \gamma_2$, see (5.14) and (5.15). We have $R \prec q^2 N^{-1} \gamma_2$. Hence, $R \sim 0$. Let us show $f_{15} \sim 0$. Expanding the exponent in powers of δ_3 we obtain $|f_{15}| \le R_1 + R_2$, where $R_1 = \mathbf{E}\, \delta_1^2 |\delta_2| Y_{\mathcal{D}}$, with $\mathcal{D} = \{m_2+1, \dots, m_3\}$ and $R_2 = \mathbf{E}\, \delta_1^2 |\delta_2\, \delta_3|$.

Now consider R_1. Put

$$S_1 = m_1^{-1} \sum_{k=1}^{m_1} \zeta^2(A_k) \quad \text{and} \quad S_2 = (m_2 - m_1)^{-1} \sum_{k=m_1+1}^{m_2} \zeta^2(A_k).$$

Notice, that the conditional expectation of δ_1^2 given all random variables, but ν_k, $1 \le k \le m_1$ is $t^2\, p\, q\, m_1\, S_1$. By Jensen's inequality the conditional expectation of $|\delta_2|$ given all random variables, but ν_k, $m_1 < k \le m_2$ is bounded by $|t|\, (p\,q)^{1/2}\, (m_2 - m_1)^{1/2}\, S_2^{1/2}$. Therefore,

$$R_1 \le |t|^3\, (p\,q)^{3/2}\, m^{3/2}\, S_1\, S_2^{1/2} Y_{\mathcal{D}}.$$

It follows from the inequality $a\sqrt{b} \ll (a^{3/2} + b^{3/2})$ with $a = S_1$ and $b = S_2$ that $S_1\, S_2^{1/2} \ll \sqrt{2}(S_1^{3/2} + S_2^{3/2})$. Let us estimate $\mathbf{E}\, S_1^{3/2}\, Y_{\mathcal{D}}$. By Jensen's inequality,

$$S_1^{3/2} \le m_1^{-1} \sum_{k=1}^{m_1} |\zeta(A_k)|^3.$$

This inequality in combination with (5.14), (5.15) and (5.9) implies $E\, S_1^{3/2}\, Y_{\mathcal{D}} \ll q^{3/2}\, N^{-3}\, F_{\mathcal{D}}\, \gamma_3$. Clearly, the same bound holds for $E\, S_2^{3/2}\, Y_{\mathcal{D}}$. We obtain

$$\mathbf{E}\, \delta_1^2\, |\delta_2|\, Y_{\mathcal{D}} \ll t^3\, q^3\, N^{-3/2}\, F_{\mathcal{D}}\, \gamma_3. \tag{4.13}$$

Therefore, $R_1 \prec q^3 N^{-3/2} \gamma_3$ and thus $R_1 \sim 0$.

Finally, consider R_2. Clearly, $R_2 \le \mathbf{E}\,\delta_1^2\delta_2^2 + \mathbf{E}\,\delta_1^2\delta_3^2$. We shall estimate the first summand only. The conditional expectation of $\delta_1^2\delta_2^2$ given all random variables, but ν_k, $1 \le k \le m_2$ is $(p\,q)^2\, t^4\, m_1\,(m_2 - m_1)\,\zeta^2(A_1)\,\zeta^2(A_2)$. Furthermore, it is easy to show that

$$\mathbf{E}\,\zeta^2(A_1)\,\zeta^2(A_2) \le p\,q\,(n-m)\,N^{-6}\,\gamma_4 + (p\,q)^2\,(n-m)^2\,N^{-6}\,\gamma_2^2.$$

Estimating $t^2/N \le 1$ we obtain

$$\mathbf{E}\,\delta_1^2\delta_2^2 \ll q\,t^2\,\varphi^2\,(N^{-2}\,\gamma_4 + q\,N^{-1}\gamma_2^2). \tag{4.14}$$

Hence, $\mathbf{E}\,\delta_1^2\delta_2^2 \prec q\,(N^{-1}\,\gamma_4 + q\,\gamma_2^2)$. Therefore, $R_2 \sim 0$ and we obtain $f_{13} \sim 0$.

Let us show $f_{14} \sim 0$. Expanding the exponent in powers of δ_3 we get $f_{14} = f_{16} + R$, where

$$f_{16} = i^3\,\mathbf{E}\,K\ \mathrm{e}\big\{\theta_1\,(\delta_1 + \delta_2 + \theta_2\,\delta_3 + \delta_4)\big\}\,\delta_1\,\delta_2\,\delta_3, \qquad |R| \le \mathbf{E}|\delta_1\,\delta_2|\,Y_{\mathcal{D}},$$

and $\mathcal{D} = \{m_2+1, \ldots, m_3\}$. We have $|R| \le \mathbf{E}\,\delta_1^2\,Y_{\mathcal{D}} + \mathbf{E}\,\delta_2^2\,Y_{\mathcal{D}}$. Obviously, the bound (4.12) applies to these expectations as well and yields $R \sim 0$.

Let us show $f_{16} \sim 0$. Expanding the exponent in f_{16} in powers of δ_4 one obtains $|f_{16}| \le R_1 + R_2$, where we denote $R_1 = \mathbf{E}\,|\delta_1\,\delta_2\,\delta_3\,\delta_4|$ and $R_2 = \mathbf{E}|\delta_1\,\delta_2\,\delta_3|\,Y_{\mathcal{D}}$, and where $\mathcal{D} = \{m_3+1,\ \ldots, m\}$. By means of the inequality $a\,b \le a^2 + b^2$, R_1 can be bounded by the sum of expectations $\mathbf{E}\delta_i^2\delta_j^2$, $i \ne j$. Clearly the bound (4.14) applies to these expectations as well. Therefore, $R_1 \sim 0$. Consider R_2. We have $R_2 \le \mathbf{E}\,\delta_1^2\,|\delta_3|\,Y_{\mathcal{D}} + \mathbf{E}\,\delta_2^2\,|\delta_3|\,Y_{\mathcal{D}}$. The bound (4.13) extends to these expectations as well. Hence, $R_2 \sim 0$. Therefore $f_{16} \sim 0$ and we obtain $f_2 \sim f_{10} + f_{11}$.

Consider $f_{11} = m\,\mathbf{E}\ \mathrm{e}\{T + \tilde{Q}\}\,\xi_1$. Expanding the exponent in powers of T_1 we get $f_{11} = f_{17} + R$, where

$$f_{17} = m\,\mathbf{E}\ \mathrm{e}\{T^{(1)} + \tilde{Q}\}\,\xi_1\,T_1, \qquad |R| \le m\,\mathbf{E}\,|\xi_1|\,T_1^2\,Y_{\mathcal{D}},$$

and where $\mathcal{D} = \{2, \ldots, m\}$. We have $|R| \le R_1 + R_2$, where $R_1 = m\,\mathbf{E}\,T_1^4\,Y_{\mathcal{D}}$ and $R_2 = m\,\mathbf{E}\,\xi_1^2\,Y_{\mathcal{D}}$. By (5.14) and (5.15), $R_1 \le p\,q\,m\,F_{\mathcal{D}}\,N^{-2}\,(t^4 + s^4/q^2)\beta_4$. and $R_2 \le t^2\,q^2\,F_{\mathcal{D}}\,N^{-1}\,\gamma_2$, cf. (4.12). It follows that $R_1 \prec q\,N^{-1}\beta_4$ and $R_2 \prec q^2\,N^{-1}\,\gamma_2$. Therefore, $R \sim 0$ and we obtain $f_{11} \sim f_{17}$.

We shall show $f_{17} \sim f_{18}$, where $f_{18} = m\,i^2\mathbf{E}\ \mathrm{e}\{T^{(1)}\}\,\xi_1\,T_1$. Expanding the exponent in powers of $\tilde{Q}$ we obtain $f_{17} = f_{18} + R$, where $|R| \le m\,\mathbf{E}\,|\xi_1\,\tilde{Q}\,T_1|\,Y_{\mathcal{D}}$ and $\mathcal{D} = \{2, \ldots, m\}$. We have $R \le R_1 + R_2$, $R_1 = m\,\mathbf{E}\,\tilde{Q}^2\,T_1^2\,Y_{\mathcal{D}}$ and $R_2 = m\,\mathbf{E}\,\xi_1^2\,Y_{\mathcal{D}}$. It is shown above that $m\,\mathbf{E}\,\xi_1^2\,Y_{\mathcal{D}} \sim 0$. Hence, it suffices to prove $R_1 \sim 0$. We have $R_1 = m\,C_{n-m}^2\,\mathbf{E}\,Q_{n-1,n}^2 T_1^2 Y_{\mathcal{D}}$. By (5.14) and (5.15), $E\,Q_{n-1,n}^2 T_1^2 Y_{\mathcal{D}} \le F_{\mathcal{D}} E\,Q_{n-1,n}^2 T_1^2$.

Simple calculations show $R_1 \ll q^3\, t^2\, (t^2+s^2/q)\, F_{\mathcal{D}}\, N^{-1}\, \gamma_2$. Hence, $R_1 \prec q^3\, N^{-1}\, \gamma_2$ and thus $R_1 \sim 0$ and we obtain $f_{11} \sim f_{18}$.

Finally expanding the exponent in $f_{18} = m\,(n-m)\, i^2 \mathbf{E}\ \mathrm{e}\{T^{(1)}\}\, Q_{1,2}\, T_1$ in powers of T_2 we get $f_{18} = f_{19} + R$, where

$$f_{19} = m\,(n-m)\, i^3\, \mathbf{E}\ \mathrm{e}\{T^{(1,2)}\}\, Q_{1,2}\, T_1\, T_2, \qquad |R| \le m\,(n-m)\, \mathbf{E}\,|V_3|\, Y_{\mathcal{D}},$$

and where $\mathcal{D} = \{2,\dots,m\}$. We have $\mathbf{E}\,|V_3|\, Y_{\mathcal{D}} \le F_{\mathcal{D}}\, \mathbf{E}\,|V_3|$. Using (4.10) it is easy to show that $R \prec q^2\, N^{-1}\,(\gamma_4 + \beta_4)$. Therefore $R \sim 0$ and we obtain $f_{11} \sim f_{19}$.

We shall complete the proof of (4.7) by showing that

$$f_{10} \sim h_1 + f_{20}, \qquad f_{20} = i^3\, C_{n-m}^2\, \mathbf{E}\ \mathrm{e}\{T^{(1,2)}\}\, Q_{1,2}\, T_1\, T_2.$$

Expanding the exponent in f_{10} in powers of $\tilde{Q}$ we get $f_{10} = h_1 + f_{20} + R$, where $|R| \le \mathbf{E}\tilde{Q}^2 Y_{\mathcal{D}}$ and where $\mathcal{D} = \{1,\dots,m\}$. We have $\mathbf{E}\tilde{Q}^2 Y_{\mathcal{D}} = C_{n-m}^2 \mathbf{E} Q_{n-1,n}^2 Y_{\mathcal{D}}$. By (5.14), (5.15), $\mathbf{E} Q_{n-1,n}^2 Y_{\mathcal{D}} \le F_{\mathcal{D}}\, E Q_{n-1,n}^2$. Therefore, $|R| \ll q^2\, t^2\, F_{\mathcal{D}}\, N^{-1}\, \gamma_2$. Hence, $R \sim 0$ and we obtain $f_{10} \sim h_1 + f_{20}$ which completes the proof.

Lemma 4.2. *Assume that $\beta_2 = 1$ and that (3.1) holds. For any $C_1 > 0$ there exists a constant $c > 0$ which depends only on C_1 such that*

$$\int_{|t| \le C_1/q} \lambda \int_{|s| \le \pi\,\tau} \big|\mathbf{E}\ \mathrm{e}\{\tilde{U}\} - (h_1 + h_2)\big| ds\, \frac{dt}{|t|} \ll N^{-1}\,(\beta_4/q + \gamma_2). \tag{4.15}$$

Proof of Lemma 4.2. For a function $f(s,t)$ we write $f \prec L$ if

$$\int_{|t| \le C_1/q} \lambda \int_{|s| \le \pi\,\tau} |f| ds\, \frac{dt}{|t|} \ll L.$$

Given two functions $f(s,t)$ and $g(s,t)$ we write $f \sim g$ if $f - g \prec N^{-1}\,(\beta_4/q + \gamma_2)$.

Let $m = m(s)$ be an integer approximately equal to $C_3\, n\, \varphi$, where

$$\varphi = \ln\big(s^2 + C_4\big)/\big(s^2 + C_4\big), \tag{4.16}$$

and where C_3 and C_4 are large, but absolute constants. As in proof of Lemma 4.1 we may and shall assume that $10 < m < n$.

Define W, ξ, $\tilde{T}$, $\tilde{Q}$, Λ and f_1, f_2 in the same way as in proof of Lemma 4.1. Then $\mathbf{E}\ \mathrm{e}\{\tilde{U}\} = f_1 + f_2 + R$, with $|R| \le \mathbf{E}\,\Lambda^2$. We have $\mathbf{E}\,\Lambda^2 \ll q^2\, t^2\, \varphi^2\, N^{-1}\, \gamma_2$. It is easy to show that $\Lambda \prec N^{-1}\, \gamma_2$ and we obtain $\mathbf{E}\ \mathrm{e}\{\tilde{U}\} \sim f_1 + f_2$.

We shall prove the estimate (4.18) for f_2, see below. Expanding the exponent in powers of ξ we obtain $f_2 = f_3 + f_4 + R$, where $f_3 = \mathbf{E}\,\mathrm{e}\{W + \tilde{T}\}$, $f_4 =$

$i\,\mathbf{E}\,\mathrm{e}\big\{W+\tilde T\big\}\,\xi$ and $|R|\le \mathbf{E}\,\xi^2$. We have $\mathbf{E}\,\xi^2 \ll q^2\,t^2\,\varphi\,N^{-1}\,\gamma_2$. Now it is easy to show $R\prec N^{-1}\,\gamma_2$. Hence, $f_2\sim f_3+f_4$.

Consider f_4. Let m_1 be an integer approximately equal $m/2$. Put $\delta_1=\sum_{k=1}^{m_1}\xi_k$ and $\delta_2=\xi-\delta_1$. We have $f_4=f_5+f_6$, where $f_j=i\,\mathbf{E}\,\mathrm{e}\big\{W+\tilde T\big\}\,\delta_{j-4}$, $j=5,6$. Expanding the exponent in f_5 in powers of $\tilde Q$ one obtains $f_5=i\,\mathbf{E}\,\mathrm{e}\{T\}\,\delta_1+R$, where $|R|\le \mathbf{E}\,|\tilde Q\,\delta_1|\,Y_{\mathcal B}$ and $\mathcal B=\{m_1+1\dots,m\}$. We have $|R|\le \mathbf{E}\,\tilde Q^2\,Y_{\mathcal B}+\mathbf{E}\,\delta_1^2$. Clearly, $\mathbf{E}\,\delta_1^2\prec N^{-1}\,\gamma_2$, cf. the estimate of $\mathbf{E}\,\xi^2$ above. Furthermore, by (5.14) and (5.15),

$$\mathbf{E}\,\tilde Q^2\,Y_{\mathcal B}=t^2\,(p\,q)^2\,C_{n-m}^2\,\mathbf{E}\,g_2^2(A_{n-1},A_n)\,Y_{\mathcal B}\ll q^2\,t^2\,F_{\mathcal B}\,N^{-1}\,\gamma_2. \tag{4.17}$$

It follows that $\mathbf{E}\tilde Q^2 Y_{\mathcal B}\prec N^{-1}\gamma_2$ and we obtain $R\sim 0$. Hence, $f_5\sim i\,\mathbf{E}\,\mathrm{e}\{T\}\,\delta_1$ and similarly $f_6\sim \mathbf{E}\;\mathrm{e}\{T\}\,\delta_2$. Therefore, $f_4\sim i\,\mathbf{E}\;\mathrm{e}\{T\}\,\xi$.

Expanding the exponent in f_3 in powers of $\tilde Q$ one obtains $f_3=h_1+f_7+R$, where

$$f_7=i\,C_{n-m}^2\,\mathbf{E}\;\mathrm{e}\{T\}\,Q_{1,2},\qquad |R|\le \mathbf{E}\,\tilde Q^2\,Y_{\mathcal C}\quad\text{and}\quad \mathcal C=\{1,\dots,m\}.$$

Clearly, the bound (4.17) applies to $\mathbf{E}\,\tilde Q^2\,Y_{\mathcal C}$ as well. Therefore, $R\sim 0$. Hence, $f_3\sim h_1+f_7$ and we obtain

$$f_2\sim h_1+i\,\big(C_{n-m}^2+m\,(n-m)\big)\mathbf{E}\;\mathrm{e}\{T\}\,Q_{1,2}. \tag{4.18}$$

It remains to bound f_1. Expanding f_1 in powers of ξ one obtains $f_1=i\,\mathbf{E}\;\mathrm{e}\big\{W+\tilde T\big\}\,\Lambda+R$, where $|R|\le \mathbf{E}\,\big|\Lambda\,\xi\big|$. Since $|R|\le \mathbf{E}\,\Lambda^2+\mathbf{E}\,\xi^2$ and both expectations can be estimated as above, we have $R\sim 0$. Furthermore, expanding the exponent in powers of $\tilde Q$ we get

$$i\,\mathbf{E}\;\mathrm{e}\big\{W+\tilde T\big\}\,\Lambda=i\,\mathbf{E}\;\mathrm{e}\{T\}\,\Lambda+R,\qquad\text{where}\qquad |R|\le \mathbf{E}\,\big|\Lambda\,\tilde Q\big|.$$

The inequality $ab\le a^2\varphi+b^2\varphi^{-1}$ with $a=\tilde Q$, $b=\Lambda$ gives $|R|\le \mathbf{E}\big(\varphi^{-1}\Lambda^2+\varphi\,\tilde Q^2\big)$. We have $|R|\ll q^2\,t^2\,\varphi\,N^{-1}\,\gamma_2$. It is easy to show that $R\prec N^{-1}\,\gamma_2$. Therefore, $R\sim 0$. Hence, we obtain $f_1\sim i\,\mathbf{E}\;\mathrm{e}\{T\}\,\Lambda$. This relation in combination with (4.18) implies $f_1+f_2\sim h_1+f_9$, where $f_9=i\,C_n^2\,\mathbf{E}\;\mathrm{e}\{T\}\,Q_{1,2}$.

It remains to show $f_9\sim h_2$. Expanding the exponent in in powers of T_1 and then in powers of T_2 one obtains $f_{10}=h_2+R$, with $|R|\le n^2\,\mathbf{E}\,|Q_{1,2}\,T_1|\,T_2^2\,Y_{\mathcal B}$, where $\mathcal B=\{3,\dots,n\}$. Let us show $R\sim 0$. By (5.14), (5.15), we get $|R|\le n^2 F_{\mathcal B} E\,|Q_{1,2}T_1|\,T_2^2$. Put $\overline{T}_1=(\nu_2-p)\,t\,g_1(A_2)$ and $\overline{T}_2=(\nu_2-p)(s/\tau)$. The inequality $T_2^2\le 2\,\overline{T}_1^2+2\,\overline{T}_2^2$ gives

$$E\,|Q_{1,2}\,T_1|\,T_2^2\le 2\,(R_1+R_2),\qquad R_1=E\,|Q_{1,2}\,T_1|\,\overline{T}_1^2,\qquad R_2=\,E\,|Q_{1,2}\,T_1|\,\overline{T}_2^2.$$

Further, calculations show that

$$R_2 \le (p\,q)^2\,N^{-3}\,(s^2/q)\big(t^2(\gamma_2+1)+\gamma_2^{1/2}\,|t\,s|/\sqrt{q}\big).$$

The inequality $a\,b \le a^2+b^2$ with $a=|g_2(A_1,A_2)|$, $b=|t\,g_1(A_1)|\,g_1^2(A_2)$ yields

$$R_1 \le (p\,q)^2\,N^{-3}\,|t|^3\,(\gamma_2+(t^2+s^2/q)\beta_4\big).$$

Furthermore, proceeding as in proof of Lemma 5.4 one may show that $F_{\mathcal{B}} \le \exp\{-c\,(t^2\,q+s^2)\}$, where $c>0$ is an absolute constant. Therefore, $F_{\mathcal{B}}\,t^r\,s^k \prec c(r,k)/q^{(r+1)/2}$, for $r,\ k>0$, where the constant $c(r,k)$ depends on r and k only. It follows that $n^2\,F_{\mathcal{B}}\,R_1 \prec N^{-1}\,(\beta_4/q+\gamma_2)$ and $n^2\,F_{\mathcal{B}}\,R_2 \prec N^{-1}\,(1+\gamma_2^{1/2}+\gamma_2)$. Hence, we obtain $R\sim 0$ thus completing the proof of the lemma.

Write $\beta(t)=\mathbf{E}\,\mathrm{e}\{t\,(\nu_0-p)\}$, where ν_0 denotes a Bernoulli random variable with probabilities $\mathbf{P}\{\nu_0=1\}=p$ and $\mathbf{P}\{\nu_0=0\}=q$.

Lemma 4.3. *Assume that $\beta_2=1$ and suppose that (3.1) holds. For $i=1,2$, we have*

$$\int_{|t|\le H_1}\Big|\lambda\int_{|s|\le\pi\,\tau} h_i\,ds-\hat{G}_i\Big|\,\frac{dt}{|t|} \ll N^{-1}\,\big(q^{-1}\,\beta_4+\gamma_2\big). \tag{4.19}$$

Proof of Lemma 4.3. For a function $f(s,t)$ write $f\prec L$ if

$$\tau\,\lambda\int_{|t|\le H_1}\Big|\int_{|s|\le\pi} f ds\Big|\,\frac{dt}{|t|} \ll L.$$

Given two functions $f(s,t)$ and $g(s,t)$ write $f\sim g$ if $f-g\prec N^{-1}\,\big(q^{-1}\,\beta_4+\gamma_2\big)$. By Erdös and Renyi (1959), the function

$$t\to\lambda\int_{|s|\le\pi\,\tau} h_1\,ds=\tau\,\lambda\int_{|s|\le\pi} f_1(s,t)ds,\quad f_1(s,t)=\prod_{i=1}^{n}\beta\big(t\,g_1(a_i)+s\big),$$

is the Fourier transform of the probability distribution function of the sum $g_1(A_1)+\dots+g_1(A_N)$. In particular,

$$\tau\,\lambda\int_{|s|\le\pi}\beta^n(s)ds=1.$$

Therefore, (4.19) can be rewritten as $f_i\sim\varphi_i$, $i=1,2$, where $\varphi_i(s,t)=\beta^n(s)\hat{G}_i(t)$ and where

$$f_2(s,t)=\mathbf{E}\,\omega(s,t)\,\prod_{k=3}^{n}\beta\big(s+t\,g_1(A_k)\big),\quad \omega\big(\tfrac{s}{\tau},t\big)=i^3\,C_n^2\,Q_{1,2}\,T_1\,T_2.$$

In order to prove $f_1 \sim \varphi_1$ we shall show $f_1 - \varphi_3 \sim \varphi_4$, where

$$\varphi_3 = \beta^n(s)\,\alpha(t^2/p), \qquad \varphi_4 = \beta^n(s)\,N^{-1/2}\,\alpha(t^2/p)\,t^3\,\gamma^{(3)}(0)\,\alpha_3/(6\,p).$$

Here $\alpha(\eta) = \exp\{-p\,q\,\eta/2\}$ and $\gamma(\eta) = \beta(\eta)\,\exp\{p\,q\,\eta^2/2\}$. Write $\zeta = s^2 n + t^2/p$,

$$y_k = g_1(a_k), \quad k = 1, 2, \dots, n, \quad \text{and} \quad v(r) = \prod_{k=1}^{r} \gamma(s + t y_k), \quad r = 1, 2, \dots, n.$$

It follows from $\mathbf{E} g_1(A_1) = 0$ and $\beta_2 = 1$ that $f_1 = \alpha(\zeta)\,v(n)$ and $\varphi_3 = \alpha(\zeta)\,\gamma^n(s)$. Consider the difference $f_1 - \varphi_3$. The identity

$$\prod_1^n a_j - \prod_1^n b_j = \sum_{k=1}^{n} \Big(\prod_1^{k-1} a_j\Big)(a_k - b_k) \prod_{k+1}^{n} b_j, \tag{4.20}$$

with $a_k = \gamma(s + t\,y_k)$, $b_k = \gamma(s)$, combined with the mean value theorem,

$$\gamma(s + t\,y_k) - \gamma(s) = t\,y_k\,\mathbf{E}\gamma'\big(\xi\big), \quad \xi = s + \theta_1\,t\,y_k, \tag{4.21}$$

and the Taylor expansion, $\gamma'(\xi) = \gamma^{(3)}(0)\,\xi^2/2 + \mathbf{E}\gamma^{(4)}(\theta_2\,\xi)\xi^3\,(1-\theta_2)^2/2$, implies

$$v(n) - \gamma^n(s) = v_1 + v_2, \quad v_i = \sum_{k=1}^{n} \psi_k\,\delta_i(k), \quad \psi_k = \gamma^{n-k}(s)\,v(k-1), \quad i = 1, 2,$$

$$\delta_1(k) = \gamma^{(3)}(0)\,t\,y_k\,\mathbf{E}\,\xi^2/2, \qquad \delta_2(k) = t\,y_k\,\mathbf{E}\gamma^{(4)}(\theta_2\,\xi)\xi^3\,(1-\theta_2)^2/2.$$

Thus we obtain $f_1 - \varphi_3 = f_3 + f_4$, where $f_i = \alpha(\zeta)\,v_{i-2}$, $i = 3, 4$.

Let us show $f_4 \sim 0$. First, we shall estimate ψ_k uniformly over $k = 1, 2, \dots, n$,

$$|\psi_k| \le \exp\{p\,q\,W\,\zeta/2\}, \quad W = \frac{3}{2}\,(1 - \Theta(b_2)) + 2\Theta(b_2)(b_1/b_2). \tag{4.22}$$

Using $|\beta(s)| \le \alpha(\Theta(b_2)\,s^2)$, $|\beta(s + t y_k)| \le \alpha(\Theta(b_2)(s + t y_k)^2)$, for $k \notin \mathcal{K}$, see (5.27), and $|\beta(s + t y_k)| \le 1$, for $k \in \mathcal{K}$, we obtain

$$|\psi_k| \le \exp\{p\,q\,W_1/2\}, \quad \text{with} \quad W_1 = (1 - \Theta(b_2))\,(s^2(n-k) + B_1) + \Theta(b_2)\,B_2,$$

where $B_1 = \sum_{j=1}^{k-1}(s + t y_j)^2$ and $B_2 = \sum_{j=1, j\in\mathcal{K}}^{k-1}(s + t y_j)^2$. The inequalities

$$\sum_{j=1}^{k-1} y_k^2 \le \beta_2/p = 1/p \quad \text{and} \quad \Big|2 \sum_{j=1}^{k-1} y_j\Big| \le (n/p)^{1/2},$$

for the second one see Höglund (1978), yield $B_1 \le s^2(k-1)+t^2/p+s\,t\,(n/p)^{1/2}$. An application of $2\,a\,b \le a^2+b^2$ with $a=n^{1/2}s$, $b=p^{-1/2}t$ yields $B_1 \le s^2(k-1)+t^2/p+\zeta/2$. Furthermore, by (5.21) and (5.24), we have $B_2 \le 2\,(b_1/b_2)\,\zeta$. Collecting the bounds for B_1 and B_2 we obtain $W_1 \le W\,\zeta$, thus proving (4.22).

Consider $\delta_2(k)$. Simple calculations show that

$$|\gamma^{(4)}(\eta)| \ll (p\,q + (p\,q\,\eta)^4\,)\exp\{p\,q\,\eta^2/2\},$$

for any $\eta \in R$. Using $|s| \le c$ and $y_k^2 \le 1/p$, $k=1,2,\dots,n$ (the last inequality follows from $y_k^2 \le \beta_2/p$ and $\beta_2=1$) it follows

$$\big|\gamma^{(4)}\big(\theta_2\,\xi)\big)\big| \ll d\,\exp\big\{p\,q\,(ty_k)^2\big\}, \quad d=\big(p\,q+(t\,p^{1/2}q)^4\big).$$

Furthermore, we have

$$\exp\{p\,q\,(t\,y_k)^2\} \ll \hat e, \qquad \hat e=\exp\{q\,t^2\,(b_1/b_2)\}. \tag{4.23}$$

Indeed, if $k \notin \mathcal{K}$ then (4.23) follows from the inequality $(t\,y_k)^2 \le b_2^2 \le c$. For $k \in \mathcal{K}$, (5.24) gives $y_k^2 \le (b_1/b_2)/p$, thus proving (4.23). We obtain $|\gamma^{(4)}(\theta_2\,\xi)| \ll \hat e\,d$. Furthermore, an application of the inequality $|\xi|^3 \ll \big(|s|^3+|ty_k|^3\big)$ yields

$$|\delta_2(k)| \ll |t|\,d\,\hat e\,|y_k|\,(|s|^3+|ty_k|^3) \tag{4.24}$$

and we obtain $|f_4| \ll \alpha((1-W)\,\zeta)\,t\,\hat e\,d\,B_3$, where $B_3=\sum_1^n|y_k|\,(|s|^3+|ty_k|^3)$. An application of Lemma 5.2 gives $B_3 \ll N^{-1}\,\beta_4\,\big(\,(s^2n)^{3/2}\,p^{1/2}+|t|^3\,p\big)$. Write $V=(1-W)\,\zeta-2(b_1/b_2)\,t^2/p$. Then

$$|f_4| \ll N^{-1}\,\beta_4\,\alpha(V)\,t\,d\,\big(\,(s^2n)^{3/2}\,p^{1/2}+|t|^3\,p\big). \tag{4.25}$$

The constants b_1 and b_2 are choosen so that $W+2\,(b_1/b_2)<1$. Therefore,

$$\alpha(V) \le \alpha\big(c\,\zeta\big), \tag{4.26}$$

where $c>0$ is a small, but absolute constant. It is easy to show that (4.25) combined with (4.26) implies $f_4 \prec q^{-1}\,N^{-1}\,\beta_4$. We obtain $f_4 \sim 0$ and therefore, $f_1-\varphi_3 \sim f_3$.

Let us prove $f_3 \sim \varphi_4$. It follows from the identities

$$\psi_j=\gamma^{n-1}(s)+\gamma^{n-j}(s)\,\big(v(j-1)-\gamma^{j-1}(s)\big),$$

$$v(j-1)-\gamma^{j-1}(s)=\sum_{k=1}^{j-1}v(k-1)\big(\gamma(s+ty_k)-\gamma(s)\big)\gamma^{j-k-1}(s)$$

that $f_3 = f_5 + f_6$, where $f_5 = \alpha(\zeta)\,\gamma^{n-1}(s)\sum_{j=1}^n \delta_1(j)$ and

$$f_6 = \alpha(\zeta)\sum_{j=1}^{n}\delta_1(j)\sum_{k=1}^{j-1}\bigl(\gamma(s+ty_k)-\gamma(s)\bigr)\,\phi_k, \qquad \phi_k = \gamma^{n-k-1}(s)v(k-1).$$

We shall show that $f_6 \sim 0$. It is easy to see that (4.22) applies to ϕ_k as well, i.e. $|\phi_k| \le \exp\{p\,q\,W\,\zeta/2\}$. Furthermore, the mean value theorem (4.21), combined with the Taylor expansion $\gamma'(\xi) = \mathbf{E}(1-\theta_2)\gamma^{(3)}\bigl(\theta_2\xi\bigr)\xi^2$, yields

$$\bigl|\gamma(s+ty_k)-\gamma(s)\bigr| \le |ty_k|\mathbf{E}\bigl|\gamma^{(3)}\bigl(\theta_2\xi\bigr)\bigr|\,\xi^2. \tag{4.27}$$

Some calculations show that $|\gamma^{(3)}(\eta)| \ll \exp\{p\,q\eta^2/2\}\,\bigl(p\,q + |p\,q\,\eta|^3\bigr)$, for $\eta \in R$. The argument which leads to the inequality $|\gamma^{(4)}(\theta_2\,\xi)| \ll \hat{e}\,d$ above, yields $|\gamma^{(3)}(\theta_2\,\xi)| \ll \hat{e}\,d'$ where $d' = p\,q + |t\,p^{1/2}\,q|^3$. We obtain

$$\bigl|\gamma(s+ty_k)-\gamma(s)\bigr| \ll |ty_k|\,\hat{e}\,d'\,(s^2+(ty_k)^2). \tag{4.28}$$

Furthermore, it is easy to show that

$$|\delta_1(j)| \ll |ty_j|\,p\,q\,(s^2+(ty_j)^2). \tag{4.29}$$

Combining the bound $|\phi_k|\,\hat{e} \le \exp\{p\,q\,W\,\zeta/2\}\,\hat{e} \le \alpha(V)$, with the inequalities (4.28) and (4.29) one obtains

$$|f_6| \ll \alpha(V)\,t^2\,p\,q\,d'\,B_4^2, \quad \text{where} \quad B_4 = \sum_1^n |y_k|\,(s^2+(ty_k)^2).$$

By Lemma 5.2, $B_4 \le N^{-1/2}\,\beta_3\,\zeta$. Therefore,

$$|f_6| \ll N^{-1}\,\beta_3^2\,\alpha(V)\,p\,q\,d'\,t^2\,\zeta^2,$$

and using (4.26) one can show that $f_6 \prec q^{-1}\,N^{-1}\,\beta_3^2/N$. Therefore, $f_6 \sim 0$.

Let us show that $f_5 \sim \varphi_4$. We have $f_5 = f_7 + f_8$, where

$$f_7 = \alpha(\zeta)\gamma^n(s)\sum_{k=1}^n\delta_1(k), \qquad f_8 = \alpha(\zeta)\bigl(\gamma^{n-1}(s)-\gamma^n(s)\bigr)\sum_{k=1}^n\delta_1(k).$$

Let us show that $f_8 \sim 0$. By the mean value theorem, $1-\gamma(s) = \gamma(0)-\gamma(s) = \mathbf{E}\gamma'(\theta_1 s)\,s$. This combined with the bound $|\gamma'(\eta)| \ll p\,q$, for $|\eta| \le c$, gives

$|1-\gamma(s)| \ll p\,q\,|s|$, for $|s| \le \pi$. Furthermore, by (4.22), $|\alpha(\zeta)\gamma^{n-1}(s)| \le \alpha(V)$. Therefore,

$$\alpha(\zeta)\,|\gamma^{n-1}(s)-\gamma^{n}(s)| \ll \alpha(V)\,p\,q\,|s|.$$

It follows from (4.29) that

$$|f_8| \ll \alpha(V)\,(p\,q)^2\,|s\,t|\,B_4 \ll N^{-1/2}\,\beta_3\,(p\,q)^2\,\alpha(V)\,|s\,t|\,\zeta.$$

Using (4.26) it easily follows that $f_8 \prec p\,\beta_3/N$. Hence, $f_8 \sim 0$.

Let us show that $f_7 \sim \varphi_4$. Observe, that $\alpha(\zeta)\gamma^n(s) = \alpha(t^2/p)\,\beta^n(s)$. It follows from $\mathbf{E}g_1(A_1)=0$ and $\beta_2 = 1$ that

$$\sum_{k=1}^{n}\delta_1(k) = \gamma^{(3)}(0)\Big(s\,t^2/(2\,p) + (t^3/6)\sum_{k=1}^{n} y_k^3\Big).$$

Therefore, $f_7-\varphi_4 = f_9$, where $f_9 = \alpha(t^2/p)\,\gamma^{(3)}(0)\,s\,t^2\,(2\,p)^{-1})\,\beta^n(s)$. Since

$$\int_{|s|\le\pi} s\,\exp\{-p\,q\,s^2 n/2\}ds = 0,$$

we have $f_9 \sim f_{10}$, where $f_{10} = \alpha(t^2/p)\,\gamma^{(3)}(0)s\,t^2\,(2\,p)^{-1}\,\big(\beta^n(s)-\alpha^n(s)\big)$. Let us show that $f_{10}\sim 0$. The identity (4.20) with $a_k = \beta(s)$, $b_k = \alpha(s)$, combined with the Taylor expansion, $\beta(s)-\alpha(s) = \mathbf{E}(1-\theta_1)^2\,\big(\beta^{(3)}(\theta_1\,s)-\alpha^{(3)}(\theta_1\,s)\big)\,s^3/2$, and the estimate $|\beta^{(3)}(\theta_1 s)| + |\alpha^{(3)}(\theta_1 s)| \ll p\,q$, for $|s|\le\pi$, yields

$$|\beta^n(s)-\alpha^n(s)| \ll p\,q\,|s|^3\sum_{k=1}^{n}\big|\beta^{k-1}(s)\,\alpha^{n-k}(s)\big|.$$

Furthermore the inequality $|\gamma^{(3)}(0)| \ll p\,q$ combined with the inequalities

$$\big|\beta^{k-1}(s)\alpha^{n-k}(s)\big| \le \alpha\big(s^2\,(k-1)\Theta(b_2) + s^2\,(n-k)\big) \le \alpha\big(\Theta(b_2)\,s^2(n-1)\big),$$

see (5.27), yields $|f_{10}| \ll \alpha(c_1\,\zeta)\,p\,q^2\,t^2\,s^4\,n$ for a small, but absolute constant c_1. Hence, it follows that $f_{10} \prec q^{-1}\,N^{-1}$. We obtain $f_{10}\sim 0$ which completes the proof of $f_1 \sim \varphi_1$.

Let us prove $f_2 \sim \varphi_2$. By (1.2), $\mathbf{E}\,\omega(s,t) = i^3\,t^3\,C_n^2\,(p\,q)^2\,N^{-5/2}\,\kappa$. Hence

$$\alpha(t^2/p)\,\mathbf{E}\,\omega(s,t) = (1-n^{-1})\,\hat{G}_2$$

and therefore, $(1-n^{-1})\,\varphi_2 = \alpha(\zeta)\,\gamma^n(s)\,\mathbf{E}\,\omega(s,t)$. Put $f_{11} = \alpha(\zeta)\,v(n)\,\mathbf{E}\,\omega(s,t)$. Notice that $f_{11} = \mathbf{E}\,\omega(s,t)\,\prod_1^n \beta(s+t\,g_1(A_k)\,)$. Therefore, $f_2 = f_{11}+f_{12}$, where $f_{12} = \mathbf{E}\,\omega(s,t)\,u_1\,(1-u_2)$,

$$u_1 = \prod_{3}^{n}\beta(s+t\,g_1(A_k)\,) \qquad \text{and} \qquad u_2 = \prod_{1}^{2}\beta(s+t\,g_1(A_k)\,).$$

Let us prove $f_2 - (1-n^{-1})\varphi_2 \sim 0$. We shall show $f_{12} \sim 0$ and $f_{13} \sim 0$, where $f_{13} = f_{11} - (1-n^{-1})\,\varphi_2$.

Consider $f_{13} = \alpha(\zeta)\,(v(n) - \gamma^n(s))\,\mathbf{E}\,\omega(s,t)$. The identity (4.20) with $a_k = \gamma(s + t\,y_k)$, $b_k = \gamma(s)$, combined with (4.22), (4.28) and (4.23) yields $|f_{13}| \ll \alpha(V)\,|t|\,d'\,|\mathbf{E}\,\omega(s,t)|\,B_4$. Substitution of the bound $B_4 \le N^{-1/2}\beta_3\zeta$ gives $|f_{13}| \ll N^{-1}\beta_3|\kappa|\,\alpha(V)\,t^4\zeta\,q^2 d'$. Using (4.26) we easily obtain $f_{13} \prec N^{-1}\beta_3\,|\varkappa|$. Hence, $f_{13} \sim 0$.

Let us show $f_{12} \sim 0$. Write

$$1 - u_2 = 1 - \beta\big(s + tg_1(A_1)\big) + \beta\big(s + tg_1(A_1)\big)\big(1 - \beta\big(s + tg_1(A_2)\big)\big).$$

By the mean value theorem, we have $1 - \beta(\eta) = \beta(0) - \beta(\eta) = \mathbf{E}\beta'(\theta_1\eta)\,\eta$. This combined with the bounds $|\beta\big(\eta)\big)| \le 1$ and $|\beta'(\eta)| \le p\,q$ yields

$$|1 - u_2| \ll p\,q\,\big(|s| + |t\,g_1(A_1)| + |t\,g_1(A_2)|\big).$$

Using this inequality it is easy to show that

$$\mathbf{E}|\omega|\,|1 - u_2| \ll n^2\,(p\,q)^3\,|t|\,\Big(|s|^3\,N^{-3/2}\,\gamma_1 + \big(|s|\,t^2\,N^{-5/2} + |t|^3\,N^{-3}\big)(\gamma_2 + \beta_4)\Big).$$

Now consider u_1. Estimating $\big|\beta\big(s + t\,g_1(A_k)\big)\big| \le \alpha\big(\Theta(b_2)\,(s + tg_1(A_k))^2\big)$, for $|g_1(A_k)| < b_2/H_1$, see (5.27), and $\big|\beta\big(s + t\,g_1(A_k)\big)\big| \le 1$ otherwise, we obtain

$$|u_1| \le \alpha\big(\Theta(b_2)\,W_2\big), \qquad W_2 = \sum_{k=3}^{n} \mathbb{I}\big\{|g_1(A_k)| < b_2/H_1\big\}\,(s + t\,g_1(A_k)\,)^2.$$

We have $W_2 \ge C_1 - C_2 - C_3$, where

$$C_1 = \sum_1^n (s + t\,y_k)^2, \quad C_2 = \sum_1^2 (s + tg_1(A_k))^2 \quad \text{and} \quad C_3 = \sum_{k\in\mathcal{K}} (s + t\,y_k)^2.$$

Clearly, $C_1 = \zeta$. By (5.21) and (5.24), $C_3 \le 2\,(b_1/b_2)\,\zeta$. Furthermore, combining the inequalities $(a+b)^2 \ll a^2 + b^2$ and (5.6) we get $C_2 \le 4\,s^2 + 4\,t^2\,(\beta_4/N)^{1/2}/p$. By (3.1), β_4/N is bounded by a sufficiently small absolute constant. Therefore we can write $C_2 \le 4\,s^2 + t^2/(2p)$. We obtain $W_2 \ge \zeta\,b'$, $b' = 1 - 2\,(b_1/b_2) - 4/n - 1/2$. Some calculations show that $b' \ge c > 0$. Therefore, $|u_1| \le \alpha(c\,\zeta)$, where $c > 0$ is an absolute constant. Hence, we get $|f_{12}| \le \alpha(c\,\zeta)\mathbf{E}|\omega||1 - u_2|$. Invoking the bound for $\mathbf{E}|\omega|\,|1 - u_2|$ we get $f_{12} \prec p\,q\,N^{-1}\,(\gamma_1 + \gamma_2 + \beta_4)$. Therefore $f_{12} \sim 0$ and hence $f_2 \sim (1 - 1/n)\varphi_2$. In view of $\varphi_2/n \prec p\,N^{-3/2}\,|\varkappa|$ this completes the proof of the lemma.

5. Auxiliary results

The section is organized as follows. In the first part we collect the necessary moment inequalities. The second part provides bounds for conditional characteristic functions. In the last part we state for easier reference some known inequalities.

5.1. Moment inequalities.

Lemma 5.1. *The random variables v_i and Θ_i, $i = 1, 2, 3$, defined in (3.3) and (3.18) satisfy the following inequalities.*

$$\mathbf{E}\Theta_i^2 \le N^{-2}\,\gamma_2, \qquad i = 1, 2, 3, \tag{5.1}$$

$$\mathbf{E}^*|v_i(A_0^*)| \le q^{-1}\,\Theta_i, \qquad i = 1, 2. \tag{5.2}$$

Let $2 \le m \le n$. For $\Lambda = \sum_{1\le i<j\le m} g_2(A_i, A_j)$, we have

$$\mathbf{E}\,\Lambda^2 = \frac{m\,(m-1)}{2\,N^3}\,(1 - c_\Lambda)\,\gamma_2, \quad c_\Lambda = \frac{2\,(m-2)}{n-2} - \frac{(m-2)(m-3)}{(n-2)(n-3)}\,. \tag{5.3}$$

Proof of Lemma 5.1. We shall prove (5.1), for the case where $i = 1$ only since the proof of (5.1) for $i = 2$, 3 is similar. By Hölder's inequality,

$$\Theta_1^2 \le \mathbf{E}^* v_1^2(A_1^*) = \sum_{k+1\le i,j\le N} \mathbf{E}^* g_2(A_1^*, A_i)\,g_2(A_1^*, A_j).$$

Notice, that for $k = 1 \le i, j \le N$,

$$\mathbf{E}\,\mathbf{E}^* g_2(A_1^*, A_i)\,g_2(A_1^*, A_j) = \mathbf{E}g_2(A_1, A_i)\,g_2(A_1, A_j).$$

We obtain

$$\mathbf{E}\Theta_1^2 \le (N-k)\,\mathbf{E}g_2^2(A_1, A_2) + (N-k)\,(N-k-1)\,\mathbf{E}g_2(A_1, A_2)\,g_2(A_1, A_3).$$

Using (1.2) it is easy to show that

$$\mathbf{E}g_2(A_1, A_2)\,g_2(A_1, A_3) = -(n-2)^{-1}\,\mathbf{E}g_2^2(A_1, A_2). \tag{5.4}$$

This completes the proof.

We shall prove (5.2) in the case $i = 1$ since the proof in the case $i = 2$ is similar. Since

$$\mathbf{E}^*\big|v_1(A_0^*)\big| = |\mathcal{J}_0|^{-1} \sum_{k\in\mathcal{J}_0} |v_1(A_k)| \le |\mathcal{J}_0|^{-1} \sum_{k\in\mathcal{J}_1} |v_1(A_k)| \le \Theta_1\,\frac{|\mathcal{J}_1|}{|\mathcal{J}_0|}\,,$$

the inequality $|\mathcal{J}_1|/|\mathcal{J}_0| \le q^{-1}$ completes the proof.

In order to show (5.3) write $\mathbf{E}\Lambda^2 = 2^{-1}(m-1)m\,\mathbf{E}\,g_2(A_1,A_2)\Lambda$. We have

$$\mathbf{E}g_2(A_1,A_2)\,\Lambda = \mathbf{E}g_2^2(A_1,A_2) + 2\,(m-2)\,\mathbf{E}g_2(A_1,A_2)\,g_2(A_1,A_3) + 2^{-1}\,(m-2)\,(m-3)\,\mathbf{E}g_2(A_1,A_2)\,g_2(A_3,A_4).$$

Now (5.3) follows from (5.4) and the identity

$$\mathbf{E}g_2(A_1,A_2)\,g_2(A_3,A_4) = \frac{2}{(n-2)\,(n-3)}\,\mathbf{E}\,g_2^2(A_1,A_2),$$

which is a consequence of (1.2). This proves the lemma.

Lemma 5.2. *Assume that $\beta_2 = 1$. Then*

$$\sum_1^n |g_1(a_k)| \le n^{3/2}\,p^{1/2}\,\beta_4/N, \qquad \sum_1^n |g_1(a_k)| \le N^{-1/2}\,n\,\beta_3, \tag{5.5}$$

$$g_1^2(A_1) + g_2^2(A_2) \le \bigl(2\,\beta_4\,p^{-1}\,N^{-1}\bigr)^{1/2}, \tag{5.6}$$

$$\mathbf{E}\,|y_{1,2}|\,x_2^2 \ll N^{-1/2}\,Z, \qquad \mathbf{E}\,|y_{1,2}\,x_1|\,x_2^2 \ll N^{-1}\,Z, \tag{5.7}$$

where $y_{1,2} = g_2(A_1,A_2)$, $x_i = g_1(A_i)$, $i = 1,2$ and $Z = N^{-2}\,(\beta_4^{1/2} + \gamma_4^{1/2})$.

Proof of Lemma 5.2. Write $y_k = g_1(a_k)$, $k = 1,\dots,n$. The inequalities (5.5) follow from

$$p^{-1} = \sum_{k=1}^n y_k^2 \le n^{(r-2)/r}\bigl(\sum_{k=1}^n |y_k|^r\bigr)^{2/r}, \quad \sum_{k=1}^n |y_k| \le n^{(r-1)/r}\bigl(\sum_{k=1}^n |y_k|^r\bigr)^{1/r}, \quad r = 3,\,4.$$

The inequality (5.6) is a consequence of

$$g_1^2(A_1) + g_2^2(A_2) \le \sqrt{2}\,\Bigl(\sum_{k=1}^2 g_1^4(A_k)\Bigr)^{1/2} \quad \text{and} \quad \sum_{k=1}^2 g_1^4(A_k) \le \sum_{k=1}^n y_k^4 = \beta_4/(p\,N).$$

The inequalities (5.7) follows from the inequalities

$$\mathbf{E}\,|y_{1,2}|\,x_2^2 \le \bigl(\mathbf{E}\,Y\bigr)^{1/2}\mathbf{E}\,x_2^2, \quad E\,|y_{1,2}x_1|\,x_2^2 \le \bigl(\mathbf{E}\,Y\bigr)^{1/2}\bigl(\mathbf{E}\,x_1^2x_2^2\bigr)^{1/2}, \quad Y = y_{1,2}^2x_2^2,$$

combined with the estimates $\mathbf{E}\,x_2^2 \le N^{-1}$ and $\mathbf{E}\,x_1^2\,x_2^2 \le N^{-2}$ and the inequality $Y \le N^2\,y_{1,2}^4 + N^{-2}\,x_2^4$ which completes the proof of the lemma.

Let $m \le N$ and $\mathcal{B} \subset \{1,2,\dots,m\}$. For $k = 1,2,\dots,m$, put

$$\zeta(A_k) = \sum_{m+1 \le l \le n} (\nu_l - p)\,g_2(A_k,A_l).$$

Here ν_l, $l = 1,2,\dots,n$ denote i.i.d. Bernoulli random variables independent of $\overline{A}$ and such that $\mathbf{P}\{\nu_1 = 1\} = 1 - \mathbf{P}\{\nu_1 = 0\} = p$. Given $\{A_i,\,i \in \mathcal{B}\}$ let $A_{\mathcal{B}}^*$ denote a random variable uniformly distributed in the set $\{A_i,\,i \in \mathcal{B}\}$. By $\mathbf{E}_{\mathcal{B}}^*$ we denote the conditional expectation given all random variables but $A_{\mathcal{B}}^*$.

Lemma 5.3. *Assume that $\mathcal{B} \cap \{1,2,3,4\} = \emptyset$. Then*

$$\mathbf{E}^{1,2,3,4} \big(\mathbf{E}^*_{\mathcal{B}} \, \zeta^2(A^*_{\mathcal{B}})\big)^{3/2} \ll q^{3/2} \, N^{-3} \, \gamma_3. \tag{5.8}$$

For any $k = 1, 2, \dots m$ and $i = 3, 4, \dots, m$, we have

$$\mathbf{E}^* |\zeta(A_k)|^3 \ll p \, q \, (1 + n \, p \, q)^{1/2} \sum_{l=m+1}^{n} |g_2(A_k, A_l)|, \tag{5.9}$$

$$\mathbf{E}^{1,2} |\zeta(A_i)|^3 \ll q^{3/2} N^{-3} \gamma_3, \qquad \mathbf{E}^{1,2,i} |\zeta(A_i)|^3 \ll q^{3/2} N^{3/2} \mathbf{E}^i |g_2(A_i, A_n)|^3. \tag{5.10}$$

Proof of Lemma 5.3. Let $Z_1, \dots, Z_j$ denote independent and centered random variables. In what follows we apply the inequality

$$\mathbf{E}\big|Z_1 + \cdots + Z_j\big|^r \le c(r) \sum_{l=1}^{j} \mathbf{E} \, |Z_l|^r + c(r) \big(\sum_{l=1}^{j} \mathbf{E} \, Z_l^2\big)^{r/2}, \quad r \ge 2. \tag{5.11}$$

An application of Hölder's inequality yields

$$\Big(\sum_{l=m+1}^{n} g_2^2(a, A_l) \Big)^{3/2} \le (n - m)^{1/2} \sum_{l=m+1}^{n} |g_2(a, A_l)|^3, \qquad a \in \mathcal{A}. \tag{5.12}$$

In order to prove (5.9) we shall apply (5.11) with $r = 3$ to $Y_l = (\nu_l - p) \, g_2(A_k, A_l)$, $l = m+1, \dots, n$ conditionally, given $\overline{A}$. Furthermore, since $\mathbf{E}^* Y_l^2 = p \, q \, g_2^2(A_k, A_l)$, using (5.12) with $a = A_k$ we obtain (5.9).

Finally, let us show (5.8). By Jensen's inequality, $\big(\mathbf{E}^*_{\mathcal{B}} \zeta^2(A^*_{\mathcal{B}})\big)^{3/2} \le \mathbf{E}^*_{\mathcal{B}} |\zeta(A^*_{\mathcal{B}})|^3$. Applying (5.11) to $Y_l = (\nu_l - p) \, g_2(A^*_{\mathcal{B}}, A_l)$ conditionally, given $\overline{A}$ and $A^*_{\mathcal{B}}$, and then using (5.12) with $a = A^*_{\mathcal{B}}$ we get

$$\mathbf{E}^{1,2,3,4} \big(\mathbf{E}^*_{\mathcal{B}} \, \zeta^2(A^*_{\mathcal{B}})\big)^{3/2} \ll p \, q \, \big(1 + (p \, q \, n)^{1/2}\big) \sum_{l=m+1}^{n} \mathbf{E}^{1,2,3,4} |g_2(A^*_{\mathcal{B}}, A_l)|^3.$$

The inequality (5.8) follows from

$$\mathbf{E}^{1,2,3,4} |g_2(A^*_{\mathcal{B}}, A_l)|^3 = \mathbf{E}^{1,2,3,4} |g_2(5, A_l)|^3 \ll \mathbf{E} \, |g_2(A_5, A_l)|^3.$$

Since (5.10) is a corollary of (5.9) the proof of the lemma is complete.

5.2. Bounds for conditional characteristic functions.

Let $m \le N$, $\mathcal{B} \subset \{1, 2, \dots, m\}$. For t, s, x, $\tau \in R$ and $0 < d < \pi$, write

$$(5.13)\ \alpha = 2\,\pi\,\Big(\frac{4}{\Theta(1)} + 1\Big),\ \ \delta = \Theta(1)/40,\ \ \varkappa_{\mathcal{B}} = \alpha\,N\,\mathbf{E}^*_{\mathcal{B}}\zeta^2(A^*_{\mathcal{B}}),\ \ \mathbb{I}_{\mathcal{B}} = \mathbb{I}\big\{\varkappa_{\mathcal{B}} > \delta\big\},$$

$$\Psi_{\mathcal{B}} = g_{\mathcal{B}}(t)\,Z_{1,\mathcal{B}}^{1/2}, \quad g_{\mathcal{B}}(t) = \exp\{p\,q\,\delta\,\frac{|\mathcal{B}|}{N}\,t^2\}, \quad Z_{1,\mathcal{B}} = \prod_{k\in\mathcal{B}} u_1\big(t\,g_1(A_k) + s/\tau\big),$$

$$u_d(x) = 1 - 2^{-1}\,p\,q\,\Theta(d)\,x^2\,\chi_d(x), \quad v_d(x) = 1 + 2\,\pi\,p\,q\,d^{-1}\Big(\frac{4}{\Theta(d)} + 1\Big)\,x^2,$$

where $\chi_d(x) = \mathbb{I}\big\{\,|x| < \pi + d\big\}$. Recall that the expectation $\mathbf{E}^*_{\mathcal{B}}\zeta^2(A^*_{\mathcal{B}})$ has been defined before Lemma 5.3 above. The next lemma provides bounds for conditional expectations of random variables $Z_{\mathcal{B}}$ and $Y_{\mathcal{B}}$ (defined in Section 4).

Lemma 5.4. *Let $1 < m < n$ and let $\mathcal{B}, \mathcal{C} \subset \{1, \dots, m\}$. Assume that $\mathcal{C} \cap \mathcal{B} = \emptyset$ and write $\mathcal{D} = \mathcal{B} \cup \mathcal{C}$. Suppose that $\beta_2 = 1$. Then, for $|t| \le H_1$ and $|s/\tau| \le \pi$,*

$$Z_{\mathcal{B}} \le \mathbb{I}_{\mathcal{B}} + \Psi_{\mathcal{B}}, \qquad Y_{\mathcal{B}} \le \Psi_{\mathcal{B}}, \qquad Z_{\mathcal{B}}\,Y_{\mathcal{C}} \le \mathbb{I}_{\mathcal{B}} + \Psi_{\mathcal{D}}. \tag{5.14}$$

Furthermore,

$$\mathbf{E}^{n-3,\dots,n}\,\Psi_{\mathcal{B}} \le F_{\mathcal{B}} \qquad \textit{and} \qquad \mathbf{E}^{n-3,\dots,n}\,\Psi^2_{\mathcal{B}} \le F^2_{\mathcal{B}}. \tag{5.15}$$

Here $F_{\mathcal{B}} = \exp\Big\{-|\mathcal{B}|\,p\,q\,N^{-1}\,\big(t^2 + s^2/q\big)\,z\Big\}$, with $z = 4^{-1}\,\Theta(1)\,(9/10 - \eta)$ and η defined in (5.20).

Remark. Assume that (3.1) holds. If c_0 is sufficiently small, we have $\eta < 1/10$ and therefore, $z > \Theta(1)/5$. If $|\mathcal{B}| \ge m/4$ and $m = C_2\,q^{-1}\,n\,\varphi$, where φ is defined in (4.4) then for $C_1/q \le |t| \le H_1$ and $|s| \le \pi\,\tau$,

$$(5.16) \qquad F_{\mathcal{B}} \le \exp\big\{-r\,(t^2 + s^2/q)\,\varphi\big\} = (t^2 + s^2/q)^r, \quad r = C_2\,\Theta(1)/20.$$

Hence for any fixed $A > 0$ we can choose a (sufficiently large) C_2 such that $F_{\mathcal{B}} \le (t^2 + s^2/q)^A$.

If $|\mathcal{B}| \ge m/4$ where $m = C_3\,n\,\varphi$, and where φ is defined in (4.16) then

$$F_{\mathcal{B}} \le \exp\big\{-\frac{C_3\,z}{4}\,\frac{s^2}{s^2 + C_4}\,\ln(s^2 + C_4)\big\},$$

for $|t| \le C_1/q$ and $|s| \le \pi\,\tau$. Therefore

$$(5.17)\ \ F_{\mathcal{B}} \le \big(s^2 + C_4\big)^r,\ \ r = C_3\,\Theta(1)/40,\ \ \text{if } s^2 > C_4 \ \text{and}\ F_{\mathcal{B}} \le 1,\ \ \text{for } s^2 < C_4.$$

Hence, for any fixed $A > 0$ one may choose (sufficiently large) C_3 such that $F_{\mathcal{B}} \le C(A)\,(s^2 + C_4)^A$. Here $C(A)$ is a positive constant depending only on A.

Proof of Lemma 5.4. Take $d = 1$, $x = t\,g_1(A_k) + s/\tau$ and $y = t\,\zeta(A_k)$ in Lemma 5.5 to get

$$\prod_{k\in\mathcal{B}} \Big|\beta\big(t\,g_1(A_k) + s/\tau + t\,\zeta(A_k)\big)\Big| \le Z_{1,\mathcal{B}}^{1/2}\, Z_{2,\mathcal{B}}^{1/2}, \qquad Z_{2,\mathcal{B}} = \prod_{k\in\mathcal{B}} v_1\big(t\,\zeta(A_k)\big).$$

Hence, $Z_{\mathcal{B}} \le Z_{1,\mathcal{B}}^{1/2}\, Z_{2,\mathcal{B}}^{1/2}$. Similarly,

$$Y_{\mathcal{B}} = \prod_{k\in\mathcal{B}} \Big|\beta\big(t\,g_1(A_k) + s/\tau\big)\Big| \le Z_{1,\mathcal{B}}^{1/2}. \tag{5.18}$$

An application of the inequality $1 + x \le e^x$ yields $Z_{2,\mathcal{B}} \le \exp\big\{p\,q\,|\mathcal{B}|\,N^{-1}\,t^2\,\varkappa_{\mathcal{B}}\big\}$. We have $\big(1 - \mathbb{I}_{\mathcal{B}}\big)\, Z_{2,\mathcal{B}}^{1/2} \le g_{\mathcal{B}}(t)$ and $Z_{\mathcal{B}} \le 1$. Therefore

$$Z_{\mathcal{B}} = \mathbb{I}_{\mathcal{B}}\, Z_{\mathcal{B}} + \big(1 - \mathbb{I}_{\mathcal{B}}\big)\, Z_{\mathcal{B}} \le \mathbb{I}_{\mathcal{B}} + g_{\mathcal{B}}(t)\, Z_{1,\mathcal{B}}^{1/2},$$

and the first inequality of (5.14) follows. Similarly,

$$Z_{\mathcal{B}} Y_{\mathcal{C}} \le \mathbb{I}_{\mathcal{B}} + \big(1 - \mathbb{I}_{\mathcal{B}}\big)\, Z_{\mathcal{B}} Y_{\mathcal{C}} \le \mathbb{I}_{\mathcal{B}} + g_{\mathcal{B}}(t)\, Z_{1,\mathcal{B}\cup\mathcal{C}}^{1/2} \le \mathbb{I}_{\mathcal{B}} + g_{\mathcal{B}\cup\mathcal{C}}(t)\, Z_{1,\mathcal{B}\cup\mathcal{C}}^{1/2},$$

thus proving the last inequality of (5.14). The second one follows from (5.18), since $g_{\mathcal{B}}(t) \ge 1$.

Finally, we shall prove (5.15). Given A_j, $j = n-3, \dots, n$ let $\widehat{A}^*$ denote a random variable which is uniformly distributed in the set $\mathcal{A} \setminus \{A_j,\, j = n-3, \dots, n\}$ and let $\widehat{\mathbf{E}}^*$ denote the conditional expectation given all random variables but $\widehat{A}^*$. By Hölder's inequality and Theorem 4 of Hoeffding (1963),

$$\mathbf{E}^{n-3,\dots,n}\, \Psi_{\mathcal{B}} \le g_{\mathcal{B}}(t)\, \Big(\mathbf{E}^{n-3,\dots,n}\, Z_{1,\mathcal{B}}\Big)^{1/2} \le g_{\mathcal{B}}(t)\, \big(w^{|\mathcal{B}|}\big)^{1/2}, \tag{5.19}$$

where $w = \widehat{\mathbf{E}}^*\, u_1\big(t\,g_1(\widehat{A}^*) + s/\tau\big)$. Note, that for $|t| < H_1$ and $\big|s/\tau\big| \le \pi$, we have

$$\chi_1\big(t\,g_1(\widehat{A}^*) + s/\tau\big) \ge \mathbb{I}_1(\widehat{A}^*), \qquad \text{where} \qquad \mathbb{I}_1(a) = \mathbb{I}\big\{H_1\,\big|g_1(a)\big| < 1\big\}.$$

Therefore, $w \le 1 - 2^{-1}\,p\,q\,\Theta(1)\,S$. Here $S = \widehat{\mathbf{E}}^*\,\big(t\,g_1(\widehat{A}^*) + s/\tau\big)^2\, \mathbb{I}_1(\widehat{A}^*)$. Observe, that $S = n\,(n-4)^{-1}(S_1 - S_2)$, where

$$S_1 = \mathbf{E}\,\big(t\,g_1(A_1) + s/\tau\big)^2\, \mathbb{I}_1(A_1), \ \text{ and } \ S_2 = n^{-1} \sum_{n-3\le k\le n} \big(t\,g_1(A_k) + s/\tau\big)^2\, \mathbb{I}_1(A_k).$$

We shall bound S_2 from above, i.e.,

$$S_2 \le 8\, n^{-1}\, s^2/\tau^2 + 2\, n^{-1}\, t^2\, L, \quad \text{where} \quad L = \sum_{n-3 \le k \le n} g_1^2(A_k).$$

By Hölder's inequality,

$$L \le 4^{1/3} \Big(\sum_{n-3 \le k \le n} \big| g_1(A_k) \big|^3 \Big)^{2/3} \le 2\, N^{-1}\, n^{2/3}\, \beta_3^{2/3}.$$

Hence,

$$S_2 \le \big(N^{-1}\, t^2 + (s/\tau)^2 \big) \big(8\, n^{-1} + 4\, \big(\beta_3^2\, n^{-1} \big)^{1/3} \big).$$

An application of lemma 5.6 yields $S_1 \ge \big(N^{-1} t^2 + s^2/\tau^2 \big)\,(1 - 2\, b_1)$. Substituting the lower and upper bounds for S_1 and S_2 respectively we obtain

$$S \ge \big(N^{-1}\, t^2 + s^2/t^2 \big)\,(1 - \eta), \qquad \eta = 2\, b_1 + 8\, n^{-1} + 4\, \big(\beta_3^2\, n^{-1} \big)^{1/3}. \tag{5.20}$$

Therefore,

$$w \le 1 - 2^{-1}\, p\, q\, \Theta(1)\, \big(N^{-1}\, t^2 + s^2/\tau^2 \big)\,(1 - \eta).$$

This inequality in combination with the inequality $1 + x \le e^x$ and (5.19) yields the first inequality of (5.15). To prove the second inequality of (5.15) we directly apply Hoeffding's (1963) Theorem 4, i.e., instead of (5.19) we use $\mathbf{E}^{n-3,\ldots,n}\, \Psi_{\mathcal{B}}^2 \le g^2(t)\, w^{|\mathcal{B}|}$. This completes the proof of Lemma 5.4.

Lemma 5.5. *For each $0 \le d < \pi$ and $x,\, y \in R$, we have*

$$\big|\, \beta(x+y) \big|^2 \le u_d(x)\, v_d(y).$$

Proof of Lemma 5.5. Observe, that for $|x| \ge \pi + d$, the lemma follows from the inequality $\big|\beta(x+y)\big| \le 1$.

Assume that $|x| < \pi + d$. We have $\big|\beta(\eta)\big|^2 = 1 - 2\, p\, q\, \big(1 - \cos(\eta)\big)$. Put $\eta = x + y$. The mean value theorem combined with the inequality $|x\, y| \le c\, x^2 + c^{-1}\, y^2$, $c > 0$, yields

$$\begin{aligned} \big|\cos(x+y) - \cos(x)\big| &\le \big|\mathbf{E}\, \sin(x + \theta_1\, y)\, y\big| \\ &\le \big(|x| + |y|\big)\, |y| \le c\, x^2 + \big(c^{-1} + 1\big)\, y^2. \end{aligned}$$

Therefore,

$$\big|\beta(x+y)\big|^2 \le 1 - 2\, p\, q\, \big(1 - \cos(x) - c\, x^2 - \big(c^{-1} + 1\big)\, y^2\big).$$

Using the inequality $1 - 2\,p\,q\,\big(1-\cos(x)\big) \le 1 - p\,q\,x^2\,\Theta(d)$, see (5.26), we obtain

$$\big|\beta(x+y)\big|^2 \le w + z, \quad w = 1 - p\,q\,\big(\Theta(d) - 2\,c\big)\,x^2, \quad z = \big(c^{-1}+1\big)\,2\,p\,q\,y^2.$$

But $1 - p\,q\,x^2\,\Theta(d) \ge d/\pi$, for $|x| \le \pi + d$. Hence, $w > d/\pi$ and therefore $w + z \le w\,\big(1 + \pi\,d^{-1}\,z\big)$. Choosing $c = \Theta(d)/4$ completes the proof of the lemma.

The inequality given in the next lemma as well as inequalities (5.21) and (5.24) below can be found in Höglund (1978). For easier reference we include a proof of these inequalities.

Lemma 5.6. *Assume that $\beta_2 = 1$. For any $0 \le d < \pi$, write*

$$S_1 = \mathbf{E}\,\big(t\,g_1(A_1) + s/\tau\big)^2\,\mathbb{I}_d(A_1), \qquad \mathbb{I}_d(a) = \mathbb{I}\big\{H_1\,\big|g_1(a)\big| < d\big\}.$$

We have $S_1 \ge \big(N^{-1}\,t^2 + s^2/\tau^2\big)\,\big(1 - 2\,\max\{\,\frac{b_1}{d}\,,\,\frac{b_1^2}{d^2}\,\}\big)$.

Remark. Since $b_1 < b_2 < 1$, we have $\max\{b_1, b_1^2\} = b_1$ and $\max\{\,\frac{b_1}{b_2}\,,\,\frac{b_1^2}{b_2^2}\,\} = \frac{b_1}{b_2}$.

Proof of Lemma 5.6. Write $\mathcal{K}(d) = \big\{k\,:\,H_1\,|g_1(a_k)| \ge d\big\}$. We have

$$|\mathcal{K}(d)| = \sum_{k\in\mathcal{K}(d)} 1 \le \sum_{k\in\mathcal{K}(d)} \big|d^{-1}\,g_1(a_k)\,H_1\big|^r \le n\,\beta_r\,\beta_3^{-r}\,b_1^r/d^r, \quad r = 2,3. \tag{5.21}$$

Furthermore, it follows from $\beta_2 = 1$ and $\mathbf{E}\,g_1(A_1) = 0$ that

$$S_1 = N^{-1}\,t^2 + s^2/\tau^2 - W, \quad \text{where} \quad W = n^{-1}\sum_{k\in\mathcal{K}(d)} \big(t\,g_1(a_k) + s/\tau\big)^2. \tag{5.22}$$

We have $W \le W_1 + W_2$, where

$$W_1 = 2\,n^{-1}|\mathcal{K}(d)|\,s^2/\tau^2 \quad \text{and} \quad W_2 = 2\,t^2\,n^{-1}\sum_{k\in\mathcal{K}(d)} g_1^2(a_k). \tag{5.23}$$

An application of (5.21) with $r = 2$ yields $W_1 \le 2\,\beta_3^{-2}(b_1/d)^2\,(s/\tau)^2$. Combining Hölder's inequality and (5.21) with $r = 3$, we get

$$\sum_{k\in\mathcal{K}(d)} g_1^2(a_k) \le \Big(\sum_{k\in\mathcal{K}(d)} \big|g_1(a_k)\big|^3\Big)^{2/3}\,|\mathcal{K}(d)|^{1/3} \le N^{-1}\,n\,(b_1/d). \tag{5.24}$$

Using the bounds for W_1 and W_2 in (5.22) and $\beta_3 \ge \beta_2^{3/2} = 1$ we obtain the lemma.

5.3. Hoeffding's Theorem 4 and Höglund's inequalities.

Let a population $\mathcal{P}$ consists of n values $p_1, \dots, p_n$. Let $\mathcal{X}_1, \dots, \mathcal{X}_N$ denote a random sample *without* replacement from $\mathcal{P}$ and let $\mathcal{Y}_1, \dots, \mathcal{Y}_N$ denote a random sample *with* replacement from $\mathcal{P}$ (which are in fact i.i.d.).

Theorem (Hoeffding (1963)). *If the function $f(x)$ is continuous and convex then*

$$\mathbf{E}\, f\big(\sum_{k=1}^{N} \mathcal{X}_k\big) \le \mathbf{E}\, f\big(\sum_{k=1}^{N} \mathcal{Y}_k\big).$$

The following inequalities are proved in Höglund (1978). Given integers $1 \le N \le n$, write $s = N/n$ and $t = 1 - s$. We have

$$\sqrt{\pi}/2 \le C_n^N\, s^N\, t^{n-N}\, \big(2\,\pi\, s\, t\, n\big)^{1/2} \le 1. \tag{5.25}$$

Furthermore,

$$1 - \cos v \ge \frac{1}{2}\, v^2\, \Theta(u), \quad \text{for} \quad |v| \le \pi + u \quad \text{and} \quad 0 \le u \le \pi. \tag{5.26}$$

In particular, for $|v| \le \pi + u$ and $0 \le u \le \pi$, we have

$$|\beta(v)| \le \exp\big\{-\tfrac{p\,q}{2}\, v^2\, \Theta(u)\big\}. \tag{5.27}$$

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.

Bentkus. V. and Götze, F., *On minimal moment assumptions in Berry–Esseen theorems for U-statistics*, Teor. Veroyatnost. i Primenen. **40** (1995), 596–614.

Bentkus. V., Götze, F. and Zitikis, R., *Lower estimates of the convergence rate for U-statistics*, Ann. Probab. **22** (1994), 1707–1714.

Bentkus. V., Götze, F. and van Zwet, W. R., *An Edgeworth expansion for symmetric Statistics* (1994), Preprint 94 - 020, SFB 343, Universität Bielefeld. To appear in Ann. Statist.

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

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.

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.

Bikelis, A., *On the estimation of the remainder term in the central limit theorem for samples from finite populations.*, Studia Sci. Math. Hungar. **4** (1974), 345–354. (In Russian).

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

Callaert, H. and Janssen, P., *The Berry–Esseen theorem 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.

Chan, Y. -K. and Wierman, J., *On the Berry–Esseen theorem for U-statistics*, Ann. Probab. **5** (1977), 136–139.

Does, R.J.M.M., *An Edgeworth expansion for simple linear rank statistics under the null-hypothesis*, Ann. Statist. **11** (1983), 607–624.

Erdös P. and Rényi A., *On the central limit theorem for samples from a finite population*, Publ. Math. Inst. Hungar. Acad. Sci. **4** (1989), 49–61.

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

Helmers, R., *Edgeworth expansions for linear combinations of order statistics* (1982), Mathematical Centre Tracts 105,, Mathematisch 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 (S.S. Gupta and J.O. Berger, eds), Academic Press, New York,, 497–512.

Hoeffding, W., *A class of statistics with asymptotically normal distributions*, Ann. Math. Statist. **19** (1948), 293–325.

Hoeffding, W., *Probability inequalities for sums of bounded random variables*, J. Amer. Stat. Assoc. **58** (1963), 13–30.

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

Ivanoff,J.B. and Weber, N.C., *Weak convergence of row and column exchangeable arrays*, Stochastics Rep. **40** (1992), 1–22.

G. Jogesh Babu and Kesar Singh, *Edgeworth expansions for sampling without replacement from finite populations*, J. Multivar. Analysis. **18** (1990), 390–404.

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.

Kokic P.N. and Weber N.C., *Rates of strong convergence for U-statistics in finite populations*, J. Austral. Math. Soc. Ser. A **50** (1991), 468–480.

Nandi, H. K. and Sen, P.K., *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* (1975), Springer, New York.

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

Robins J., *An asymptotic expansion for samples from a finite population*, Ann. Statist. **6** (1978), 1005–10011.

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

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

Zhao Lin Cheng and Chen Xi Ru, *Berry–Esseen bounds for finite-population U-statistics*, Sci.Sinica Ser. A **30** (1987), 113–127.

Zhao Lin Cheng and Chen Xi Ru, *Normal approximation for finite-population U-statistics*, Acta mathematicae applicatae Sinica **6** (1990), 263–272.

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

van Zwet, W.R., *On the Edgeworth expansion for the simple linear rank statistic*, Coll. Math. Soc. J. Bolyai **32** (1982), B.V. Gnedenko, M.L.Puri, I.Vincze (eds)., North-Holland, Amsterdam, 889–909.

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

Friedrich Götze
Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
33501 Bielefeld 1
Germany
E-mail address: goetze@mathematik.uni-bielefeld.de