Probability Theory and Related Fields (2022) 183:1153–1235
https://doi.org/10.1007/s00440-022-01144-x

Edgeworth approximations for distributions of symmetric statistics

Mindaugas Bloznelis[1] · **Friedrich Götze**[2]

*In memoriam Willem Rutger van Zwet *March 31, 1934 †July 2, 2020.*

Received: 29 March 2021 / Revised: 29 January 2022 / Accepted: 7 May 2022 /
Published online: 17 June 2022

Abstract

We study the distribution of a general class of asymptotically linear statistics which are symmetric functions of N independent observations. The distribution functions of these statistics are approximated by an Edgeworth expansion with a remainder of order $o(N^{-1})$. The Edgeworth expansion is based on Hoeffding's decomposition which provides a stochastic expansion into a linear part, a quadratic part as well as smaller higher order parts. The validity of this Edgeworth expansion is proved under Cramér's condition on the linear part, moment assumptions for all parts of the statistic and an optimal dimensionality requirement for the non linear part.

Keywords Edgeworth expansion · Littlewood–Offord problem · Concentration in Banach spaces · Symmetric statistic · U-statistic · Hoeffding decomposition

Mathematics Subject Classification Primary 62E20; Secondary 60F05

Research funded in part by the German Research Foundation—SFB 1283/2 2021—317210226.

✉ Friedrich Götze
goetze@math.uni-bielefeld.de

Mindaugas Bloznelis
mindaugas.bloznelis@mif.vu.lt

1 Faculty of Mathematics and Informatics, Vilnius University, Vilnius, Lithuania

2 Faculty of Mathematics, University of Bielefeld, Bielefeld, Germany

1 Introduction and results

1.1 Introduction

Let $X,\ X_1, X_2, \ldots, X_N$ be independent and identically distributed random variables taking values in a measurable space $(\mathcal{X}, \mathcal{B})$. Let P_X denotes the distribution of X on $(\mathcal{X}, \mathcal{B})$. We assume that $\mathbb{T}(X_1, \ldots, X_N)$ is a symmetric function of its arguments (symmetric statistic, for short). Furthermore, we assume that the moments $\mathbf{E}\mathbb{T}$ and $\sigma_{\mathbb{T}}^2 := \mathbf{Var}\mathbb{T}$ are finite. A function of observations $X_1, \ldots, X_N$ is called *linear statistic* if it can be represented by a sum of functions depending on a single observation only. Many important statistics are non linear, but can be approximated by a linear statistic. We call these statistics *asymptotically linear*. The central limit theorem and the normal approximation with rate $O(N^{-1/2})$ extend to the class of asymptotically linear statistics as well. Our approach in studying the distribution of this class of statistics in the statistically relevant case of asymptotic normal $\mathbb{T}$ is based on Hoeffding's decomposition of $\mathbb{T}$, see Hoeffding [31], Efron and Stein [21] and van Zwet [37]. Hoeffding's decomposition expands $\mathbb{T}$ into the series of centered and mutually uncorrelated U-statistics of increasing order

$$\mathbb{T} = \mathbf{E}\mathbb{T} + \frac{1}{N^{1/2}} \sum_{1 \le i \le N} g(X_i) + \frac{1}{N^{3/2}} \sum_{1 \le i < j \le N} \psi(X_i, X_j) + \frac{1}{N^{5/2}} \sum_{1 \le i < j < k \le N} \chi(X_i, X_j, X_k) + \ldots.$$

Let L, Q and K denote the first, the second and the third sum. We call L the linear part, Q the quadratic part and K the cubic part of the decomposition. We shall consider a general situation where the kernel $\mathbb{T} = \mathbb{T}^{(N)}$, the space $(\mathcal{X}, \mathcal{B}) = (\mathcal{X}^{(\mathcal{N})}, \mathcal{B}^{(\mathcal{N})})$ and the distribution $P_X = P_X^{(N)}$ all depend on N as $N \to \infty$. In order to keep the notation simple we drop the subscript N in what follows. An improvement over the normal approximation is obtained by using Edgeworth expansions for the distribution function $\mathbb{F}(x) = \mathbf{P}\{\mathbb{T} - \mathbf{E}\mathbb{T} \le \sigma_{\mathbb{T}} x\}$. For this purpose we write Hoeffding's decomposition in the form

$$\mathbb{T} - \mathbf{E}\mathbb{T} = L + Q + K + R, \tag{1}$$

where R denotes the remainder. For a number of important examples of asymptotically linear statistics we have $R/\sigma_{\mathbb{T}} = o_P(N^{-1})$ (in probability) as $N \to \infty$. Therefore, the U-statistic $\sigma_{\mathbb{T}}^{-1}(L + Q + K)$ can be viewed as a stochastic expansion of $(\mathbb{T} - \mathbf{E}\mathbb{T})/\sigma_{\mathbb{T}}$ up to the order $o_P(N^{-1})$.

Furthermore, a so-called Edgeworth expansion of $\sigma_{\mathbb{T}}^{-1}(L + Q + K)$ can be used to approximate $\mathbb{F}(x)$ by a smooth distribution function $G(x)$ as defined in (2) below depending on N and moments of $\mathbb{T}$. A two term Edgeworth expansion of the distribution function of $\sigma_{\mathbb{T}}^{-1}(L + Q + K)$ is given by

$$G(x) = \Phi(x) - \frac{1}{\sqrt{N}}\frac{\kappa_3}{6}(x^2-1)\Phi'(x)$$
$$-\frac{1}{N}\Big(\frac{\kappa_3^2}{72}(x^5-10x^3+15x)\Phi'(x) + \frac{\kappa_4}{24}(x^3-3x)\Phi'(x)\Big). \qquad (2)$$

Here Φ respectively Φ' denote the standard normal distribution function and its derivative. Furthermore, we introduce $\sigma^2 = \mathbf{E}g^2(X_1)$ and

$$\kappa_3 = \sigma^{-3}\Big(\mathbf{E}g^3(X_1) + 3\mathbf{E}g(X_1)g(X_2)\psi(X_1, X_2)\Big),$$
$$\kappa_4 = \sigma^{-4}\Big(\mathbf{E}g^4(X_1) - 3\sigma^4 + 12\mathbf{E}g^2(X_1)g(X_2)\psi(X_1, X_2)$$
$$+ 12\mathbf{E}g(X_1)g(X_2)\psi(X_1, X_3)\psi(X_2, X_3)$$
$$+ 4\mathbf{E}g(X_1)g(X_2)g(X_3)\chi(X_1, X_2, X_3)\Big).$$

Our main result, Theorem 1 below, establishes a bound $o(N^{-1})$ for the Kolmogorov distance between $\mathbb{F}(x)$ and $G(x)$:

$$\Delta = \sup_{x\in\mathbb{R}} |\mathbb{F}(x) - G(x)| = o(N^{-1}). \qquad (3)$$

Valid expansions of this type were shown by Cramér [19] for sums of independent random variables X_j and later on for the Student statistic (which is of type (1)) by Kai-Lai Chung [18]. A new impetus for studying higher order approximations in statistic was given by the fundamental paper of Hodges and Lehmann on deficiency [30], where they compared the power of two tests based on N and N' observations respectively and where $N' - N = o(N)$ as $N \to \infty$. They suggested a program of comparisons of the power of tests, estimators and confidence regions based on classical parametric and non parametric symmetric statistics e.g. using ranks and ordered samples. They noted that this would require going beyond Gaussian limit theorems to asymptotic expansions to order N^{-1}. For more details on the statistical relevance and the related development of asymptotic methods we refer to the review paper in memory of Willem van Zwet [12].

Now we discuss the principal contribution of this paper: the minimal smootness and structural conditions under which approximation (3) holds. Let us emphasize that any $\mathbb{F}$ satisfying (3) cannot have fluctuations/increments of order $\Theta(N^{-1})$ in the intervals of size $o(N^{-1})$ because G is a differentiable function with all derivatives bounded. We focus on the conditions that guarrantee the necessary level of smoothness of the distribution of $\mathbb{T}$. In the case of linear statistic $\mathbb{T} = \mathbf{E}\mathbb{T} + L$ the necessary smoothness of $\mathbb{F}$ is ensured by the classical Cramer condition

$$\limsup_{|t|\to\infty} |\mathbf{E}\exp\{itg(X_1)\}| < 1. \qquad (C)$$

This condition excludes, in particular, lattice distributions, for which approximation (3) obviously fails. We note that condition (C) can be weakened to cover some special classes of discrete distribution which are sufficiently non-lattice distributed, see e.g. Bickel and Robinson [13], Angst and Poly [1] or Bobkov [14] for almost sure choices of such non-lattice discrete distributions.

Since the class of symmetric statistic should include the case of linear statistics we require a Cramér type condition but on the linear part of the statistic only, see (7). Interestingly, this condition together with appropriate moment conditions on various parts of the decomposition (1) guarantees already an approximation error $\Delta = O(N^{-1})$ for general symmetric statistics (see [4]). But (7) is not sufficient for the desired error bound $o(N^{-1})$ even for U statistics of degree two, see Example 1 below. The reason why (7) alone is not sufficient for the approximation accuracy $\Delta = o(N^{-1})$ is due to the potential occurrence of a very special relation between the linear and quadratic parts L and Q that fosters an approximate lattice structure as shown in Example 1. Namely, the quadratic part of the U statistic in Example 1 has a factorizable kernel ψ of the form $\psi_h(X_1, X_2) = h(X_1)g(X_2)+g(X_1)h(X_2)$, $h-$measurable. The following structural condition (4) (introduced in the unpublished manuscript by Götze and van Zwet [25]) avoids such counterexamples by separating (in L^2 distance) the random variable $\psi(X_1, X_2)$ from any random variable of the form $\psi_h(X_1, X_2)$. Note that the L^2 distance $\mathbf{E}(\psi(X_1, X_2) - \psi_h(X_1, X_2))^2$ is minimized by $h(x) = b(x)$, where

$$b(x) = \sigma^{-2}\mathbf{E}\big(\psi(X_1, X_2)g(X_2)\big|X_1 = x\big) - (\kappa/2\sigma^4)g(x).$$

Here $\kappa = \mathbf{E}\psi(X_1, X_2)g(X_1)g(X_2)$. Therefore, we will assume that, for some absolute constant $\delta_* > 0$, we have

$$\mathbf{E}\Big(\psi(X_1, X_2) - \big(b(X_1)g(X_2) + b(X_2)g(X_1)\big)\Big)^2 \geq \delta_*^2\sigma_{\mathbb{T}}^2. \tag{4}$$

The main contribution of the present paper consist of a proof that condition (4) will indeed ensure the desired bound $\Delta = o(N^{-1})$. The proof is based on a careful investigation of the size distribution for $|t| > N^{1-\nu}$ of the absolute values of conditional Fourier transforms of symmetric statistics that is the landscape of its maxima when imposing Cramer's condition (7) and the structural condition (4). Here new methods are used for studying this landscape in the frequency t as well in the random function representing the conditioning. For the latter variable a combinatorial argument of Kleitman on symmetric partitions for the Littlewood–Offord problem in Banach spaces (see [15]) is used.

A short outline of the approach is given at the beginning of Sect. 2, where we focus on the use of condition (4).

1.2 Results

Let us state our main result Theorem 1.

Moment conditions We will assume that, for some absolute constants $0 < A_* < 1$ and $M_* > 0$ and numbers $r > 4$ and $s > 2$, we have

$$
\begin{aligned}
&\mathbf{E}g^2(X_1) > A_*\sigma_{\mathbb{T}}^2, \quad \mathbf{E}|g(X_1)|^r < M_*\sigma_{\mathbb{T}}^r,\\
&\mathbf{E}|\psi(X_1, X_2)|^r < M_*\sigma_{\mathbb{T}}^r, \quad \mathbf{E}|\chi(X_1, X_2, X_3)|^s < M_*\sigma_{\mathbb{T}}^s.
\end{aligned} \tag{5}
$$

These moment conditions refer to the linear, quadratic and cubic part of $\mathbb{T}$. In order to control the remainder R of the approximation (1) we use moments of differences introduced in Bentkus, Götze and van Zwet [4], see also van Zwet [37]. Define, for $1 \le i \le N$,

$$
D_i\mathbb{T} = \mathbb{T} - \mathbf{E}_i\mathbb{T}, \qquad \mathbf{E}_i\mathbb{T} := \mathbf{E}(\mathbb{T}|X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_N).
$$

A subsequent application of difference operations $D_i, D_j, \dots,$ (the indices $i, j, \dots,$ are all distinct) produce higher order differences, like

$$
D_iD_j\mathbb{T} := D_i(D_j\mathbb{T}) = \mathbb{T} - \mathbf{E}_i\mathbb{T} - \mathbf{E}_j\mathbb{T} + \mathbf{E}_i\mathbf{E}_j\mathbb{T}.
$$

For $m = 1, 2, 3, 4$ write $\Delta_m^2 = \mathbf{E}|N^{m-1/2}D_1D_2\cdots D_m\mathbb{T}|^2$.

We will assume that for some absolute constant $D_* > 0$ and number $\nu_1 \in (0, 1/2)$ we have

$$
\Delta_4^2/\sigma_{\mathbb{T}}^2 \le N^{1-2\nu_1}D_* \tag{6}
$$

For a number of important examples of asymptotically linear statistics the moments Δ_m^2 are evaluated or estimated in [4]. Typically we have $\Delta_m^2/\sigma_{\mathbb{T}}^2 = O(1)$ for some m. Therefore, assuming that (6) holds uniformly in N as $N \to \infty$, we obtain from the inequality $\mathbf{E}R^2 \le N^{-3}\Delta_4^2$, see (167) (see "Appendix"), that $R/\sigma_{\mathbb{T}} = O_P(N^{-1-\nu_1})$. Furthermore, assuming that (5), (6) hold uniformly in N as $N \to \infty$, we obtain from (167), (166), see "Appendix", that $\sigma^2/\sigma_{\mathbb{T}}^2 = (1 - O(N^{-1}))$.

Cramér type smoothness condition We introduce the function

$$
\rho(a, b) = 1 - \sup\{|\mathbf{E}\exp\{itg(X_1)/\sigma\}| : a \le |t| \le b\}
$$

and assume that, for some $\delta > 0$ and $\nu_2 > 0$, we have

$$
\rho(\beta_3^{-1}, N^{\nu_2+1/2}) \ge \delta. \tag{7}
$$

Here $\beta_3 = \sigma^{-3}\mathbf{E}|g(X_1)|^3$. Define $\nu = 600^{-1}\min\{\nu_1, \nu_2, s-2, r-4\}$.

Theorem 1 *Assume that for some absolute constants $A_*, M_*, D_* > 0$ and numbers $r > 4, s > 2, \nu_1, \nu_2 > 0$ and $\delta, \delta_* > 0$, the conditions (5), (6), (7), (4) hold. Then there exists a constant $C_* > 0$ depending only on $A_*, M_*, D_*, r, s, \nu_1, \nu_2, \delta, \delta_*$ such that*

$$
\Delta \le C_*N^{-1-\nu}\big(1 + \delta_*^{-1}N^{-\nu}\big).
$$

Remark 1 The value of $\nu = 600^{-1}\min\{\nu_1, \nu_2, s-2, r-4\}$ is far from being optimal. Furthermore, the moment conditions (5) and (6) are not the weakest possible that would ensure the approximation of order $o(N^{-1})$. The condition (5) can likely be reduced to the moment conditions that are necessary to define Edgeworth expansion terms κ_3 and κ_4, similarly, (6) can be reduced to $\Delta_4^2/\sigma_{\mathbb{T}}^2 = o(N^{-1})$. No effort was made to obtain the result under the optimal conditions. This would increase the complexity of the proof which is already rather involved.

Remark 2 Condition (4) can be relaxed. Assume that for some absolute constant G_* we have

$$\mathbf{E}\Big(\psi(X_1, X_2) - \big(b(X_1)g(X_2) + b(X_2)g(X_1)\big)\Big)^2 \geq N^{-2\nu} G_* \sigma_{\mathbb{T}}^2. \tag{8}$$

The bound of Theorem 1 holds if we replace (4) by this weaker condition. In this case we have $\Delta \leq C_* N^{-1-\nu}$, where the constant C_* depends on $A_*, D_*, G_*, M_*, r, s, \nu_1, \nu_2, \delta$.

In the particular case of U statistics of degree three (the case where $R \equiv 0$ in (1)) the proof of Theorem 1 has been outlined in the unpublished manuscript by Götze and van Zwet [25]. We provide a complete and more readable version of the arguments sketched in that preprint and extend them to a general class of symmetric statistics. In the same paper [25], see as well [4], it was shown that moment conditions (like (5), (6)) together with Cramér's condition (like (7)) do not suffice for the bound $\Delta = o(N^{-1})$. For convenience we state this result in Example 1 below.

Example 1 Let $X_1, X_2, \ldots$ be independent random variables uniformly distributed on the interval $(-1/2, 1/2)$. Define $T_N = (W_N + N^{-1/2} V_N)(1 - N^{-1/2} V_N)$, where $V_N = N^{-1/2} \sum \{N^{1/2} X_j\}$ and $W_N = N^{-1} \sum [N^{1/2} X_j]$. Here $[x]$ denotes the nearest integer to x and $\{x\} = x - [x]$.

Assume that $N = m^2$, where m is odd. We have, by the local limit theorem,

$$\mathbf{P}\{W_N = 1\} \geq c N^{-1} \qquad \text{and} \qquad \mathbf{P}\{|V_N| < \delta\} > c\delta, \qquad 0 < \delta < 1,$$

where $c > 0$ is an absolute constant. From these inequalities it follows by the independence of W_N and V_N, that $\mathbf{P}\{1 - \delta^2 N^{-1} \leq T_N \leq 1\} \geq c^2 \delta N^{-1}$.

The example defines a sequence of U-statistics $\mathbb{T}_N$ whose distribution functions $\mathbb{F}_N$ have $O(N^{-1})$ sized increments in a particular interval of length $o(N^{-1})$. These fluctuations of magnitude $O(N^{-1})$ appear as a result of a nearly lattice structure induced by the interplay between the (smooth) linear part and the quadratic part.

1.3 Earlier work

There is a rich literature devoted to normal approximation and Edgeworth expansions for various classes of asymptotically linear statistics (see e.g. Babu and Bai [2], Bai and Rao [3], Bentkus, Götze and van Zwet [4], Bhattacharya and Ghosh [8, 9], Bhattacharya and Rao [7], Bickel [10], Bickel, Götze and van Zwet [11], Callaert, Janssen and Veraverbeke [16], Chibisov [17], Hall [28], Helmers [29], Petrov [33], Pfanzagl [34], Serfling [35], etc.

A wide class of statistics can be represented as functions of sample means of vector variables. Edgeworth expansions of such statistics can be obtained by applying the multivariate expansion to corresponding functions, see Bhattacharya and Ghosh [8, 9]. In their work the crucial Cramér condition (C) is assumed on the joint distribution of all the components of a vector which may be too restrictive in cases where some components have a negligible influence on the statistic. More often only one or a few of the components satisfy a conditional version of condition (C). Bai and Rao [3], Babu and Bai [2] established Edgeworth expansions for functions of sample means under such a conditional Cramér condition. This approach exploits the smoothness of the distribution of a random vector as well as the smoothness of the function defining the statistic. In particular this approach needs a class of statistics which are smooth functions of observations or can be approximated by such functions via Taylor's expansion, see also Chibisov [17]. The respective condition (6) of the present paper is expressed in terms of moments of iterated differences Δ_m and does not assume Taylor's expansion.

Let us note that generally the smoothness of the distribution function of $\mathbb{T}$ may have little to do with the smoothness of the function $\mathbb{T}(X_1, \ldots, X_N)$ of observations $X_1, \ldots, X_N$. Just take Gini's mean difference $\sum_{i<j} |X_i - X_j|$ with absolutely continuous X_i for example. Another interesting example is about Studentization, when it enchances the smoothness of the distribution function of a sum of lattice random variables dramatically, see [26]. Our Theorem 1 shows, in particular, that structural condition (4) together with (7) guarantee the smoothness of the distribution of $\mathbb{T}$ necessary for the bound $\Delta = o(N^{-1})$.

In order to compare Theorem 1 with earlier results of *similar* nature let us consider the case of U-statistics of degree two

$$\mathbb{U} = \frac{\sqrt{N}}{2}\binom{N}{2}^{-1} \sum_{1\le i<j\le N} h(X_i, X_j), \tag{9}$$

where $h(\cdot, \cdot)$ denotes a (fixed) symmetric kernel. Assume for simplicity of notation and without loss of generality that $\mathbf{E}h(X_1, X_2) = 0$. Write $h_1(x) = \mathbf{E}(h(X_1, X_2)|X_1 = x)$ and assume that $\sigma_h^2 > 0$, where $\sigma_h^2 = \mathbf{E}h_1^2(X_1)$. In this case Hoeffding's decomposition (1) reduces to $\mathbb{U} = L+Q$, where, by the assumption $\sigma_h^2 > 0$, we have $\mathbf{Var}L > 0$. Since the cubic part vanishes we remove the moment $\mathbf{E}g(X_1)g(X_2)g(X_3)\chi(X_1, X_2, X_3)$ from the expression for κ_4. In this way we obtain the two term Edgeworth expansion (2) for the distribution function $\mathbb{F}_U(x) = \mathbf{P}\{\mathbb{U} \le \sigma_{\mathbb{U}} x\}$ with $\sigma_{\mathbb{U}}^2 := \mathbf{Var}\mathbb{U}$.

We call h reducible if for some measurable functions $u, v : \mathcal{X} \to \mathcal{R}$ we have $h(x, y) = v(x)u(y) + v(y)u(x)$ for $P_X \times P_X$ almost sure $(x, y) \in \mathcal{X} \times \mathcal{X}$. A simple calculation shows that for a sequence of U-statistics (9) with a fixed non-reducible kernel condition (4) is satisfied, for some $\delta_* > 0$, uniformly in N. A straightforward consequence of Theorem 1 is the following corollary. Write $\tilde{\nu} = 600^{-1} \min\{\nu_2, r - 4, 1\}$.

Corollary 1 *Assume that* $\mathbf{E}h(X_1, X_2) = 0$ *and for some* $r > 4$

$$\mathbf{E}|h(X_1, X_2)|^r < \infty. \tag{10}$$

Assume that $\sigma_h^2 > 0$ *and the kernel* h *is non-reducible and that for some* $\delta > 0$

$$\sup\{|\mathbf{E}e^{it\sigma_h^{-1}h_1(X_1)}| : |t| \ge \beta_3^{-1}\} \le 1 - \delta. \tag{11}$$

Then there exist a constant $C_* > 0$ *such that*

$$\sup_{x \in \mathbb{R}} |\mathbb{F}_U(x) - G(x)| \le C_* N^{-1-\tilde{\nu}}.$$

For U-statistics with a fixed kernel the validity of the Edgeworth expansion (2) up to the order $o(N^{-1})$ was established by Callaert, Janssen and Veraverbeke [16] and Bickel, Götze and van Zwet [11]. In addition to the moment conditions (like (10)) and Cramér's condition (like (11)) Callaert, Janssen and Veraverbeke [16] imposed the following rather implicit condition. They assumed that for some $0 < c < 1$ and $0 < \alpha < 1/8$ the event

$$\Big|\mathbf{E}\big(\exp\{it\sigma_{\mathbb{U}}^{-1}\sum_{j=m+1}^{N} h(X_1, X_j)\}\,\big|X_{m+1}, \dots, X_N\big)\Big| \le c \tag{12}$$

has probability $1 - o(1/N \log N)$ uniformly for all $t \in [N^{3/4}/\log N,\ N \log N]$. Here $m \approx N^{\alpha}$, for a small positive α. Bickel, Götze and van Zwet [11] more explicitly required that the linear operator, $f(\cdot) \to \mathbf{E}\psi(X, \cdot) f(X)$ defined by ψ has a sufficiently large number of non-zero eigenvalues (the number depending on the existing moments, but always larger than 4). Correspondingly the eigenvalue condition is stronger than the non-reducibility condition of Corollary 1 since for a reducible kernel h the linear operator $f(\cdot) \to \mathbf{E}\psi(X, \cdot) f(X)$ has at most two eigenvalues. On the other hand, it is difficult to compare the structural non-reducibility condition with condition (12) whose technical nature is discussed in the outline of the proof at the beginning of Sect. 2.

The remaining parts of the paper (Sects. 2–5) contain the proof of Theorem 1. Auxiliary results are placed in the "Appendix".

2 Proof of Theorem 1

2.1 Proof highlights

After the seminal paper of Esseen [22] a standard proof of the validity of the normal approximation and its refinements proceeds in two steps. In the first step, with the aid of a smoothing inequality, the Kolmogorov distance between the distribution function and its approximation G is upper bounded by a (weighted) average difference of the respective Fourier transforms, see (25). In the second step one performs a careful analysis of the Fourier transforms: for frequencies $t = O(\sqrt{N})$ one shows the closeness between the respective Fourier transforms, while for the remaining range $\Omega(\sqrt{N}) \le |t| \le O(T)$ one establishes their exponential decay. The cut-of T is defined by the desired approximation accuracy level $O(T^{-1})$ (in our case $T = N^{1+\nu}$)). The approach, initially developed for sums of independent random variables [22, 33], was later applied to non-degenerate U-statistics [11, 16] and general asymptotically linear symmetric statistics [4, 37].

One particular problem related to the implementation of the proof strategy outlined above is about establishing exponential decay of the (absolute value of the) Fourier transform in the range of large frequencies. For a linear statistic this problem is elegantly resolved by introducing Cramér's condition. Indeed, in view of the multiplicity property of the Fourier transform, the Cramér condition implies the desired exponential decay. Consequently, Cramér's condition together with moment conditions ensure the validity of an Edgeworth expansion of an arbitrary order. But the multiplicity property can not be used any more (at least directly) when we turn to general symmetric statistics because various parts (linear, quadratic etc.) are mutually dependent. This fact leads to considerable difficulties in estimating the respective Fourier transforms in the range of large frequencies $t \gg N$ and requires new conditions to control the above mentioned dependencies. The present paper suggests a novel approach to estimation of the Fourier transform of a symmetric statistic for large frequencies.

As our general setup of symmetric statistics covers linear ones, we keep assuming the Cramér condition, but on the linear part of the statistic only, see (7). In view of **Example 1**, condition (7) is not enough. We introduce the additional structural condition (4), which together with (7) guarantees the desired $O(N^{-1-\nu})$ upper bound on the weighted average of the Fourier transform over the frequency range $N^{1-\nu} \leq |t| \leq N^{1+\nu}$, see (26) below. Condition (4) is optimal and natural in the sense that it matches the counterexample. It has first appeared in the unpublished manuscript [25] by Götze and van Zwet in the case of U statistics.

Let us compare (4) with alternative conditions introduced in earlier papers by Callaert, Janssen and Veraverbeke [16] and Bickel, Götze and van Zwet [11] in the case of U statistics of degree two. The conditional Cramér condition (12) of [16] forces the multiplicity property of the Fourier transform in a formal way thus circumventing the problem of establishing relation between the structure of the kernel (of U statistic) and the smoothness of the distribution. Therefore (4) and (12) are not comparable. This is not the case with the eigenvalue condition of [11], which is stronger than (4). In their proof Bickel, Götze and van Zwet [11] have used for the frequencies $t \in [N^{(r-1)/r}/\log N,\ N \log N]$ a symmetrization technique of [23] which essentially estimates the absolute value of the Fourier transform of U by that of a bilinear version of Q thus neglecting L and its smoothess properties implied by Cramér's condition (7). The approach of the present paper instead makes use of the smoothness of L and Q simultaneously.

The main contribution of this paper is in showing that condition (4) suggested by the counterexample (Example 1) is sufficient to prove the bound of Theorem 1. This condition is used in constructing estimates of weighted averages of the Fourier transform (26) that we briefly comment below. In fact, after initial "linearization" step we turn to slightly modified statistic $\tilde{\mathbb{T}}(X_1, \dots, X_N)$, where the nonlinear terms in $X_1, \dots, X_m$ are removed (see (19)), and then switch to $T' = \tilde{\mathbb{T}}(X_1, \dots, X_m, Y_{m+1}, \dots, Y_N)$, where $Y_{m+1}, \dots, Y_N$ are truncated versions of $X_{m+1}, \dots, X_N$, see (42). Let $\mathbf{E}_{\mathbb{Y}}$ denote the conditional expectation given $\mathbb{Y} = (Y_{m+1}, \dots, Y_N)$. The conditional

Fourier transform $\mathbf{E}_{\mathbb{Y}} \exp\{itT'\} = \mathbf{E}\big(\exp\{itT'\}\big|Y_{m+1}, \ldots, Y_N\big)$ contains a multiplicative component α_t^m, where

$$\alpha_t = \mathbf{E}_{\mathbb{Y}} \exp\left\{itN^{-1/2}g(X_1) + itN^{-3/2}\sum_{l=m+1}^{N}\psi(X_1, Y_l)\right\}. \tag{13}$$

For t satisfying $|\alpha_t|^2 \le 1 - m^{-1}\ln^2 N$ the bound $|\mathbf{E}_{\mathbb{Y}} \exp\{itT'\}| \le \exp\{-0.5\ln^2 N\}$ follows immediately. We then look carefully at the set of remaining t. We show that this set is a union of non-intersecting intervals (depending on $\mathbb{Y}$) each of size $O(\sqrt{N/m}\ln N)$. While estimating the weighted averages of the Fourier transform over these intervals we split the frequency domain $N^{1-\nu} \le |T| \le N^{1+\nu}$ into a deterministic sequence J_p, $p = 1, 2, \ldots$, of consecutive intervals of size $\Theta(N^{1-\nu})$ so that each 'singular' set $\{t \in J_p : |\alpha_t|^2 > 1 - m^{-1}\ln^2 N\}$ is either empty or an interval $[a_N, a_N + b_N^{-1}]$ of size $b_N^{-1} = O\big(\sqrt{N/m}\ln N\big)$, (see (51) and (56) based on Lemma 12). At the very last step, using Kleitman's concentration inequalities for sums of random variables with values in a function space, we upper bound the probability of the event that each particular singular set is non-empty, that is, the event that $\sup_{t\in J_p}|\alpha_t|^2 > 1 - m^{-1}\ln^2 N$ thus obtaining an extra factor $N^{-k\nu}$, $k \ge 5$ to arrive to the error bound $o(N^{-1})$.

More precisely, the non-zero projection to the g orthogonal part of $\sum_{l=m+1}^{N}\psi(\cdot, Y_l)$ which is non zero by condition (4) is used in the crucial Lemma 2. Via conditioning and randomization we represent it as a sum $S_\alpha := \sum_{j=1}^{n}\alpha_j f_j$ of independent $\alpha_j = 0, 1$ variables with vectors f_j with $||f_j|| > \epsilon$ and estimate the combinatorial probability for those $\alpha = (\alpha_1, \ldots, \alpha_n)$ that a value larger than $1 - m^{-1}ln^2 N$ of the conditional Fourier transform, say $\tilde{\phi}_t(\alpha)$, of $f + S_\alpha$ occurs at some 'singular' frequency $t \in J_p$. This is achieved by Kleitman's partition of the 2^n α's into at most $\binom{n}{n/2}$ disjoint sets, say C_d, $1 \le d \le \binom{n}{n/2}$, such that for different $\alpha, \alpha' \in C_d$, S_α and $S_{\alpha'}$ are separated by a distance of at least ϵ. This separation implies by Lemma 2 that the event that t is singular somewhere in the interval J_p can be witnessed by at most one $\alpha \in C_d$ for each C_d. Hence the singular event among the $\alpha's$ has combinatorial probability at most $\binom{n}{n/2}2^{-n} = O(n^{-1/2})$.

The crucial arguments in Lemma 2 rest upon the observation on harmonics (see (118)) that two singular values $\tilde{\phi}_t(\alpha), \tilde{\phi}_s(\alpha') \ge 1 - m^{-1}\ln^2 N$ imply a similar high value of $\mathbf{E}\exp\{i(t(f + S_\alpha) - s(f + S_{\alpha'}))\}$. If here t and s are close, say $|t - s| \le \delta_2$, such a high value is excluded by the separation of S_α and S_α which dominate $(t - s)f$ (see step 4.2.1 in Lemma 2), whereas for $\delta_2 < |t - s| < N^{\nu - 1/2}$, Cramér's condition for $(t - s)f$ applies which together with size bounds on tS_α and $sS_{\alpha'}$ again prevents a high value (see step 4.2.2 in Lemma 2).

Note that this method of width bounds and separation of singular sets of Fourier transforms has been successfully employed for optimal approximation results for U-statistics with non-Gaussian limits by Bentkus, Götze and Zaitsev, see [5] and [27] and is strongly related to results on the distribution of quadratic forms on lattices by Bentkus and Götze, see [6] and [24], the latter providing a solution of the Davenport-Lewis conjecture for positive definite forms.

Finally, we mention that in the case of U statistics of degree three ($\mathbb{T} = \mathbf{E}\mathbb{T} + L + Q + K$) the proof is outlined in the unpublished manuscript of Götze and van Zwet [25]. We extend these arguments to general symmetric statistics using stochastic expansions by means of Hoeffding's decomposition and bounds for various parts of the decomposition.

2.2 Outline of the proof

Firstly, using the linear structure induced by Hoeffding's decomposition we replace $\mathbb{T}/\sigma_{\mathbb{T}}$ by the statistic $\tilde{\mathbb{T}}$ which is conditionally linear given $X_{m+1}, \dots, X_N$. Secondly, invoking a smoothing inequality we pass from distribution functions to Fourier transforms. In the remaining steps we bound the difference $\delta(t) = \mathbf{E}e^{it\tilde{\mathbb{T}}} - \hat{G}(t)$, for $|t| \le N^{1+\nu}$. For "small frequencies" $|t| \le CN^{1/2}$, we expand the characteristic function $\mathbf{E}e^{it\tilde{\mathbb{T}}}$ in order to show that $\delta(t) = o(N^{-1})$. Here we combine various techniques developed in earlier papers [4, 11, 16]. For remaining range of frequencies, that is $CN^{1/2} \le |t| \le N^{1+\nu}$, we bound the summands $\mathbf{E}e^{it\tilde{\mathbb{T}}}$ and $\hat{G}(t)$ separately. The cases of "large frequencies" $N^{1-\nu} \le |t| \le N^{1+\nu}$ and "medium frequencies" $C\sqrt{N} \le |t| \le N^{1-\nu}$ are treated in a different manner. For medium frequencies the Cramér type condition (7) ensures an exponential decay of $|\mathbf{E}e^{it\tilde{\mathbb{T}}}|$. For large frequencies we combine conditions (7) and (4).

2.3 Hoeffding's decomposition

Before starting the proof we introduce some notation. By c_* we shall denote a positive constant which may depend only on $A_*, D_*, M_*, r, s, \nu_1, \nu_2, \delta$, but it does not depend on N. In different places the values of c_* may be different.

It is convenient to write the decomposition in the form

$$\mathbb{T} = \mathbf{E}\mathbb{T} + \sum_{1\le k\le N} \mathbb{U}_k, \qquad \mathbb{U}_k = \sum_{1\le i_1<\dots<i_k\le N} g_k(X_{i_1}, \dots, X_{i_k}), \tag{14}$$

where, for every k, the symmetric kernel g_k is centered, i.e., $\mathbf{E}g_k(X_1, \dots, X_k) = 0$, and satisfies, see, e.g., [4],

$$\mathbf{E}\big(g_k(X_1, \dots, X_k)\big|X_2, \dots, X_k\big) = 0 \qquad \text{almost surely.} \tag{15}$$

Here we write $g_1 := N^{-1/2}g$, $g_2 := N^{-3/2}\psi$ and $g_3 := N^{-5/2}\chi$. Furthermore, for an integer $k > 0$ we write $\Omega_k := \{1, \dots, k\}$. Given a subset $A = \{i_1, \dots, i_k\} \subset \Omega_N$ we write, for short, $T_A := g_k(X_{i_1}, \dots, X_{i_k})$. Put $T_\emptyset := \mathbf{E}\mathbb{T}$. Now the decomposition (14) can be written as follows

$$\mathbb{T} = \mathbf{E}\mathbb{T} + \sum_{1\le k\le N} \mathbb{U}_k = \sum_{A\subset\Omega_N} T_A, \qquad \mathbb{U}_k = \sum_{|A|=k,\, A\subset\Omega_N} T_A.$$

2.4 Proof of Theorem 1

Throughout the proof we assume without loss of generality that

$$4 < r \le 5, \qquad 2 < s \le 3 \qquad \text{and} \qquad \mathbf{E}\mathbb{T} = 0, \qquad \sigma_{\mathbb{T}}^2 = 1. \tag{16}$$

Denote, for $t > 0$,

$$\beta_t = \sigma^{-t}\mathbf{E}|g(X_1)|^t, \qquad \gamma_t = \mathbf{E}|\psi(X_1, X_2)|^t, \qquad \zeta_t = \mathbf{E}|\chi(X_1, X_2, X_3)|^t.$$

Linearization. Choose number $\nu > 0$ and integer m such that

$$\nu = 600^{-1}\min\{\nu_1,\ \nu_2,\ s-2,\ r-4\}, \qquad m \approx N^{100\nu}. \tag{17}$$

Split

$$\mathbb{T} = \mathbb{T}_{[m]} + \mathbb{W}, \qquad \mathbb{T}_{[m]} = \sum_{A:\, A\cap\Omega_m \neq \emptyset} T_A, \qquad \mathbb{W} = \sum_{A:\, A\cap\Omega_m = \emptyset} T_A. \tag{18}$$

Furthermore, write

$$\begin{aligned}
&\mathbb{T}_{[m]} = \mathbb{U}_1^* + \mathbb{U}_2^* + \Lambda, \qquad \Lambda = \Lambda_1 + \Lambda_2 + \Lambda_3 + \Lambda_4 + \Lambda_5,\\
&\mathbb{U}_1^* = \sum_{i=1}^m T_{\{i\}}, \qquad \mathbb{U}_2^* = \sum_{i=1}^m \sum_{j=m+1}^N T_{\{i,j\}},\\
&\Lambda_1 = \sum_{1\le i<j\le m} T_{\{i,j\}}, \qquad \Lambda_2 = \sum_{|A|\ge 3, |A\cap\Omega_m|=2} T_A,\\
&\Lambda_3 = \sum_{A:\, |A\cap\Omega_m|\ge 3} T_A, \qquad \Lambda_4 = \sum_{|A|=3,\ |A\cap\Omega_m|=1} T_A,\\
&\Lambda_5 = \sum_{i=1}^m \eta_i, \qquad \eta_i = \sum_{|A|\ge 4,\ A\cap\Omega_m=\{i\}} T_A.
\end{aligned}$$

Before applying a smoothing inequality we replace $\mathbb{F}(x)$ by

$$\tilde{\mathbb{F}}(x) := \mathbf{P}\{\tilde{\mathbb{T}} \le x\}, \qquad \text{where} \qquad \tilde{\mathbb{T}} = \mathbb{U}_1^* + \mathbb{U}_2^* + \mathbb{W} = \mathbb{T} - \Lambda. \tag{19}$$

In order to show that Λ can be neglected we apply a simple Slutzky type argument. Given $\varepsilon > 0$, we have

$$\Delta \le \sup_{x\in\mathbb{R}} |\tilde{\mathbb{F}}(x) - G(x)| + \varepsilon \sup_{x\in\mathbb{R}} |G'(x)| + \mathbf{P}\{|\Lambda| > \varepsilon\}. \tag{20}$$

From Lemma 5 we obtain via Chebyshev's inequality, for $\varepsilon = N^{-1-\nu}$,

$$\begin{aligned}\mathbf{P}\{|\Lambda| > \varepsilon\} &\le \sum_{i=1}^{5} \mathbf{P}\left\{|\Lambda_i| > \frac{\varepsilon}{5}\right\} \\ &\le \left(\frac{5}{\varepsilon}\right)^3 \mathbf{E}|\Lambda_1|^3 + \left(\frac{5}{\varepsilon}\right)^2 (\mathbf{E}\Lambda_2^2 + \mathbf{E}\Lambda_3^2 + \mathbf{E}\Lambda_5^2) + \left(\frac{5}{\varepsilon}\right)^s \mathbf{E}|\Lambda_4|^s \\ &\le c_* N^{-1-\nu}.\end{aligned}$$

In the last step we used conditions (5), (6) and the inequality (168). Furthermore, using (5) and (6) one can show that

$$\sup_{x\in\mathbb{R}} |G'(x)| \le c_*. \tag{21}$$

Therefore, (20) implies

$$\Delta \le \tilde{\Delta} + c_* N^{-1-\nu}, \qquad \text{where} \qquad \tilde{\Delta} := \sup_{x\in\mathbb{R}} |\tilde{\mathbb{F}}(x) - G(x)|.$$

It remains to show that $\tilde{\Delta} \le c_* N^{-1-\nu}$.

A smoothing inequality. Given $a > 0$ and even integer $k \ge 2$ consider the probability density function, see (10.7) in Bhattacharya and Rao [7],

$$x \to g_{a,k}(x) = a\, c(k)(ax)^{-k} \sin^k(ax), \tag{22}$$

where $c(k)$ is the normalizing constant. Its characteristic function

$$\hat{g}_{a,k}(t) = \int_{-\infty}^{+\infty} e^{itx} g_{a,k}(x)dx = 2\pi\, a\, c(k) u^{*k}_{[-a,a]}(t)$$

vanishes outside the interval $|t| \le ka$. Here $u^{*k}_{[-a,a]}(t)$ denotes the probability density function of the sum of k independent random variables each uniformly distributed in $[-a, a]$. It is easy to show that the function $t \to \hat{g}_{a,k}(t)$ is unimodal and symmetric around $t = 0$.

Let μ be the probability distribution with the density $g_{a,2}$, where a is chosen to satisfy $\mu([-1, 1]) = 3/4$. Given $T > 1$ define $\mu_T(\mathcal{A}) = \mu(T\mathcal{A})$, for a Borel set $\mathcal{A} \subset \mathbb{R}$. Let $\hat{\mu}_T$ denote the characteristic function corresponding to μ_T.

We apply Lemma 12.1 of [7]. It follows from (21) and the identity $\mu_T([-T^{-1}, T^{-1}]) = 3/4$ that

$$\tilde{\Delta} \le 2 \sup_{x\in\mathbb{R}} \left|(\tilde{\mathcal{F}} - \mathcal{G}) * \mu_T(-\infty, x]\right| + c_* T^{-1}. \tag{23}$$

Here $\tilde{\mathcal{F}}$ and $\mathcal{G}$ denote the probability distribution of $\tilde{\mathbb{T}}$ and the signed measure with density function $G'(x)$ respectively. Furthermore, $*$ denotes the convolution operation. Proceeding as in the proof of Lemma 12.2 ibidem we obtain

$$(\tilde{\mathcal{F}} - \mathcal{G}) * \mu_T(-\infty, x] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \left(\mathbf{E}e^{it\tilde{\mathbb{T}}} - \hat{G}(t)\right) \frac{\hat{\mu}_T(t)}{-it} dt, \tag{24}$$

where $\hat{G}$ denotes the Fourier transform of $G(x)$. Note that $\hat{\mu}_T(t)$ vanishes outside the interval $|t| \le 2aT$. Finally, we obtain from (23) and (24) that

$$\tilde{\Delta} \le \frac{1}{\pi} \sup_{x \in \mathbb{R}} |I(x)| + c_* \frac{2a}{T}, \qquad I(x) := \int_{-T}^{T} e^{-itx} \big(\mathbf{E} e^{it\tilde{\mathbb{T}}} - \hat{G}(t) \big) \frac{\hat{\mu}_{T'}(t)}{-it} dt, \quad (25)$$

where $T' = T/2a$. Here we use the fact that $\hat{\mu}_{T'}(t) = 0$ for $|t| > T$. Denote $K_N(t) = \hat{\mu}_{T'}(t)$ and observe that $|K_N(t)| \le 1$ (since $\mu_{T'}$ is a probability measure). Let

$$T = N^{1+\nu}.$$

We have

$$\begin{aligned}
&|I(x)| \le I_1 + I_2 + |I_3| + |I_4|, \\
&I_1 = \int_{|t| \le t_1} \big| \mathbf{E} e^{it\tilde{\mathbb{T}}} - \hat{G}(t) \big| \frac{dt}{|t|}, \quad I_2 = \int_{t_1 < |t| < T} |\hat{G}(t)| \frac{dt}{|t|}, \\
&I_3 = \int_{t_1 < |t| < t_2} e^{-itx} \mathbf{E} e^{it\tilde{\mathbb{T}}} \frac{K_N(t)}{-it} dt, \quad I_4 = \int_{t_2 < |t| < T} e^{-itx} \mathbf{E} e^{it\tilde{\mathbb{T}}} \frac{K_N(t)}{-it} dt.
\end{aligned}$$

Here we denote $t_1 = N^{1/2} 10^{-3} / \beta_3$ and $t_2 = N^{1-\nu}$. In view of (25) the bound $\tilde{\Delta} \le c_* N^{-1-\nu}$ follows from the bounds

$$|I_k| \le c_* N^{-1-\nu}, \quad k = 1, 2, 3, \quad \text{and} \quad |I_4| \le c_* N^{-1-\nu} (1 + \delta_*^{-1} N^{-\nu}). \quad (26)$$

The bound $I_2 \le c_* N^{-1-\nu}$ is a consequence of the exponential decay of $|\hat{G}(t)|$ as $|t| \to \infty$. In Sect. 3 we show (26) for $k = 3, 4$. The proof of (26), for $k = 1$, is based on careful expansions and is given Sect. 5.

3 Large frequencies

Here we prove inequalities (26) for I_3 and I_4. The proof of $|I_3| \le c_* N^{-1-\nu}$ is relatively simple and is deferred to the end of the section.

Let us upper bound $|I_4|$. We will show that

$$\Big| \int_{N^{1-\nu} < |t| < N^{1+\nu}} e^{-itx} \mathbf{E} e^{it\tilde{\mathbb{T}}} \frac{K_N(t)}{t} dt \Big| \le c_* \frac{1 + \delta_*^{-1}}{N^{1+2\nu}}. \quad (27)$$

In what follows we assume that N is sufficiently large, say $N > C_*$, where C_* depends only on $A_*, D_*, M_*, r, s, \nu_1, \nu_2, \delta$. We use this inequality in several places below, where the constant C_* can be easily specified. Note that for small N such that $N \le C_*$ the inequality (27) becomes trivial.

3.1 Notation

Let us first introduce some notation. Introduce the number

$$\alpha = 3/(r+2) \tag{28}$$

and note that for $r \in (4, 5]$ and ν defined by (17) we have

$$2/r < \alpha < 1/2 \qquad \text{and} \qquad 80\nu < \min\{r\alpha - 2,\ 1 - 2\alpha\}.$$

Given N introduce the integers

$$n \approx N^{50\nu}, \qquad M = \lfloor (N-m)/n \rfloor. \tag{29}$$

We have $N - m = M\,n + s$, where the integer $0 \le s < n$. Observe, that the inequalities $\nu < 600^{-1}$ and $m < N^{1/2}$, see (17), imply $M > n$. Therefore $s < M$. Split the index set

$$\begin{aligned} &\{m+1, \ldots, N\} = O_1 \cup O_2 \cup \cdots \cup O_n, \\ &O_i = \{j : m + (i-1)M < j \le m + iM\}, \quad 1 \le i \le n-1, \\ &O_n = \{j : m + (n-1)M < j \le N\}. \end{aligned} \tag{30}$$

Clearly, $O_1, \ldots, O_{n-1}$ are of equal size (=M) and $|O_n| = M + s < 2M$.

We shall assume that the random variable $X : \Omega \to \mathcal{X}$ is defined on the probability space (Ω, P) and P_X is the probability distribution on $\mathcal{X}$ induced by X. Given $p \ge 1$ let $L^p = L^p(\mathcal{X}, \mathcal{P}_{\mathcal{X}})$ denote the space of real functions $f : \mathcal{X} \to \mathbb{R}$ with $\mathbf{E}|f(X)|^p < \infty$. Denote $\|f\|_p = (\mathbf{E}|f(X)|^p)^{1/p}$. With a random variable $f(X)$ we associate an element (vector) $f = f(\cdot)$ of L^p, $p \le r$. Let $p_g : L^2 \to L^2$ denote the projection onto the subspace orthogonal to the vector $g(\cdot)$ in L^2. Given $h \in L^2$, decompose

$$h = a_h g + h^*, \qquad \text{where} \qquad a_h = \langle h, g\rangle \, \|g\|_2^{-2} \qquad \text{and} \qquad h^* = p_g(h). \tag{31}$$

Here $\langle h, g\rangle = \int h(x) g(x) P_X(dx)$. For $h \in L^r$ we have

$$\|h\|_r \ge \|h\|_2 \ge \|h^*\|_2. \tag{32}$$

Furthermore, for $r^{-1} + v^{-1} = 1$ (here $r \ge 2 \ge v > 1$) we have, by Hölder's inequality,

$$|\,\langle h, g\rangle\,| \le \|h\|_r \|g\|_v \le \|h\|_r \|g\|_2.$$

In particular,

$$|a_h| \le \|h\|_r / \|g\|_2. \tag{33}$$

Denote

$$c_g := 1 + \|g\|_r/\|g\|_2, \qquad c_g^* := 1 + M_*^{1/r} A_*^{-1/2}$$

and observe that $c_g \le c_g^*$. It follows from the decomposition (31) and (33) that

$$\|h^*\|_r \le \|h\|_r + |a_h|\,\|g\|_r \le \|h\|_r(1 + \|g\|_r/\|g\|_2) = c_g\|h\|_r. \tag{34}$$

Introduce the numbers

$$a_1 = \frac{1}{4}\min\left\{\frac{1}{12c_g^*}, \frac{(c_r A_*/2^r M_*)^{1/(r-2)}}{1+4A_*^{-1/2}}\right\}, \qquad c_r = \frac{7}{24}\frac{1}{2^{r-1}}. \tag{35}$$

It follows from (7) that there exist $\delta', \delta'' > 0$ depending on A_*, M_*, δ such that (uniformly in N) Cramér's characteristic ρ satisfies the inequalities

$$\rho(a_1, 2N^{-\nu+1/2}) \ge \delta', \qquad \rho((2\beta_3)^{-1}, N^{\nu_2+1/2}) \ge \delta''. \tag{36}$$

We shall prove the first inequality only. In view of (7) it suffices to prove that $\rho(a_1, \beta_3^{-1}) \ge \delta'$. Expanding the exponent in powers of $itg(X_1)/\sigma$ we show the inequality

$$|\mathbf{E}e^{it\sigma^{-1}g(X_1)}| \le 1 - 2^{-1}t^2(1 - 3^{-1}|t|\beta_3).$$

For $|t| \le \beta_3^{-1}$ this inequality implies

$$|\mathbf{E}e^{it\sigma^{-1}g(X_1)}| \le 1 - t^2/3.$$

Therefore, $\rho(a_1, \beta_3^{-1}) \ge a_1^2/3$ and we can choose $\delta' = \min\{\delta, a_1^2/3\}$ in (36).

Introduce the constant (depending only on A_*, M_*, δ)

$$\delta_1 = \delta'/(10c_g^*). \tag{37}$$

Note that $0 < \delta_1 < 1/10$. Given $f \in L^r$ and $T_0 \in \mathbb{R}$ such that

$$N^{-\nu+1/2} \le |T_0| \le N^{\nu+1/2}, \tag{38}$$

denote

$$\begin{aligned} I(T_0) &= [T_0,\ T_0 + \delta_1 N^{-\nu+1/2}],\\ u_t(f) &= \int \exp\bigl\{it\bigl(g(x) + N^{-1/2}f(x)\bigr)\bigr\}P_X(dx),\\ v(f) &= \sup_{t\in I(T_0)} |u_t(f)|, \qquad \tau(f) = 1 - v^2(f). \end{aligned} \tag{39}$$

Given a random variable η with values in L^r and number $0 < s < 1$, define

$$d_s(\eta, I(T_0)) = \mathbb{I}_{\{v^2(\eta) > 1 - s^2\}} \mathbb{I}_{\{\|\eta\|_r \le N^\nu\}}, \qquad \delta_s(\eta, I(T_0)) = \mathbf{E} d_s(\eta, I(T_0)). \quad (40)$$

Introduce the function

$$\psi^{**}(x, y) = \psi(x, y) - b(x)g(y) - b(y)g(x) \quad (41)$$

and the number

$$\delta_3^2 = \mathbf{E}|\psi^{**}(X_1, X_2)|^2.$$

It follows from (4) and our assumption $\sigma_{\mathbb{T}}^2 = 1$, see (16), that $\delta_3^2 \ge \delta_*^2$.

3.2 Proof of (27)

We write $\mathbf{E}_{\mathbb{Y}} \exp\{itT'\}$ in the form $\mathbf{E}_{\mathbb{Y}} \exp\{itT'\} = \alpha_t^m \exp\{itW'\}$, where α_t is defined in (13) and where the random variable W' is defined in the same way as $\mathbb{W}$ in (18), but with $T_A = g_k(X_{i_1}, \dots, X_{i_k})$ replaced by $g_k(Y_{i_1}, \dots, Y_{i_k})$ for each $A = \{i_1, \dots, i_k\}$. A standard way to upper bound a quantity like $|\mathbf{E}_{\mathbb{Y}} e^{itT'}|$ is to show an exponential decay (in m) of the product $|\alpha_t^m|$ using a Cramér type condition. This task can be accomplished for medium frequencies. Indeed, for $|t| = o(N)$ the quadratic part $itN^{-3/2} \sum_{j=m+1}^N \psi(X_1, Y_j)$ can be neglected and Cramér's condition implies $|\alpha_t| \le 1 - v'$ for some $v' > 0$. This leads to an exponential bound $|\alpha_t^m| \le e^{-mv'}$. For large frequencies $|t| \approx N$, the contribution of the quadratic part becomes significant. To upper bound $|\alpha_t^m|$ we use condition (4). We show that, for a large set of values $t \in J_p$, see (51), Cramér's condition (7) yields the desired decay of $|\alpha_t^m|$, while the measure of the set of remaining t is small with high probability.

Step 1. Truncation. Recall that the random variable $X : \Omega \to \mathcal{X}$ is defined on the probability space (Ω, P). Let X' be an independent copy so that (X, X') is defined on $(\Omega \times \Omega', P \times P)$, where $\Omega' = \Omega$. It follows from $\mathbf{E}|\psi(X, X')|^r < \infty$, by Fubini, that for P almost all $\omega' \in \Omega'$ the function $\psi(\cdot, X'(\omega')) = \{x \to \psi(x, X'(\omega')),\ x \in \mathcal{X}\}$ is an element of L^r. Furthermore, one can define an L^r-valued random variable $Z' : \Omega' \to L^r$ such that $Z'(\omega') = \psi(\cdot, X'(\omega'))$, for P almost all ω'. Consider the event $\tilde{\Omega} = \{\|Z'\|_r \le N^\alpha\} \subset \Omega'$ and denote $q_N = P(\tilde{\Omega})$. Here $\|Z'\|_r = (\int |\psi(x, X'(w'))|^r P_X(dx))^{1/r}$ denotes the L^r norm of the random vector Z' and α is defined in (28). Let $Y : \tilde{\Omega} \to \mathcal{X}$ denote the random variable X' conditioned on the event $\tilde{\Omega}$. Therefore Y is defined on the probability space $(\tilde{\Omega}, \tilde{P})$, where $\tilde{P}$ denotes the restriction of $q_N^{-1} P$ to the set $\tilde{\Omega}$ and, for every $\omega' \in \tilde{\Omega}$, we have $Y(\omega') = X'(\omega')$. Let Z denote the L^r- valued random element $\{x \to \psi(x, Y(\omega'))\}$ defined on the probability space $(\tilde{\Omega}, \tilde{P})$.

We can assume that $\mathbb{X} := (X_1, \dots, X_N)$ is a sequence of independent copies of X defined on the probability space (Ω^N, P^N). Let $\overline{\omega} = (\omega_1, \dots, \omega_N)$ denote an element of Ω^N. Every X_j defines random vector $Z'_j = \psi(\cdot, X_j)$ taking values in L^r. Introduce events $A_j := \{\|Z'_j\|_r \le N^\alpha\} \subset \Omega^N$ and let $\mathbb{X}' = (X_1, \dots, X_m, Y_{m+1}, \dots, Y_N)$ denote the sequence $\mathbb{X}$ conditioned on the event $\Omega^* = \cap_{j=m+1}^N A_j = \Omega^m \times \tilde{\Omega}^{N-m}$. Clearly, $\mathbb{X}'(\overline{\omega}) = \mathbb{X}(\overline{\omega})$ for every $\overline{\omega} \in \Omega^*$ and $\mathbb{X}'$ is defined on the space $\Omega^m \times \tilde{\Omega}^{N-m}$ equipped with the probability measure $P^m \times \tilde{P}^{N-m}$. In particular, the random variables $X_1, \dots, X_m, Y_{m+1}, \dots, Y_N$ are independent and Y_j, for $m+1 \le j \le N$, has the same distribution as Y. Let Z_j denote the L^r- valued random element $\{x \to \psi(x, Y_j),\ x \in \mathcal{X}\}$, for $m+1 \le j \le N$. Let

$$T' := \tilde{\mathbb{T}}(X_1, \dots, X_m, Y_{m+1}, \dots, Y_N). \tag{42}$$

We are going to replace $\mathbf{E}e^{it\tilde{\mathbb{T}}}$ by $\mathbf{E}e^{itT'}$. For $s > 0$ we have almost surely

$$1 - \mathbb{I}_{A_j} \le N^{-\alpha s}\|Z'_j\|_r^s, \qquad \|Z'_j\|_r^r = \mathbf{E}\big(|\psi(X, X_j)|^r \,\big|\, X_j\big). \tag{43}$$

From (43) with $s = r$ we obtain, by Chebyshev's inequality, that

$$0 \le 1 - q_N \le N^{-r\alpha}\mathbf{E}|\psi(X, X_j)|^r \le N^{-r\alpha}M_* \le c_* N^{-2-3\nu}. \tag{44}$$

Consequently, for $k \le N$ we have

$$\begin{aligned} q_N^{-k} &\le (1 - N^{-r\alpha}M_*)^{-k} \le (1 - N^{-2}M_*)^{-N} \le c_*, \\ q_N^{-k} - 1 &\le k q_N^{-k}(1 - q_N) \le c_* k N^{-2-3\nu} \le c_* N^{-1-3\nu}. \end{aligned} \tag{45}$$

Using the identity, which holds for a measurable function $f : \mathcal{X}^{\mathcal{N}} \to \mathbb{R}$,

$$\mathbf{E}f(X_1, \dots, X_m, Y_{m+1}, \dots, Y_N) = \mathbf{E}f(X_1, \dots, X_N)\frac{\mathbb{I}_{A_{m+1}} \cdots \mathbb{I}_{A_N}}{q_N^{(N-m)}} \tag{46}$$

we obtain from (45) and (46) for $f \ge 0$ that

$$\mathbf{E}f(X_1, \dots, X_m, Y_{m+1}, \dots, Y_N) \le c_*\mathbf{E}f(X_1, \dots, X_N). \tag{47}$$

Furthermore, (45) and (46) imply

$$\begin{aligned} |\mathbf{E}e^{it(T'-x)} - \mathbf{E}e^{it(\tilde{\mathbb{T}}-x)}| &\le \big(q_N^{-(N-m)} - 1\big) + \big(1 - \mathbf{P}\{A_{m+1} \cap \cdots \cap A_N\}\big) \\ &= (q_N^{-(N-m)} - 1) + (1 - q_N^{N-m}) \le c_* N^{-1-3\nu}. \end{aligned} \tag{48}$$

Now we can replace the integral in (27) by the integral

$$I := \int_{N^{1-\nu} \le |t| \le N^{1+\nu}} \mathbf{E}e^{it\hat{T}} v_N(t)dt, \qquad \text{where} \qquad v_N(t) = t^{-1}K_N(t), \qquad \hat{T} = T' - x. \tag{49}$$

In view of (48) and the simple inequality $|K_N(t)| \le 1$ the error of this replacement is $c_* N^{-1-2\nu}$. Hence in order to prove (27) it remains to show that

$$|I| \le c_* \frac{1+\delta_3^{-1}}{N^{1+2\nu}}. \tag{50}$$

Step 2. Here we prove (50). We split the integral

$$I = \sum_p I_p, \qquad I_p = \mathbf{E} \int_{t \in J_p} e^{it\hat{T}} v_N(t) dt, \tag{51}$$

where $\{J_p,\ p = 1, 2, \dots\}$ is a sequence of consecutive intervals of length $\approx \delta_1 N^{1-\nu}$ each and $\cup_p J_p = [N^{1-\nu}, N^{1+\nu}]$. Recall that δ_1 is defined in (37). To prove (50) we show that for every p

$$|I_p| \le c_* N^{-2} + c_* N^{-1-4\nu}\big(1 + \delta_3^{-1}\big). \tag{52}$$

We fix p and prove (52). Firstly, we replace I_p by $\mathbf{E} J_*$, where

$$J_* = \int \mathbb{I}_{\{t \in I_*\}} v_N(t) \mathbf{E}_{\mathbb{Y}} e^{it\hat{T}} dt$$

and where $I_* = I_*(Y_{m+1}, \dots, Y_N) \subset J_p$ is a random subset:

$$I_* = \{t \in J_p : |\alpha_t|^2 > 1 - \varepsilon_m^2\}, \qquad \varepsilon_m^2 = m^{-1} \ln^2 N. \tag{53}$$

Note that for $t \in J_p \setminus I_*$, we have

$$|\mathbf{E}_{\mathbb{Y}} e^{itT'}| \le |\alpha_t|^m \le (1 - \varepsilon_m^2)^{m/2} \le c_* N^{-3}.$$

These inequalities imply the bound

$$|I_p - \mathbf{E} J_*| \le c_* N^{-2}. \tag{54}$$

Secondly, we show that with a high probability the set $I_* \subset J_p$ is an interval. This fact and the fact that $v_N(t)$ is monotone will be used latter to bound the integral J_*. Introduce the L^r- valued random element

$$S = N^{-1/2}(Z_{m+1} + \dots + Z_N) = N^{-1/2} \sum_{j=m+1}^{N} \psi(\cdot, Y_j). \tag{55}$$

We apply Lemma 12 (see below) to the set $N^{-1/2} I_*$ conditionally given the event $\mathbb{S} = \{\|S\|_r < N^{\nu/10}\}$. This lemma shows that $N^{-1/2} I_*$ is an interval of size at most

$c_*\varepsilon_m$. Hence we can write $I_* = (a_N, a_N + b_N^{-1})$ and

$$\mathbb{I}_{\mathbb{S}} J_* = \mathbb{I}_{\mathbb{S}} \mathbf{E}_{\mathbb{Y}} \tilde{J}_*, \qquad \tilde{J}_* = \int_{a_N}^{a_N + b_N^{-1}} v_N(t) e^{it\hat{T}} dt, \tag{56}$$

where the random variables a_N, b_N (functions of $Y_{m+1}, \dots, Y_N$) satisfy

$$a_N \in J_p \qquad \text{and} \qquad b_N^{-1} \le c_* \varepsilon_m \sqrt{N} = c_* \sqrt{N} m^{-1/2} \ln N.$$

Furthermore, by Lemma 13 below we have $\mathbf{P}\{\mathbb{S}\} \ge 1 - c_* N^{-3}$. Therefore,

$$|\mathbf{E} J_* - \mathbf{E} \mathbb{I}_{\mathbb{S}} J_*| \le c_* N^{-2}. \tag{57}$$

Next, we observe that $I_* \neq \emptyset$ if and only if $\tilde{\alpha}^2 > 1 - \varepsilon_m^2$, where

$$\tilde{\alpha} = \sup\{|\alpha_t| : t \in J_p\}.$$

Therefore we can write (56) in the form

$$\mathbb{I}_{\mathbb{S}} J_* = \mathbb{I}_{\mathbb{B}} J_* = \mathbb{I}_{\mathbb{B}} \mathbf{E}_{\mathbb{Y}} \tilde{J}_*, \qquad \text{where} \qquad \mathbb{B} = \{\tilde{\alpha}^2 > 1 - \varepsilon_m^2\} \cap \mathbb{S}.$$

This identity together with (54) and (57) imply

$$|I_p| \le |\mathbf{E} \mathbb{I}_{\mathbb{B}} \mathbf{E}_{\mathbb{Y}} \tilde{J}_*| + c_* N^{-2}. \tag{58}$$

Using the integration by parts formula we shall show below that

$$|\mathbf{E} \mathbb{I}_{\mathbb{B}} \mathbf{E}_{\mathbb{Y}} \tilde{J}_*| \le \frac{c}{N^{1-\nu}} \Big(\mathbf{P}\{\mathbb{B}\} + \int_{b_N}^{1} \frac{\mathbf{P}\{\mathbb{B}_\varepsilon\}}{\varepsilon^2} d\varepsilon \Big), \qquad \text{where} \qquad \mathbb{B}_\varepsilon := \mathbb{B} \cap \{|\hat{T}| \le \varepsilon\}. \tag{59}$$

Moreover, we shall show that

$$\int_{b_N}^{1} \frac{\mathbf{P}\{\mathbb{B}_\varepsilon\}}{\varepsilon^2} d\varepsilon \le c_* \frac{1 + \delta_3^{-1}}{N^{5\nu}} \qquad \text{and} \qquad \mathbf{P}\{\mathbb{B}\} \le c_* \frac{1 + \delta_3^{-1}}{N^{5\nu}}. \tag{60}$$

The latter inequalities in combination with (58) and (59) yield (52). We prove (60) in Sect. 3.3.

Let us prove (59). Firstly, we show that

$$|\tilde{J}_*| \le c(|\hat{T}| + b_N)^{-1} a_N^{-1}. \tag{61}$$

From the integration by parts formula we obtain the identity

$$i\hat{T} \tilde{J}_* = v_N(t) e^{it\hat{T}} \Big|_{a_N}^{a_N + b_N^{-1}} - \int_{a_N}^{a_N + b_N^{-1}} v_N'(t) e^{it\hat{T}} dt =: a' - a''. \tag{62}$$

By our choice of the smoothing kernel the function $v_N(t)$ is monotone on J_p. Therefore

$$|a''| \le \int_{a_N}^{a_N+b_N^{-1}} \left|v_N'(t)\right| dt = \left|\int_{a_N}^{a_N+b_N^{-1}} v_N'(t)dt\right| = \left|v_N(a_N) - v_N(a_N + b_N^{-1})\right|.$$

Invoking the simple inequality $|a'| \le |v_N(a_N)| + |v_N(a_N + b_N^{-1})|$ and using $|v_N(t)| \le |t|^{-1}$ we obtain from (62) that

$$|\hat{T}\tilde{J}_*| \le c\left(a_N^{-1} + (a_N + b_N^{-1})^{-1}\right) \le c\, a_N^{-1}.$$

For $|\hat{T}| > b_N$, this inequality implies (61). For $|\hat{T}| \le b_N$ the inequality (61) follows from the inequalities

$$|\tilde{J}_*| \le \int_{a_N}^{a_N+b_N^{-1}} |v_N(t)|dt \le \int_{a_N}^{a_N+b_N^{-1}} \frac{c}{|t|}dt \le c\, a_N^{-1} b_N^{-1}.$$

The proof of (61) is complete. Now from (61) and the inequality $a_N \ge N^{1-\nu}$ we obtain that

$$|\tilde{J}_*| \le c(|\hat{T}| + b_N)^{-1} N^{-1+\nu}.$$

Finally, using the inequality (which holds for arbitrary real number v)

$$\frac{1}{|v| + b_N} \le 2 + 2\int_{b_N}^{1} \frac{d\varepsilon}{\varepsilon^2} \mathbb{I}_{\{|v| \le \varepsilon\}}$$

we show that

$$|\tilde{J}_*| \le \frac{c_*}{N^{1-\nu}} \left(1 + \int_{b_N}^{1} \frac{d\varepsilon}{\varepsilon^2} \mathbb{I}_{\{|\hat{T}| \le \varepsilon\}}\right).$$

The latter inequality implies (59).

3.3 Proof of (60)

The first and second inequality of (60) are proved in steps A and B.
Step A. Proof of the first inequality of (60). Recall $\mathbb{W}$ from (18). We split

$$\mathbb{W} = W_1 + W_2 + W_3, \qquad W_1 = \frac{1}{N^{1/2}} \sum_{j=m+1}^{N} g(X_j),$$

$$W_2 = \frac{1}{N^{3/2}} \sum_{m<i<j\le N} \psi(X_i, X_j), \qquad W_3 = \sum_{|A|\ge 3: A\cap\Omega_m = \emptyset} T_A.$$

Define W_1', W_2', W_3' as W_1, W_2, W_3 above, but with X_j replaced by Y_j, for $m+1 \leq j \leq N$. We have $W' = W_1' + W_2' + W_3'$. Now we write $\hat{T}$ (see (49)) in the form $\hat{T} = L + \Delta + W_3'$, where

$$L = \frac{1}{\sqrt{N}} \sum_{j=1}^{m} g(X_j) + \frac{1}{\sqrt{N}} \sum_{j=m+1}^{N} g(Y_j) - x,$$
$$\Delta = \frac{1}{N^{3/2}} \sum_{j=1}^{m} \sum_{l=m+1}^{N} \psi(X_j, Y_l) + \frac{1}{N^{3/2}} \sum_{m+1 \leq j < l \leq N} \psi(Y_j, Y_l). \quad (63)$$

The inequalities $|\hat{T}| \leq \varepsilon$ and $|L| \geq 2\varepsilon$ imply $|\Delta + W_3'| > \varepsilon$. Therefore,

$$\mathbf{P}\{\mathbb{B}_\varepsilon\} \leq \mathbf{P}\{\mathbb{B} \cap \{|L| \leq 2\varepsilon\}\} + \mathbf{P}\{|\hat{T}| \leq \varepsilon,\ |\Delta + W_3'| \geq \varepsilon\} =: I_1(\varepsilon) + I_2(\varepsilon).$$

To prove the first inequality of (60) we show that

$$\int_{b_N}^{1} \frac{d\varepsilon}{\varepsilon^2} I_1(\varepsilon) \leq c_* N^{-5\nu}(1 + \delta_3^{-1}), \qquad \int_{b_N}^{1} \frac{d\varepsilon}{\varepsilon^2} I_2(\varepsilon) \leq c_* N^{-5\nu}. \quad (64)$$

Step A.1. Proof of the second inequality of (64). We have

$$I_2(\varepsilon) \leq \mathbf{P}\{|W_3'| \geq \varepsilon/2\} + I_3(\varepsilon), \quad \text{where} \quad I_3(\varepsilon) := \mathbf{P}\{|L + \Delta| < 3\varepsilon/2, |\Delta| > \varepsilon/2\}. \quad (65)$$

It follows from (47), by Chebyshev's inequality, that $\mathbf{P}\{|W_3'| > \varepsilon/2\} \leq c_* \varepsilon^{-2} \mathbf{E} W_3^2$. Furthermore, invoking the inequalities, see (167), (168) below,

$$\mathbf{E} W_3^2 = \sum_{|A| \geq 3: A \cap \Omega_m = \emptyset} \mathbf{E} T_A^2 \leq \sum_{|A| \geq 3} \mathbf{E} T_A^2 \leq N^{-2} \Delta_3^2 \leq c_* N^{-2}$$

we obtain from (65) that $I_2(\varepsilon) \leq I_3(\varepsilon) + c_* \varepsilon^{-2} N^{-2}$. Since

$$\int_{b_N}^{1} \frac{d\varepsilon}{\varepsilon^2} \Big(\frac{1}{\varepsilon^2 N^2} \Big) \leq c_* b_N^{-3} N^{-2} \leq c_* N^{-5\nu},$$

it suffices to show inequality (64) for $I_3(\varepsilon)$ (instead of $I_2(\varepsilon)$). Recall the notation $\Lambda_1 = N^{-3/2} \sum_{1 \leq i < j \leq m} \psi(X_i, X_j)$ and put $U = \Lambda_1 + \Delta$. We have

$$I_3(\varepsilon) \leq \mathbf{P}\{|\Lambda_1| > \varepsilon/4\} + I_4(\varepsilon), \quad \text{where} \quad I_4(\varepsilon) := \mathbf{P}\{|L + U| < 2\varepsilon,\ |U| > \varepsilon/4\}.$$

Invoking the inequality, which follows by Chebyshev's inequality,

$$\mathbf{P}\{|\Lambda_1| > \varepsilon/4\} \leq 16 \varepsilon^{-2} \mathbf{E} \Lambda_1^2 \leq c_* \varepsilon^{-2} m^2 N^{-3}$$

we upper bound the integral

$$\int_{b_N}^{1} \frac{d\varepsilon}{\varepsilon^2} \mathbf{P}\{|\Lambda_1| > \varepsilon/4\} \leq \int_{b_N}^{1} \frac{d\varepsilon}{\varepsilon^2} \Big(\frac{m^2}{\varepsilon^2 N^3}\Big) \leq c_* b_N^{-3} m^2 N^{-3} \leq c_* N^{-5\nu}.$$

Hence, it remains to show the second inequality of (64) for $I_4(\varepsilon)$.

Let $I_4'(\varepsilon)$ be the same probability as $I_4(\varepsilon)$ but with X_i replaced by Y_i, for $1 \leq i \leq m$. That is,

$$I_4'(\varepsilon) = \mathbf{P}\{|L' + U'| < 2\varepsilon,\ |U'| > \varepsilon/4\},$$
$$L' = \frac{1}{N^{1/2}} \sum_{1 \leq i \leq N} g(Y_i) - x, \qquad U' = \frac{1}{N^{3/2}} \sum_{1 \leq i < j \leq N} \psi(Y_i, Y_j).$$

By the same reasoning as in (48) we obtain that $|I_4(\varepsilon) - I_4'(\varepsilon)| \leq c_* N^{-1-3\nu}$. Now, in view of the bound

$$\int_{b_N}^{1} \frac{d\varepsilon}{\varepsilon^2} N^{-1-3\nu} \leq c_* b_N^{-1} N^{-1-3\nu} \leq c_* N^{-5\nu}$$

we conclude that it suffices to show (the second) inequality (64) for $I_4'(\varepsilon)$.

Let us show the second inequality of (64) for $I_4'(\varepsilon)$. We split the sample

$$\mathbb{Y} := \{Y_1, \dots, Y_N\} = \mathbb{Y}_1 \cup \mathbb{Y}_2 \cup \mathbb{Y}_3,$$

into three groups of nearly equal size. Next, we split $U' = \sum_{i \leq j} U_{ij}'$ so that the sum U_{ij}' depends on the observations from the groups $\mathbb{Y}_i$ and $\mathbb{Y}_j$ only. We have

$$I_4'(\varepsilon) \leq \sum_{i \leq j} \mathbf{P}\{|L' + U'| \leq 2\varepsilon,\ |U_{ij}'| \geq \varepsilon/24\}. \tag{66}$$

Now we show that the second inequality of (64) holds for every summand in the right of (66). Let $\tilde{U}$ denote a summand U_{ij}', say, not depending on $\mathbb{Y}_3$. Let

$$\tilde{I}(\varepsilon) := \mathbf{P}\{|L' + U'| \leq 2\varepsilon,\ |\tilde{U}| \geq \varepsilon/24\}, \qquad \mathcal{U} = \{|\tilde{U}| \geq \varepsilon/24\},$$
$$\mathcal{V} = \{|L' + U'| \leq 2\varepsilon\}, \qquad \overline{S}(x) := N^{-1/2} \sum_{Y_i \in \mathbb{Y} \setminus \mathbb{Y}_3} \psi(x, Y_i), \qquad x \in \mathcal{X}.$$

We observe that

$$\tilde{I}(\varepsilon) = \mathbf{E}\mathbb{I}_{\mathcal{U}}\mathbb{I}_{\mathcal{V}} \tag{67}$$

and note that the random function $x \to \overline{S}(x)$ is a sum of iid random variables with values in L^r such that, for every i, we have $\|\psi(\cdot, Y_i)\|_r \leq N^{\alpha}$ for almost all values of Y_i. By Lemma 13,

$$\mathbf{P}\{\|\overline{S}\|_r > N^{\nu}\} \leq N^{-3}.$$

Therefore in (67) we can replace the event $\mathcal{V}$ by $\mathcal{V}_1 = \mathcal{V} \cap \{\|\overline{S}\|_r \le N^{\nu}\}$. Furthermore, since $\tilde{U}$ does not depend on $\mathbb{Y}_3$, we have $\mathbf{E}\mathbb{I}_{\mathcal{U}}\mathbb{I}_{\mathcal{V}_1} = \mathbf{E}\mathbb{I}_{\mathcal{U}}p'$, where $p' := \mathbf{E}\big(\mathbb{I}_{\mathcal{V}_1}|\mathbb{Y}_1, \mathbb{Y}_2\big)$. The concentration bound for the conditional probability p', which is shown below,

$$p' \le c_*(\varepsilon + N^{-1/2}) \tag{68}$$

implies

$$\tilde{I}(\varepsilon) \le c_*(\varepsilon + N^{-1/2})\mathbf{P}\{\mathcal{U}\} \le c_*(\varepsilon + N^{-1/2})\varepsilon^{-r}N^{-r/2}. \tag{69}$$

In the last step we applied Markov's inequality

$$\mathbf{P}\{\mathcal{U}\} \le (24/\varepsilon)^r N^{-r/2}\mathbf{E}|N^{1/2}\tilde{U}|^r$$

and the bound $\mathbf{E}|N^{1/2}\tilde{U}|^r \le c_*\mathbf{E}|N^{1/2}U_{ij}|^r \le c_*$. Here U_{ij} denotes the random variable obtained from $\tilde{U}$ after we replace Y_j by X_j for every j. The second last inequality follows from (47). The last inequality follows from the well known moment inequalities for U-statistics [20].

It follows from (69) and the simple inequality $\varepsilon \ge b_N \ge c_*N^{-1/2}$ that

$$\int_{b_N}^1 \frac{d\varepsilon}{\varepsilon^2}\tilde{I}(\varepsilon) \le \frac{c_*}{N^{r/2}}\int_{b_N}^1 \frac{d\varepsilon}{\varepsilon^{1+r}} \le \frac{c_*}{N^{r/2}b_N^r} = c_*m^{-r/2}\ln^r N \le c_*N^{-5\nu},$$

provided that $m^{r/2} \ge N^{6\nu}$. The latter inequality is ensured by (17). Thus we have shown (64) for $\tilde{I}(\varepsilon)$.

It remains to prove (68). We write $L' + U'$ in the form $L_* + U_* + b - x$, where

$$L_* = \frac{1}{N^{1/2}}\sum_{Y_j\in\mathbb{Y}_3}\big(g(Y_j) + N^{-1/2}\overline{S}(Y_j)\big) \qquad \text{and} \qquad U_* = \frac{1}{N^{3/2}}\sum_{\{Y_j,Y_k\}\subset\mathbb{Y}_3}\psi(Y_j, Y_k),$$

and where b is a function of $\{Y_i \in \mathbb{Y} \setminus \mathbb{Y}_3\}$. Introduce the random variables $\overline{L}$ and $\overline{U}$ which are obtained from L_* and U_* after we replace every $Y_j \in \mathbb{Y}_3$ by the corresponding observation X_j. We have

$$\begin{aligned} p' &\le \sup_{v\in R}\mathbf{E}\big(\mathbb{I}_{\{L_*+U_*\in[v,v+2\varepsilon]\}}\big|\mathbb{Y}_1, \mathbb{Y}_2\big)\mathbb{I}_{\{\|\overline{S}\|_r\le N^{\nu}\}} \\ &\le c_*\sup_{v\in R}\mathbf{E}\big(\mathbb{I}_{\{\overline{L}+\overline{U}\in[v,v+2\varepsilon]\}}\big|\mathbb{Y}_1, \mathbb{Y}_2\big)\mathbb{I}_{\{\|\overline{S}\|_r\le N^{\nu}\}}. \end{aligned}$$

In the last step we applied (47). Now an application of the Berry–Esseen bound due to van Zwet [37] shows (68). The proof of the second inequality of (64) is complete.
Step A.2. Proof of the first inequality of (64). We introduce events

$$\mathbb{A} = \{\tilde{\alpha}^2 > 1 - \varepsilon_m^2\}, \quad \mathbb{V} = \{\|S\|_r \le N^{\nu}\}, \quad \mathbb{L} = \{|L| < 2\varepsilon\}$$

(recall that ε_m is defined in (53)) and write $I_1(\varepsilon)$ in the form $I_1(\varepsilon) = \mathbf{E}\,\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{S}}\mathbb{I}_{\mathbb{L}}$. We have

$$I_1(\varepsilon) = \mathbf{E}\,\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{S}}\mathbb{I}_{\mathbb{L}} \le \mathbf{E}\,\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{V}}\mathbb{I}_{\mathbb{L}}.$$

To upper bound $I_1(\varepsilon)$ we use the following strategy. We can upper bound the probability $\mathbf{P}\{\mathbb{L}\}$ using the Berry–Esseen inequality,

$$\mathbf{P}\{\mathbb{L}\} \le c_*(\varepsilon + N^{-1/2}). \tag{70}$$

Furthemore, one can show that the probability $\mathbf{P}\{\mathbb{A}\} = O(N^{-6\nu})$. We are going to make use of both of these bounds. However, since the events $\mathbb{A}$ and $\mathbb{L}$ refer to the same set of random variables $Y_{m+1}, \dots, Y_N$, we cannot argue directly that $\mathbf{E}\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{L}} \approx \mathbf{P}\{\mathbb{A}\}\mathbf{P}\{\mathbb{L}\}$. Nevertheless, invoking a complex conditioning argument we are able to show that

$$I_1(\varepsilon) \le c_*\mathcal{R}(\varepsilon + N^{-1/2}) + c_*N^{-2}, \qquad \mathcal{R} := N^{-6\nu}(1 + \delta_3^{-1}). \tag{71}$$

The latter inequality together with the inequalities $\varepsilon \ge b_N > N^{-1/2}$ imply the first part of (64). Let us prove (71). As the proof is rather involved we start by providing an outline. Let the integers n and M be defined by (29). Split $\{1, \dots, N\} = O_0 \cup O_1 \cup \dots \cup O_n$, where $O_0 = \{1, \dots, m\}$ and where the sets O_i, for $1 \le i \le n$, are defined in (30). Split L, see (63),

$$L = \sum_{k=0}^{n} L_k - x, \qquad \text{where} \qquad L_k = N^{-1/2} \sum_{j \in O_k} g(Y_j), \qquad \text{for} \qquad k = 1, \dots, n, \tag{72}$$

and where $L_0 = N^{-1/2} \sum_{j \in O_0} g(X_j)$. Observe, that $\mathbb{I}_{\mathbb{L}}$ is a function of $L_0, L_1, \dots, L_n$. The random variables $\mathbb{I}_{\mathbb{A}}$ and $\mathbb{I}_{\mathbb{V}}$ are functions of $Y_{m+1}, \dots, Y_N$ and do not depend on $X_1, \dots, X_m$. Therefore, denoting

$$m(l_1, \dots, l_n) = \mathbf{E}(\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{V}} | L_1 = l_1, \dots, L_n = l_n) \qquad \text{and} \qquad \mathcal{M} = ess \sup m(l_1, \dots, l_n)$$

we obtain from (70)

$$\mathbf{E}\,\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{V}}\mathbb{I}_{\mathbb{L}} = \mathbf{E}\,\mathbb{I}_{\mathbb{L}} m(L_1, \dots, L_n) \le c_*(\varepsilon + N^{-1/2})\mathcal{M}. \tag{73}$$

Clearly, the bound $\mathcal{M} \le c_*\mathcal{R}$ would imply (71). Unfortunately, we are not able to establish such a bound directly. In what follows we prove (71) using a delicate conditioning which allows us to estimate quantities like $\mathcal{M}$.

Step A.2.1. Firstly we replace L_k, $1 \le k \le n$, by smooth random variables

$$g_k = \frac{1}{N}\frac{\xi_k}{n^{1/2}} + L_k, \tag{74}$$

where $\xi_1, \ldots, \xi_n$ are symmetric i.i.d. random variables with the density function defined by (22) with $k = 6$ and $a = 1/6$ so that the characteristic function $t \to \mathbf{E}\exp\{it\xi_1\}$ vanishes outside the unit interval $\{t : |t| < 1\}$. Note that $\mathbf{E}\xi_1^4 < \infty$. We assume that the sequences $\xi_1, \xi_2, \ldots$ and $X_1, \ldots, X_m, Y_{m+1}, \ldots, Y_N$ are independent. In particular, ξ_k and L_k are independent.

Introduce the event

$$\tilde{\mathbb{L}} = \left\{ \left| L_0 + \sum_{k=1}^{n} g_k - x \right| < 3\varepsilon \right\}.$$

Note that

$$\mathbb{I}_{\mathbb{L}} \le \mathbb{I}_{\tilde{\mathbb{L}}} + \mathbb{I}_{\{|\xi| \ge \varepsilon N\}}, \qquad \text{where} \qquad \xi = \frac{1}{n^{1/2}} \sum_{k=1}^{n} \xi_k.$$

Using Markov's inequality and the inequality $\mathbf{E}\xi^4 \le c$ we estimate the probability

$$\mathbf{P}\{|\xi| \ge \varepsilon N\} \le \frac{\mathbf{E}\xi^4}{\varepsilon^4 N^4} \le \frac{c}{\varepsilon^4 N^4} \le \frac{c_*}{N^2},$$

where in the last step we used $\varepsilon^2 N \ge b_N^2 N \ge c'_*$. Hence we have

$$\mathbf{E}\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{V}}\mathbb{I}_{\mathbb{L}} \le \mathbf{E}\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{V}}\mathbb{I}_{\tilde{\mathbb{L}}} + c_* N^{-2}. \tag{75}$$

In the subsequent steps of the proof we replace the conditioning on $L_1, \ldots, L_n$ (in (73)) by the conditioning on the random variables $g_1, \ldots, g_n$. Since the latter random variables have densities (their densities are analysed in Lemma 7 below) the corresponding conditional distributions are much easier to handle. Moreover, we restrict the conditioning on the event where these densities are positive.

Step A.2.2. Given $w > 0$, consider the events $\{|g_k| \le n^{-1/2} w\}$ and their indicator functions $\mathbb{I}_k = \mathbb{I}_{\{|g_k| \le n^{-1/2} w\}}$. Using the simple inequality $n\mathbf{E}g_k^2 \le c_*$ (where c_* depends on M_* and r) we obtain from Chebyshev's inequality that

$$\mathbf{P}\{\mathbb{I}_k = 1\} = 1 - \mathbf{P}\{|g_k| > n^{-1/2} w\} \ge 1 - w^{-2} n \mathbf{E}|g_k|^2 > 7/8, \tag{76}$$

where the last inequality holds for a sufficiently large constant w (depending on M_*, r). Fix w such that (76) holds and introduce the event $\mathbb{B}^* = \{\sum_{k=1}^n \mathbb{I}_k > n/4\}$. Hoeffding's inequality shows $\mathbf{P}\{\mathbb{B}^*\} \ge 1 - \exp\{-n/8\}$. Therefore,

$$\mathbf{E}\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{V}}\mathbb{I}_{\tilde{\mathbb{L}}} \le \mathbf{E}\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{V}}\mathbb{I}_{\tilde{\mathbb{L}}}\mathbb{I}_{\mathbb{B}^*} + c_* N^{-2}. \tag{77}$$

Given a binary vector $\theta = (\theta_1, \ldots, \theta_n)$ (with $\theta_k \in \{0; 1\}$) write $|\theta| = \sum_k \theta_k$. Introduce the event $\mathbb{B}_\theta = \{\mathbb{I}_k = \theta_k,\ 1 \le k \le n\}$ and the conditional expectation

$$m_\theta(z_1, \ldots, z_n) = \mathbf{E}(\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{V}}\mathbb{I}_{\mathbb{B}_\theta} \mid g_1 = z_1, \ldots, g_n = z_n), \qquad (z_1, \ldots, z_n) \in \mathbb{R}^n.$$

Note that $\mathbb{I}_{\mathbb{B}_\theta}$, the indicator of the event $\mathbb{B}_\theta$, is a function of $g_1, \ldots, g_n$. It follows from the identities

$$\mathbb{B}^* = \cup_{|\theta|>n/4}\mathbb{B}_\theta \qquad \text{and} \qquad \mathbb{I}_{\mathbb{B}^*} = \sum_{|\theta|>n/4} \mathbb{I}_{\mathbb{B}_\theta}$$

(here $\mathbb{B}_\theta \cap \mathbb{B}_{\theta'} = \emptyset$, for $\theta \neq \theta'$) that

$$\mathbf{E}\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{V}}\mathbb{I}_{\tilde{\mathbb{L}}}\mathbb{I}_{\mathbb{B}^*} = \sum_{|\theta|>n/4} \mathbf{E}\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{V}}\mathbb{I}_{\tilde{\mathbb{L}}}\mathbb{I}_{\mathbb{B}_\theta} = \sum_{|\theta|>n/4} \mathbf{E}\mathbb{I}_{\mathbb{B}_\theta}\mathbb{I}_{\tilde{\mathbb{L}}} m_\theta(g_1, \ldots, g_n).$$

We shall show below that uniformly in θ, satisfying $|\theta| > n/4$, we have

$$M_\theta \leq c_* \mathcal{R}, \qquad \text{where} \qquad M_\theta := \text{ess sup}\, m_\theta(z_1, \ldots, z_n). \tag{78}$$

This bound in combination with (70), which extends to $\tilde{\mathbb{L}}$ as well, implies

$$\begin{aligned}\mathbf{E}\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{V}}\mathbb{I}_{\tilde{\mathbb{L}}}\mathbb{I}_{\mathbb{B}^*} &\leq c_*\mathcal{R} \sum_{|\theta|>n/4} \mathbf{E}\mathbb{I}_{\mathbb{B}_\theta}\mathbb{I}_{\tilde{\mathbb{L}}} = c_*\mathcal{R}\mathbf{E}\mathbb{I}_{\mathbb{B}^*}\mathbb{I}_{\tilde{\mathbb{L}}} \\ &\leq c_*\mathcal{R}\mathbf{P}\{\tilde{\mathbb{L}}\} \leq c_*\mathcal{R}(\varepsilon + N^{-1/2}).\end{aligned}$$

Combining the latter inequalities with (75) and (77) we obtain (71).
Step A.2.3. Here we show (78). Fix $\theta = (\theta_1, \ldots, \theta_n)$ satisfying $|\theta| > n/4$. Denote, for brevity, $h = |\theta|$ and assume without loss of generality that $\theta_i = 1$, for $1 \leq i \leq h$, and $\theta_j = 0$, for $h + 1 \leq j \leq n$. Consider the $h-$dimensional random vector $\overline{g}_{[\theta]} = (g_1, \ldots, g_h)$. Note that the random vector $\overline{g}_{[\theta]}$ and the sequences of random variables

$$\mathbb{Y}_\theta = \big\{Y_j : m + hM < j \leq N\big\}, \qquad \xi_\theta = \{\xi_j : h < j \leq n\}$$

are independent. Recall S from (55) and note that the terms S_θ and S'_θ of the decomposition

$$S = S_\theta + S'_\theta, \qquad S_\theta(\cdot) = \frac{1}{\sqrt{N}} \sum_{1\leq k\leq h} \sum_{j\in O_k} \psi(\cdot, Y_j)$$

are independent as well.

For $\overline{z}_{[\theta]} = (z_1, \ldots, z_h) \in \mathbb{R}^h$ we have $m_\theta(z_1, \ldots, z_n) \leq \tilde{m}_\theta(\overline{z}_{[\theta]})$, where

$$\tilde{m}_\theta(\overline{z}_{[\theta]}) = \text{ess sup}_\theta \mathbf{E}\big(\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{V}}\mathbb{I}_{\mathbb{B}_\theta} \,\big|\, \overline{g}_{[\theta]} = \overline{z}_{[\theta]},\ \mathbb{Y}_\theta,\ \xi_\theta\big)$$

denotes the "ess sup" taken with respect to almost all values of $\mathbb{Y}_\theta$ and ξ_θ. To prove (78) we show that

$$\tilde{m}_\theta(\overline{z}_{[\theta]}) \leq c_*\mathcal{R}. \tag{79}$$

Let us prove (79). Given $\mathbb{Y}_\theta$, denote $f_\theta = S'_\theta$ (note that S'_θ is a function of $\mathbb{Y}_\theta$). Using the notation (40), we have for the interval $J'_p = N^{-1/2} J_p$,

$$\mathbf{E}\big(\mathbb{I}_{\mathbb{A}}\mathbb{I}_{\mathbb{V}}\mathbb{I}_{\mathbb{B}_\theta} \,\big|\, \overline{g}_{[\theta]} = \overline{z}_{[\theta]},\ \mathbb{Y}_\theta,\ \xi_\theta\big) = \mathbb{I}_{\mathbb{B}_\theta}\mathbf{E}\big(d_{\varepsilon_m}(f_\theta + S_\theta, J'_p) \,\big|\, \overline{g}_{[\theta]} = \overline{z}_{[\theta]},\ \mathbb{Y}_\theta,\ \xi_\theta\big). \tag{80}$$

Note that the factor $\mathbb{I}_{\mathbb{B}_\theta}$ in the right side is non zero whenever $\overline{z}_{[\theta]} = (z_1, \ldots, z_h)$ satisfies $|z_i| \le w/\sqrt{n}$, for $i = 1, \ldots, h$. Introduce L^r valued random variables

$$U_i = N^{-1/2} \sum_{j \in O_i} \psi(\cdot, Y_j), \qquad i = 1, \ldots, h,$$

and the regular conditional probability

$$P(\overline{z}_{[\theta]}; \mathcal{A}) = \mathbf{E}\big(\mathbb{I}_{\{(U_1,\ldots,U_h)\in\mathcal{A}\}} \,\big|\, \overline{g}_{[\theta]} = \overline{z}_{[\theta]}\big).$$

Here $\mathcal{A}$ denotes a Borel subset of $L^r \times \cdots \times L^r$ (h-times). By independence, there exist regular conditional probabilities

$$P_i(z_i;\ \mathcal{A}_i) = \mathbf{E}(\mathbb{I}_{U_i\in\mathcal{A}_i} \,\big|\, g_i = z_i), \qquad i = 1, \ldots, h, \tag{81}$$

such that for Borel subsets $\mathcal{A}_i$ of L^r we have

$$P(\overline{z}_{[\theta]}; \mathcal{A}_1 \times \cdots \times \mathcal{A}_h) = \prod_{1\le i\le h} P_i(z_i; \mathcal{A}_i).$$

In particular, for every $\overline{z}_{[\theta]}$, the regular conditional probability $P(\overline{z}_{[\theta]}; \cdot)$ is the (measure theoretical) extension of the product of the regular conditional probabilities (81). Therefore, denoting by ψ_i a random variable with values in L^r and with the distribution

$$\mathbf{P}\{\psi_i \in \mathcal{B}\} = P_i(z_i; \mathcal{B}), \qquad \mathcal{B} \subset L^r - \text{Borel set}, \tag{82}$$

we obtain that the distribution of the sum

$$\zeta = \psi_1 + \cdots + \psi_h \tag{83}$$

of independent random variables $\psi_1, \ldots, \psi_h$ is the regular conditional distribution of S_θ, given $\overline{g}_{[\theta]} = \overline{z}_{[\theta]}$. In particular, the expectation in the right side of (80) equals $\delta_{\varepsilon_m}(f_\theta + \zeta)$, where

$$\delta_s(f_\theta + \zeta) := \mathbf{E}_\zeta d_s(f_\theta + \zeta, J'_p), \qquad s > 0, \tag{84}$$

and where $\mathbf{E}_\zeta$ denotes the conditional expectation given all the random variables, but ζ. Recall ε_m defined by (53) and note that for any ε_* satisfying the inequality

$$\varepsilon_m \le \varepsilon_* \tag{85}$$

we have

$$\delta_{\varepsilon_m}(f_\theta+\zeta)\le\delta_{\varepsilon_*}(f_\theta+\zeta). \tag{86}$$

We put $\varepsilon_*=\mu_*|T_0|N^{-1/2}/20$ and apply Lemma 1 to upper bound $\delta_{\varepsilon_*}(f_\theta+\zeta)$ (the quantity μ_* is defined in (97) below). We will use the inequalities $c_*\delta_3^2/n\le\mu_*^2\le c_*'\delta_3^2/n$ that follow from (217) below. Note that for T_0 satisfying (38), integers m,n as in (17), (29), and the quantity δ_3 (see (41)) satisfying

$$\delta_3^2\ge N^{-8\nu}, \tag{87}$$

the inequality (85) holds, provided that N is sufficiently large ($N>C_*$). Moreover, we have

$$\varepsilon_*^2\le c_*\delta_3^2N^{-48\nu}. \tag{88}$$

Now Lemma 1 (together with the moment inequalities of Lemma 10) implies the inequality

$$\delta_{\varepsilon_*}(f_\theta+\zeta)\le c_*\kappa_*^{1/2}\varepsilon_*^{(r-2)/(2r)}+c_*N^{-2}, \tag{89}$$

where the number κ_*, defined in (97), satisfies $\kappa_*\le c_*\delta_3^{-r/(r-2)}$, by (218).

Denote $\tilde r=r^{-1}+(r-2)^{-1}$. It follows from (89), (88) and (86), for $r>4$, that

$$\delta_{\varepsilon_m}(f_\theta+\zeta)\le c_*\delta_3^{-\tilde r}N^{-6\nu}+c_*N^{-2}\le c_*(1+\delta_3^{-\tilde r})N^{-6\nu}\le c_*\mathcal{R}. \tag{90}$$

In the last step we used the simple bound $\delta_3^2\le c_*$, see (200), and the inequality $1+\delta_3^{-\tilde r}\le 2+\delta_3^{-1}$, which holds for $\tilde r<1$. Note that (90) and (80), (84) imply (79). The proof the proof of the first inequality of (60) is complete.

Step B. Here we prove the second bound of (60). It is convenient to write the L^r-valued random variable (55) in the form

$$S=U_1+\cdots+U_{n-1}+U_n=:S'+U_n,\qquad\text{where}\qquad U_i=N^{-1/2}\sum_{j\in O_i}\psi(\cdot,Y_j). \tag{91}$$

Observe that $U_1,\dots,U_{n-1}$ are independent and identically distributed. We are going to apply Lemma 1 conditionally, given U_n, to the probability

$$\mathbf{P}\{\mathbb{B}\}=\mathbf{E}\tilde p(U_n),\qquad\text{where}\qquad \tilde p(f)=\mathbf{E}\big(d_{\varepsilon_m}(S'+f,\ N^{-1/2}J_p)\big|U_n=f\big).$$

To upper bound $\tilde p(f)$ we proceed similarly as in the proof of (90). Lemma 9 shows that $U_1,\dots,U_{n-1}$ satisfy the moment conditions of Lemma 1. Note that in this case the quantity μ_* satisfies $c_*\delta_3^2/n\le\mu_*^2\le c_*'/n$ (these inequalities follow from (201)). The right inequality implies the bound $\varepsilon_*\le c_*N^{-48\nu}$ instead of (88) above. As a result we obtain a different power of δ_3 in the upper bound below. Proceeding as in proof of (90), see (86), (88), (89), we obtain

$$\tilde p(f)\le c_*(1+\delta_3^{-r/2(r-2)})N^{-6\nu}\le c_*\mathcal{R}.$$

In the last step we used the inequality $1+\delta_3^{-r/2(r-2)} \leq 2+\delta_3^{-1}$, which follows from $r/2(r-2) < 1$ (recall that $r > 4$). Therefore, we have $\mathbf{P}\{\mathbb{B}\} \leq \mathbf{E}\tilde{p}(U_n) \leq c_*\mathcal{R}$, where $\mathcal{R}$ is defined in (71). This completes the proof of the second inequality in (60).

3.4 Proof of (26) for $k = 3$

Here we prove the bound $|I_3| \leq c_* N^{-1-\nu}$, see (26). It follows from (48) and the identity $\mathbf{E}_{\mathbb{Y}} \exp\{itT'\} = \alpha_t^m \exp\{itW'\}$, see (13), that

$$|I_3| \leq \int_{t_1<|t|<t_2} \frac{\mathbf{E}|\alpha_t^m|}{|t|} dt + c_* N^{-1-\nu}. \tag{92}$$

Recall the event $\mathbb{S} = \{\|S\|_r < N^{\nu/10}\}$, where S defined in (55). We have

$$\mathbf{E}|\alpha_t^m| \leq \mathbf{E}\mathbb{I}_{\mathbb{S}}|\alpha_t^m| + \mathbf{E}(1-\mathbb{I}_{\mathbb{S}}). \tag{93}$$

Using Lemma 13 we upper bound the second term on the right: $\mathbf{P}\{\|S\|_r \geq N^{\nu/10}\} \leq c_* N^{-3}$. Furthermore, the one-term expansion of the exponent in (13) in powers of $itN^{-3/2}\sum_{j=m+1}^N \psi(X_1, Y_j)$ shows the inequality

$$\mathbb{I}_{\mathbb{S}}|\alpha_t| \leq |\mathbf{E}\exp\{itN^{-1/2}g(X_1)\}| + \mathbb{I}_{\mathbb{S}}|t|N^{-1}\|S\|_1.$$

It follows from (7) that the first summand is bounded from above by $1-v$, for some $v > 0$ depending on A_*, M_*, D_*, δ only, see the proof of (36). Furthermore, the second summand is bounded from above by $N^{-9\nu/10}$ almost surely. Therefore, for sufficiently large $N > C_*$ we have $\mathbb{I}_{\mathbb{S}}|\alpha_t| \leq 1 - v/2$ uniformly in N. Invoking this bound in (93) we obtain

$$\mathbf{E}|\alpha_t^m| \leq (1-v/2)^m + c_* N^{-3} \leq c_* N^{-3},$$

for m satisfying (17). The latter inequality implies that the integral in (92) is bounded from above by $c_* N^{-2}$ thus completing the proof.

4 Combinatorial concentration bound

We start the section by introducing some notation and collecting auxiliary inequalities. Then we formulate and prove Lemmas 1 and 2.

Introduce the number

$$\delta_2 = \min\left\{\frac{1}{12c_g}, \frac{(c_r\|g\|_2^2/2^r\|g\|_r^r)^{1/(r-2)}}{1+4/\|g\|_2}\right\}, \tag{94}$$

where $c_g = 1 + \|g\|_r/\|g\|_2$ and $c_r = (7/24)2^{-(r-1)}$. Denote

$$\rho^* = 1 - \sup\{|\mathbf{E}e^{itg(X_1)}| : 2^{-1}\delta_2 \le |t| \le N^{-\nu+1/2}\}.$$

It follows from the identity $\rho^* = \rho(2^{-1}\sigma\delta_2,\ \sigma N^{-\nu+1/2})$ and the simple inequality $a_1 \le \delta_2/4$, see (35), that $\rho^* \ge \rho(2\sigma\, a_1, \sigma N^{-\nu+1/2})$. Furthermore, it follows from (169) and the assumption $\sigma_{\mathbb{T}}^2 = 1$ that $1/2 < \sigma < 2$ for sufficiently large N ($N > C_*$). Therefore, $\rho^* \ge \rho(a_1, 2N^{-\nu+1/2}) \ge \delta'$, where the last inequality follows from (36). We obtain, for $N > C_*$,

$$1 - \sup\{|\mathbf{E}e^{itg(X_1)}| : 2^{-1}\delta_2 \le |t| \le N^{-\nu+1/2}\} \ge \delta', \tag{95}$$

where the number δ' depends on $A_*, D_*, M_*, \nu_1, r, s, \delta$ only. In what follows we use the notation $c_0 = 10$. Let $L_0^r = \{y \in L^r : \int_{\mathcal{X}} y(x)P_X(dx) = 0\}$ denotes a subspace of L^r. Observe, that $\mathbf{E}g(X_1) = 0$ implies $y^*(= p_g(y)) \in L_0^r$, for every $y \in L_0^r$.

4.1 Lemma 1

Let $\psi_1, \dots, \psi_n$ denote independent random vectors with values in L_0^r. For $k = 1, \dots, n$, write

$$\zeta_k = \psi_1 + \cdots + \psi_k \qquad \text{and} \qquad \zeta = \zeta_n.$$

Let $\overline{\psi}_i$ denote an independent copy of ψ_i. Write $\psi_i^* = p_g(\psi_i)$ and $\overline{\psi}_i^* = p_g(\overline{\psi}_i)$, see (31). Introduce random vectors

$$\tilde{\psi}_i = 2^{-1}(\psi_i - \overline{\psi}_i), \qquad \tilde{\psi}_i^* = 2^{-1}(\psi_i^* - \overline{\psi}_i^*), \qquad \hat{\psi}_i = 2^{-1}(\psi_i + \overline{\psi}_i).$$

We shall assume that, for some $c_A \ge c_D \ge c_B > 0$,

$$n^{r/2}\mathbf{E}\|\tilde{\psi}_i\|_r^r \le c_A^r, \qquad c_B^2 \le n\,\mathbf{E}\|\tilde{\psi}_i^*\|_2^2 \le c_D^2, \tag{96}$$

for every $1 \le i \le n$. Furthermore, denote $\mu_i^2 = \mathbf{E}\|\tilde{\psi}_i^*\|_2^2$ and $\tilde{\kappa}_i^{r-2} = \frac{8}{3}\frac{\mathbf{E}\|\tilde{\psi}_i\|_r^r}{\mu_i^r}$,

$$\mu_* = \min_{1\le i\le n} \mu_i, \qquad \kappa_* = \max_{1\le i\le n} \tilde{\kappa}_i. \tag{97}$$

Observe that, by Hölder's inequality and (32), we have $\tilde{\kappa}_i > 1$, for $i = 1, \dots, n$.

Lemma 1 *Let $4 < r \le 5$ and $0 < \nu < 10^{-2}(r-4)$. Assume that $n \ge N^{5\nu}$. Suppose that*

$$\kappa_*^4 \le \frac{9}{256}\frac{n}{\ln N}. \tag{98}$$

Assume that (95), (96) as well as (106), (112) (below) hold. There exist a constant $c_* > 0$ which depends on $r, s, \nu, A_*, D_*, M_*, \delta$ only such that for every T_0 satisfying

(38) *we have*

$$\delta_{\varepsilon_*}(f+\zeta, I(T_0)) \le c_*(C_D/C_B)^{1/2}\kappa_*^{1/2}\varepsilon_*{}^{(r-2)/2r} + c_* N^{-2}, \tag{99}$$

for an arbitrary non-random element $f \in L_0^r$. *Here* $\varepsilon_* = \frac{\mu_*}{2c_0}\frac{|T_0|}{\sqrt{N}}$. *The function* $\delta_s(\cdot, I(T_0))$, *is defined in* (40).

In *Step A.2.3* of Sect. 3 we apply this lemma to random vector $\zeta = \psi_1 + \cdots + \psi_h$, see (83). In *Step B* of Sect. 3 we apply this lemma to the random vector S', see (91).

Proof We shall consider the case where $T_0 > 0$. For $T_0 < 0$ the proof is the same. We can assume without loss of generality that $c_0 < N^\nu$. Denote $X = \|\tilde{\psi}_i^*\|_2$ and $Y = \|\tilde{\psi}_i\|_r$ and $\mu = \mu_i$, $\kappa = \tilde{\kappa}_i$. By (32), we have $Y \ge X$.

Step 1. Here we construct the bound (100), see below, for the probability $\mathbf{P}\{B_i\}$, where

$$B_i = \{X \ge \mu/2,\ Y < \kappa\mu\}.$$

Write

$$\mu^2 = \mathbf{E}X^2 = \mathbf{E}X^2 I_A + \mathbf{E}X^2 I_{B_i} + \mathbf{E}X^2 I_D,$$
$$A = \{X < \mu/2\}, \qquad D = \{X \ge \mu/2,\ Y \ge \kappa\mu\}.$$

Substitution of the bounds

$$\mathbf{E}X^2 I_A \le \frac{\mu^2}{4},$$
$$\mathbf{E}X^2 I_{B_i} \le \mathbf{E}Y^2 I_{B_i} \le (\kappa\mu)^2\mathbf{P}\{B_i\},$$
$$\mathbf{E}X^2 I_D \le \mathbf{E}Y^2 I_{\{Y\ge\kappa\mu\}} \le (\kappa\mu)^{2-r}\mathbf{E}Y^r$$

gives

$$\mu^2 \le 4^{-1}\mu^2 + \kappa^2\mu^2\mathbf{P}\{B_i\} + (\kappa\mu)^{2-r}\mathbf{E}Y^r.$$

Finally, invoking the identity $\kappa^{r-2} = (8/3)\mathbf{E}Y^r/\mu^r$ we obtain

$$\mathbf{P}\{B_i\} \ge \frac{3}{4\kappa^2} - \frac{\mathbf{E}Y^r}{(\kappa\mu)^r} = \frac{3}{4\kappa^2}\Big(1 - \frac{4\mathbf{E}Y^r}{3\mu^r\kappa^{r-2}}\Big) = \frac{3}{8\kappa^2} \ge \frac{3}{8\kappa_*^2} =: p. \tag{100}$$

Introduce the (random) set $J = \{i : B_i \text{ occurs}\} \subset \{1, \ldots, n\}$. Hoeffding's inequality applied to the random variable $|J| = \mathbb{I}_{B_1} + \cdots + \mathbb{I}_{B_n}$ shows

$$\mathbf{P}\{|J| \le \rho n\} \le \exp\{-np^2/2\} \le N^{-2}, \qquad \rho := p/2 = (3/16)\kappa_*^{-2}. \tag{101}$$

In the last step we invoke (98) and use (100).

Step 2. Here we introduce randomization. Note that for any $\alpha_i \in \{-1,+1\}$, $i = 1,\dots,n$, the distributions of the random vectors

$$(\psi_1,\dots,\psi_n) \qquad \text{and} \qquad \big(\alpha_1\tilde\psi_1+\hat\psi_1,\dots,\alpha_n\tilde\psi_n+\hat\psi_n\big)$$

coincide. Therefore, denoting

$$\tilde\zeta_n = \alpha_1\tilde\psi_1+\dots+\alpha_n\tilde\psi_n, \qquad \hat\zeta_n = \hat\psi_1+\dots+\hat\psi_n,$$

we have for $s > 0$,

$$\delta_s(f+\zeta, I(T_0)) = \delta_s(f+\tilde\zeta_n+\hat\zeta_n, I(T_0)),$$

for every choice of $\alpha_1,\dots,\alpha_n$. From now on let $\alpha_1,\dots,\alpha_n$ denote a sequence of independent identically distributed Bernoulli random variables independent of $\tilde\psi_i$, $\hat\psi_i$, $1 \le i \le n$, and with probabilities $\mathbf{P}\{\alpha_1 = 1\} = \mathbf{P}\{\alpha_1 = -1\} = 1/2$. Denoting by $\mathbf{E}_\alpha$ the expectation with respect to the sequence $\alpha_1,\dots,\alpha_n$ we obtain

$$\delta_s(f+\zeta, I(T_0)) = \mathbf{E}_\alpha\delta_s(f+\tilde\zeta_n+\hat\zeta_n, I(T_0)). \tag{102}$$

We are going to condition on $\tilde\psi_i$ and $\hat\psi_i$, $1 \le i \le n$, while taking expectations with respect to $\alpha_1,\dots,\alpha_n$. It follows from (101), (102) and the fact that the random variable $|J|$ does not depend on $\alpha_1,\dots,\alpha_n$ that

$$\delta_s(f+\zeta, I(T_0)) \le \mathbf{E}\mathbb{I}_{\{|J|\ge\rho n\}}\gamma_s(\tilde\psi_i,\ \hat\psi_i,\ 1\le i\le n) + N^{-2}, \tag{103}$$

where

$$\gamma_s(\tilde\psi_i,\ \hat\psi_i,\ 1\le i\le n) = \mathbf{E}_\alpha\mathbb{I}_{\{|J|\ge\rho n\}}\mathbb{I}_{\{\upsilon^2(f+\tilde\zeta_n+\hat\zeta_n)>1-s^2\}}\mathbb{I}_{\{\|f+\tilde\zeta_n+\hat\zeta_n\|_r\le N^\nu\}}$$

denotes the conditional expectation given $\tilde\psi_i$, $\hat\psi_i$, $1 \le i \le n$. Note that (99) is a consequence of (103) and of the bound

$$\gamma_{\varepsilon_*}(\tilde\psi_i,\ \hat\psi_i,\ 1\le i\le n) \le c_*\kappa_*^{1/2}\varepsilon_*^{(r-2)/(2r)}. \tag{104}$$

Let us prove this bound. Introduce the integers

$$n_0 = l-1 \qquad l = \lfloor \delta_2\varkappa^{-1}\varepsilon_*^{-(r-2)/r}\rfloor, \qquad \varkappa = 2c_0(C_D/C_B)\kappa_*.$$

Let us show that

$$n_0 \le \rho n. \tag{105}$$

It follows from the inequalities

$$\varepsilon_*^{-1} \le 2\frac{c_0}{c_B}N^\nu n^{1/2}, \quad N^{\nu(r-2)/r} \le N^\nu \le n^{1/r}, \quad \delta_2 \le \frac{3}{16}\Big(\frac{3}{8}\Big)^{1/(r-2)}$$

that

$$l \le \frac{\delta_2}{C_D}\frac{1}{k_*}\left(\frac{C_B}{2c_0}\right)^{2/r}\left(N^\nu n^{1/2}\right)^{(r-2)/r} \le \frac{3}{16}\frac{1}{k_*}\frac{C_B^{2/r}}{C_D}n^{1/2}.$$

Note that (98) implies $k_* \le n^{1/4}$. Therefore, the inequality

$$C_B^{2/r}C_D^{-1} \le n^{1/4} \tag{106}$$

implies $l \le (3/16)k_*^{-2}n = \rho n$. We obtain (105).

Given $\tilde{\psi}_i,\ \hat{\psi}_i,\ 1 \le i \le n$, consider the corresponding set J, say $J = \{i_1, \dots, i_k\}$. Assume that $k \ge \rho n$. From the inequality $\rho n \ge n_0$, see (105), it follows that we can choose a subset $J' \subset J$ of size $|J'| = n_0$. Split

$$\tilde{\zeta}_n = \sum_{i\in J'}\alpha_i\tilde{\psi}_i + \sum_{i\in J\setminus J'}\alpha_i\tilde{\psi}_i =: \zeta_* + \zeta'$$

and denote $f + \zeta' + \hat{\zeta}_n = f_*$. Note that $f_* \in L_0^r$ almost surely. Let

$$\tilde{\delta} = \mathbf{E}'\mathbb{I}_{\{\upsilon^2(f_*+\zeta_*)>1-\varepsilon_*^2\}}\mathbb{I}_{\{\|f_*+\zeta_*\|_r\le N^\nu\}},$$

where $\mathbf{E}'$ denotes the conditional expectation given all the random variables, but $\{\alpha_i,\ i \in J'\}$. The bound (104) would follow if we show that

$$\tilde{\delta} \le c_*\kappa_*^{1/2}\varepsilon_*^{(r-2)/(2r)}. \tag{107}$$

Step 3. Here we prove (107). Note that for $j \in J'$ the vectors

$$x_j = T_0N^{-1/2}\tilde{\psi}_j \qquad \text{and} \qquad x_j^* = p_g(x_j) = T_0N^{-1/2}\tilde{\psi}_j^*$$

satisfy

$$\|x_j^*\|_2 \ge c_0\varepsilon_*, \qquad \|x_j\|_r \le \varkappa\varepsilon_*, \qquad \varkappa = 2c_0(C_D/C_B)\kappa_*. \tag{108}$$

Given $A \subset J'$ denote

$$x_A = \sum_{i\in A}x_i - \sum_{i\in J'\setminus A}x_i, \qquad x_A^* = p_g(x_A).$$

We are going to apply Kleitman's theorem on symmetric partitions (see, e.g. the proof of Theorem 4.2, Bollobas [15]) to the sequence $\{x_j^*,\ j \in J'\}$ in L^2. Since for $j \in J'$ we have $\|x_j^*\|_2 \ge c_0\varepsilon_*$, it follows from Kleitman's theorem that the collection $\mathcal{P}(J')$ of all subsets of J' splits into non-intersecting non-empty classes $\mathcal{P}(J') = \mathcal{D}_1 \cup \cdots \cup \mathcal{D}_s$, such that the corresponding sets of linear combinations $V_t = \left\{x_A^*,\ A \in \mathcal{D}_t\right\}$, $t = 1, 2, \dots, s$, are sparse, i.e., given t, for $A, A' \in \mathcal{D}_t$ and $A \ne A'$ we have

$$\|x_A^* - x_{A'}^*\|_2 \ge c_0\varepsilon_*. \tag{109}$$

Furthermore, the number of classes s is bounded from above by $\binom{n_0}{\lfloor n_0/2 \rfloor}$.

Next, using Lemma 2 we shall show that given f_* the class $\mathcal{D}_t$ may contain at most one element $A \in \mathcal{D}_t$ such that

$$v^2(f_* + \tilde{x}_A) > 1 - \varepsilon_*^2, \qquad \|f_* + \tilde{x}_A\|_r \leq N^\nu, \qquad \tilde{x}_A := N^{1/2} T_0^{-1} x_A. \qquad (110)$$

This means that there are at most $\binom{n_0}{\lfloor n_0/2 \rfloor}$ different subsets $A \subset J'$ for which (110) holds. This implies (107)

$$\tilde{\delta} \leq 2^{-n_0} \binom{n_0}{\lfloor n_0/2 \rfloor} \leq c n_0^{-1/2} = c \delta_2^{-1/2} \varkappa^{1/2} \varepsilon_*^{\frac{r-2}{2r}}.$$

Finally, (99) follows from (103), (104), (107).

Given $f_* \in L_0^r$ let us show that there is no pair A, A' in $\mathcal{D}_t$ which satisfy (110). Fix $A, A' \in \mathcal{D}_t$. We have, by (108) and the choice of n_0,

$$\|x_A - x_{A'}\|_r \leq 2 \sum_{i \in J'} \|x_i\|_r \leq 2 n_0 \varkappa \varepsilon_* < 2 \delta_2 \varepsilon_*^{2/r}.$$

Denoting $S_A = f_* + \tilde{x}_A$ and $S_{A'} = f_* + \tilde{x}_{A'}$ we obtain

$$\|S_A - S_{A'}\|_r = N^{1/2} T_0^{-1} \|x_A - x_{A'}\|_r \leq 2 \delta_2 \varepsilon_*^{2/r} N^{1/2} T_0^{-1}. \qquad (111)$$

Assume that S_A and $S_{A'}$ satisfy the second inequality of (110), i.e., $\|S_A\|_r \leq N^\nu$ and $\|S_{A'}\|_r \leq N^\nu$. We are going to apply Lemma 2 to the vectors S_A and $S_{A'}$. In order to check the conditions of Lemma 2 note that (114) and (115) are verified by (108), (109) and (111). Furthermore, the inequalities $c_0 < N^\nu$ and

$$c_B \geq 2 N^{4\nu} (n/N)^{1/2}, \qquad (112)$$

imply $N^{2\nu - 1/2} \leq \varepsilon_*$. Finally, we can assume without loss of generality that $\varepsilon_* \leq c'_*$, where $c'_* := \min\{(\delta'/4)^{r/2}, (A_*^{1/2}/6)^{r/2}\}$. Otherwise (99) follows from trivial inequalities

$$\delta_{\varepsilon_*} \leq 1 \leq (\varepsilon_*/c'_*)^{(r-2)/2r} \leq c_* \varepsilon_*^{(r-2)/2r}$$

and the inequality $\kappa_* > 1$.

Now Lemma 2 implies $\min\{v^2(S_A),\ v^2(S_{A'})\} \leq 1 - \varepsilon_*^2$ thus completing the proof of Lemma 1. □

4.2 Lemma 2

Here we formulate and prove Lemma 2. Let us introduce first some notation. Given $y \in L^r (= L^r(\mathcal{X}, \mathcal{P}_\mathcal{X}))$ define the symmetrization $y_s \in L^r(\mathcal{X} \times \mathcal{X}, \mathcal{P}_\mathcal{X} \times \mathcal{P}_\mathcal{X})$ by $y_s(x, x') = y(x) - y(x')$, for $x, x' \in \mathcal{X}$. In what follows X_1, X_2 denote independent random variables with values in $\mathcal{X}$ and with the common distribution P_X. By $\mathbf{E}$ we denote the expectation taken with respect to P_X. For $h \in L^r$ we write

$$\mathbf{E}h = \mathbf{E}h(X_1) = \int_{\mathcal{X}} h(x) P_X(dx), \qquad \mathbf{E}e^{ih} = \mathbf{E}e^{ih(X_1)} = \int_{\mathcal{X}} e^{ith(x)} P_X(dx).$$

Furthermore, for $2 \le p \le r$, denote

$$\|y_s\|_p^p = \mathbf{E}|y(X_1) - y(X_2)|^p, \qquad \|y\|_p^p = \mathbf{E}|y(X_1)|^p.$$

Note that for $y \in L_0^r$ we have $y^*(= p_g(y)) \in L_0^r$ and, therefore,

$$\mathbf{E}|y^*(X_1) - y^*(X_2)|^2 = 2\mathbf{E}|y^*(X_1)|^2. \tag{113}$$

Let $y_1, \dots, y_k, f$ be non-random vectors in L^r. We shall assume that these vectors belong to the linear subspace L_0^r. Given non random vectors $\alpha = \{\alpha_i\}_{i=1}^k$ and $\alpha' = \{\alpha_i'\}_{i=1}^k$, with $\alpha_i, \alpha_i' \in \{-1, +1\}$, denote

$$S_\alpha = f + \sum_{i=1}^k \alpha_i y_i, \qquad S_{\alpha'} = f + \sum_{i=1}^k \alpha_i' y_i.$$

Lemma 2 *Let* $\varkappa > 0$. *Assume that* (95) *holds and suppose that*

$$N^{\nu-1/2} \le \varepsilon \le \min\{(\delta'/4)^{r/2}, (\|g\|_2/6)^{r/2}\}.$$

Given T_0*, satisfying* (38)*, write* $T^* = N^{1/2} T_0^{-1}$ *and assume that*

$$\|y_j^*\|_2 > c_0 T^* \varepsilon, \quad \|y_j\|_r \le \varkappa T^* \varepsilon, \quad j = 1, \dots, k. \tag{114}$$

Suppose that $\|S_\alpha\|_r \le N^\nu$ *and* $\|S_{\alpha'}\|_r \le N^\nu$ *and*

$$\|S_\alpha^* - S_{\alpha'}^*\|_2 \ge c_0 T^* \varepsilon, \qquad \|S_\alpha - S_{\alpha'}\|_r \le 2\delta_2 T^* \varepsilon^{2/r}. \tag{115}$$

Then $\min\{v^2(S_\alpha),\ v^2(S_{\alpha'}\} \le 1 - \varepsilon^2$.

Recall that the functionals $v(\cdot)$, $\tau(\cdot)$, $u_t(\cdot)$ and the interval $I = I(T_0)$ used in proof below are defined in (39).

Proof Note that $\delta_1 < 1/10$ and $\delta_2 < 1/12$. In particular, we have

$$9/10 \le 1 - \delta_1 \le |s/T_0| \le 1 + \delta_1 \le 11/10, \quad \text{for} \quad |s - T_0| < \delta_1 N^{-\nu+1/2}. \tag{116}$$

Step 1. Assume that the inequality $\min\{v^2(S_\alpha),\ v^2(S_{\alpha'}\} \le 1 - \varepsilon^2$ fails. Then for some $s, t \in I$ we have

$$1 - |u_t(S_\alpha)|^2 < \varepsilon^2, \qquad 1 - |u_s(S_{\alpha'})|^2 < \varepsilon^2, \tag{117}$$

see (39). Fix these s, t and denote

$$\tilde{X} = s(g + N^{-1/2} S_{\alpha'}) - t(g + N^{-1/2} S_\alpha).$$

We are going to apply the inequality (256),

$$1 - |\mathbf{E} e^{i(Y+Z)}|^2 \geq 2^{-1}(1 - |\mathbf{E} e^{iZ}|^2) - (1 - |\mathbf{E} e^{iY}|^2)$$

to $Z = -\tilde{X}$ and $Y = s(g + N^{-1/2} S_{\alpha'})$. It follows from this inequality and (117) that

$$\varepsilon^2 > 1 - |u_t(S_\alpha)|^2 = 1 - |\mathbf{E} e^{i(Y+Z)}|^2 \geq 2^{-1}(1 - |\mathbf{E} e^{-i\tilde{X}}|^2) - \varepsilon^2.$$

In view of the identity $|\mathbf{E} e^{-i\tilde{X}}| = |\mathbf{E} e^{i\tilde{X}}|$ we have

$$1 - |\mathbf{E} e^{i\tilde{X}}|^2 < 4\varepsilon^2. \tag{118}$$

Step 2. Here we shall show that (118) contradicts the second inequality of (115). Firstly, we collect some auxiliary inequalities. Write the decomposition (31) for S_α and $S_{\alpha'}$,

$$S_\alpha = a\, g + S_\alpha^*, \qquad S_{\alpha'} = a'\, g + S_{\alpha'}^*. \tag{119}$$

Decompose

$$\begin{aligned} \tilde{X} &= vg + h, \\ v &= (s - t)(1 + a\, N^{-1/2}) + (a' - a)s N^{-1/2}, \\ h &= (s - t) N^{-1/2} S_\alpha^* + s N^{-1/2}(S_{\alpha'}^* - S_\alpha^*), \end{aligned}$$

where $v \in \mathbb{R}$ and where $h \in L^r$ is L^2-orthogonal to g. An application of (34) to S_α^* and $S_{\alpha'}^* - S_\alpha^*$ gives

$$\|h\|_r \leq c_g N^{-1/2} \big(|s|\, \|S_{\alpha'} - S_\alpha\|_r + |s - t|\, \|S_\alpha\|_r \big). \tag{120}$$

Furthermore, it follows from the simple inequality

$$\|x + y\|_2^2 \geq 2^{-1} \|x\|_2^2 - \|y\|_2^2$$

that

$$\|h\|_2^2 \geq 2^{-1} s^2 N^{-1} \|S_{\alpha'}^* - S_\alpha^*\|_2^2 - (s - t)^2 N^{-1} \|S_\alpha^*\|_2^2. \tag{121}$$

Note that for a and a' defined in (119) we obtain from (33) and (115) that

$$|a| \leq \|S_\alpha\|_r \|g\|_2^{-1} \leq N^\nu \|g\|_2^{-1}, \tag{122}$$

$$|a' - a| \leq \|S_{\alpha'} - S_\alpha\|_r \|g\|_2^{-1} \leq 2\delta_2 \varepsilon^{2/r} N^{1/2} T_0^{-1} \|g\|_2^{-1}. \tag{123}$$

Step 4.2.1. Consider the case where, $|s-t| < \delta_2$. Invoking the inequalities $\|S_\alpha\|_r \le N^\nu$ and (115) we obtain from (120) that

$$\|h\|_r^r \le (4c_g)^r \delta_2^r \Big(N^{\nu r - r/2} + \varepsilon^2 |s|^r T_0^{-r} \Big).$$

Furthermore, using (116), (94), and $N^{\nu-1/2} \le \varepsilon$, we obtain for $4 \le r \le 5$

$$\|h\|_r^r \le 3^{-r}(\varepsilon^r + \varepsilon^2 (11/10)^r) \le 3^{1-r}\varepsilon^2. \tag{124}$$

Note that (32) implies $\|S_\alpha^*\|_2 \le \|S_\alpha\|_r \le N^\nu$. This inequality in combination with (115) and (121) gives

$$\|h\|_2^2 \ge 2^{-1}(s/T_0)^2 c_0^2 \varepsilon^2 - \delta_2^2 N^{2\nu-1}.$$

Invoking (116) and using $c_0 > 10$, $\delta_2 < 12^{-1}$, and $N^{\nu-1/2} \le \varepsilon$ we obtain

$$\|h\|_2^2 \ge (4/10) c_0^2 \varepsilon^2. \tag{125}$$

Now we are going to apply Lemma 12 statement **a)** to $\tilde{X} = vg + h$. For this purpose we verify the conditions of this lemma. Firstly, note that (125), (113) imply, $\|h_s\|_2^2 \ge (8/10) c_0^2 \varepsilon^2$. Furthermore, it follows from the simple inequality $\mathbf{E}|h(X_1) - h(X_2)|^r \le 2^r \mathbf{E}|h(X_1)|^r$ and (124) that $\|h_s\|_r^r \le 3(2/3)^r \varepsilon^2$. Therefore, we obtain, for $4 \le r \le 5$,

$$\|h_s\|_r^r \le \frac{6}{10}\varepsilon^2 \le c_0^{-2}\|h_s\|_2^2 \le c_r \|h_s\|_2^2, \qquad c_r = (7/24)2^{-(r-1)}.$$

Furthermore, the inequalities (122), (123) and (116) imply

$$|v| \le \delta_2 + \delta_2 \|g\|_2^{-1}(N^{\nu-1/2} + 2\varepsilon^{2/r}(11/10)) \le \delta_2(1 + 4\|g\|_2^{-1}),$$

for $N^{\nu-1/2} \le \varepsilon \le 1$. Invoking (94) and using the inequality $\|g_s\|_r^r \le 2^r \|g\|_r^r$ and the identity $\|g_s\|_2^2 = 2\|g\|_2^2$ we obtain

$$|v|^{r-2} \le \frac{c_r}{2^r}\frac{\|g\|_2^2}{\|g\|_r^r} \le \frac{c_r}{2^r}\frac{2^{-1}\|g_s\|_2^2}{2^{-r}\|g_s\|_r^r} \le \frac{c_r}{2}\frac{\|g_s\|_2^2}{\|g_s\|_r^r}$$

as required by Lemma 12 **a)**. This lemma implies

$$1 - |\mathbf{E}e^{i\tilde{X}}|^2 \ge 6^{-1}\|h_s\|_2^2 = 3^{-1}\|h\|_2^2.$$

In the last step we used (113). Now (125), for $c_0 \ge 10$, contradicts (118).

Step 4.2.2. Consider the case where $\delta_2 < |s-t| \le \delta_1 N^{-\nu+1/2}$. It follows from (120), (115) and (116) that

$$\begin{aligned} \mathbf{E}|h| \le \|h\|_r &\le c_g\big(2\delta_2\varepsilon^{2/r}|s/T_0| + \delta_1\big) \\ &\le c_g(\delta_1 + 3\delta_2\varepsilon^{2/r}) \le c_g\delta_1 + \varepsilon^{2/r}. \end{aligned} \tag{126}$$

In the last step we used $\delta_2 < 1/3$. From (122), (123) and (116), we obtain for $\delta_2 \le |s-t|$ and $N^{\nu-1/2} \le \varepsilon$,

$$\begin{aligned} |v| &\ge \delta_2(1 - N^{\nu-1/2}\|g\|_2^{-1}) - 2\delta_2\varepsilon^{2/r}|s/T_0|\|g\|_2^{-1} \\ &= \delta_2(1 - \|g\|_2^{-1}(\varepsilon + \varepsilon^{2/r}(22/10))) \\ &\ge \delta_2(1 - 3\varepsilon^{2/r}\|g\|_2^{-1}) \ge \delta_2/2, \end{aligned}$$

provided that $\varepsilon^{2/r} < \|g\|_2/6$. Similarly, using in addition, $\delta_1, \delta_2 < 1/4$ and $\varepsilon < \|g\|_2$, we obtain, for $|s-t| \le \delta_1 N^{-\nu+1/2}$,

$$\begin{aligned} |v| &\le |s-t|(1 + N^{\nu-1/2}\|g\|_2^{-1}) + 2\delta_2\varepsilon^{2/r}|s/T_0|\|g\|_2^{-1} \\ &\le |s-t|(1 + \varepsilon\|g\|_2^{-1}) + (22/10)\delta_2\varepsilon^{2/r}\|g\|_2^{-1} \\ &\le 2\,|s-t| + 1 \le N^{-\nu+1/2}. \end{aligned}$$

It follows from these inequalities, see (95), that

$$1 - |\mathbf{E}e^{i\tilde{X}}|^2 \ge 1 - |\mathbf{E}e^{i\tilde{X}}| \ge 1 - |\mathbf{E}e^{ivg}| - \mathbf{E}|h| \ge \delta' - \mathbf{E}|h|.$$

Finally, invoking (126) and (37), we get

$$1 - |\mathbf{E}e^{i\tilde{X}}|^2 \ge \delta' - c_g\delta_1 - \varepsilon^{2/r} \ge \delta'/2 > 4\varepsilon^2,$$

Once again we obtain a contradiction to (118), thus completing the proof. □

5 Expansions

Here we prove the bound

$$\int_{|t|\le t_1} \Big|\mathbf{E}e^{it\tilde{\mathbb{T}}} - \hat{G}(t)\Big|\frac{dt}{|t|} \le c_* N^{-1-\nu}, \tag{127}$$

where $t_1 = N^{1/2}/10^3\beta_3$. For the definition of $\tilde{\mathbb{T}}$ and $\hat{G}$ see Sect. 2.4. Here and below c_* denotes a constant depending on $A_*, M_*, D_*, r, s, \nu_1$ only. We prove (127) for sufficiently large N, that is, we shall assume that $N > C_*$, where C_* is a number depending on $A_*, M_*, D_*, r, s, \nu_1$ only. Note that for $N < C_*$, the bound (127) becomes trivial, since in this case the integral is bounded by a constant.

Let us first introduce some notation. Denote $\Omega_m = \{1, \ldots, m\}$. For $A \subset \Omega_N$ write $\mathbb{U}_1(A) = \sum_{j \in A} g_1(X_j)$. Given complex valued functions f, h we write $f \prec \mathcal{R}$ if

$$\int_{|t| \le t_1} |t^{-1} f(t)| dt \le c_* N^{-1-\nu}$$

and write $f \sim h$ if $f - h \prec \mathcal{R}$. In particular, (127) can be written in short $\mathbf{E} e^{it\tilde{\mathbb{T}}} \sim \hat{G}(t)$.

In order to prove (127) we show that

$$\mathbf{E} e^{it\tilde{\mathbb{T}}} \sim \mathbf{E} e^{it\mathbb{T}} \qquad \text{and} \qquad \mathbf{E} e^{it\mathbb{T}} \sim \hat{G}(t). \tag{128}$$

In what follows we use the notation of Sect. 2. We denote $\alpha(t) = \mathbf{E} e^{itg(X_1)}$. We assume that (16) holds.

5.1 Proof of the first relation of (128)

We have, see (19),

$$\mathbb{T} = \tilde{\mathbb{T}} + \tilde{\Lambda}_1 + \tilde{\Lambda}_2, \qquad \tilde{\Lambda}_1 = \Lambda_1 + \Lambda_4, \qquad \tilde{\Lambda}_2 = \Lambda_2 + \Lambda_3 + \Lambda_5,$$

where the random variables Λ_j are introduced in Sect. 2.4. We shall show that

$$\mathbf{E} e^{it\tilde{\mathbb{T}}} \sim \mathbf{E} e^{it(\tilde{\mathbb{T}} + \tilde{\Lambda}_1)} \qquad \text{and} \qquad \mathbf{E} e^{it(\tilde{\mathbb{T}} + \tilde{\Lambda}_1)} \sim \mathbf{E} e^{it\mathbb{T}}. \tag{129}$$

The second relation follows from the moment bounds of Lemma 5 via Taylor expansion. We have

$$\mathbf{E} e^{it\mathbb{T}} = \mathbf{E} e^{it(\tilde{\mathbb{T}} + \tilde{\Lambda}_1)} + R, \qquad |R| \le |t| \mathbf{E} |\tilde{\Lambda}_2|,$$

By Lyapunov's inequality,

$$\mathbf{E} |\tilde{\Lambda}_2| \le (\mathbf{E} \Lambda_2^2)^{1/2} + (\mathbf{E} \Lambda_3^2)^{1/2} + (\mathbf{E} \Lambda_5^2)^{1/2}.$$

Invoking the moment bounds of Lemma 5 we obtain $|t| \mathbf{E} |\tilde{\Lambda}_2| \prec \mathcal{R}$, thus, proving the second part of (129).

In order to prove the first part we combine Taylor's expansion with bounds for characteristic functions. Expanding the exponent we obtain

$$\mathbf{E} e^{it(\tilde{\mathbb{T}} + \tilde{\Lambda}_1)} = \mathbf{E} e^{it\tilde{\mathbb{T}}} + it \mathbf{E} e^{it\tilde{\mathbb{T}}} \tilde{\Lambda}_1 + R, \qquad |R| \le t^2 \mathbf{E} |\tilde{\Lambda}_1|^2.$$

Invoking the identities

$$\mathbf{E} \Lambda_1^2 = \binom{m}{2} \frac{\gamma_2}{N^3}, \qquad \mathbf{E} \Lambda_4^2 = m \binom{N-m}{2} \frac{\zeta_2}{N^5} \tag{130}$$

we obtain, for $\gamma_2 < c_*$ and $\zeta_2 < c_*$, see (5), and $m \le N^{1/12}$, that $R \prec \mathcal{R}$. We complete the proof of (129) by showing that

$$t\mathbf{E}e^{it\tilde{\mathbb{T}}}\tilde{\Lambda}_1 \prec \mathcal{R}. \tag{131}$$

Let us prove (131). Split $\mathbb{W} = \mathbb{W}_1 + \mathbb{W}_2 + \mathbb{W}_3 + R_W$, where

$$\mathbb{W}_k = \sum_{A\subset\Omega',\,|A|=k} T_A, \qquad R_W = \sum_{A\subset\Omega',\,|A|\ge 4} T_A.$$

Here $\Omega' = \{m+1, \ldots, N\}$. Denote $\mathbb{R} = \mathbb{U}_2^* + \mathbb{W}_3 + R_W$ and $\mathbb{U}_1 = \sum_{j=1}^N g_1(X_j)$. We have $\tilde{\mathbb{T}} = \mathbb{U}_1 + \mathbb{W}_2 + \mathbb{R}$. Expanding the exponent in powers of $it\mathbb{R}$ we obtain

$$t\mathbf{E}e^{it\tilde{\mathbb{T}}}\tilde{\Lambda}_1 = t\mathbf{E}e^{it(\mathbb{U}_1+\mathbb{W}_2)}\tilde{\Lambda}_1 + t^2 R, \tag{132}$$

where

$$|R| \le \mathbf{E}|\tilde{\Lambda}_1\mathbb{R}| \le (r_1 + r_2)(r_3 + r_4 + r_5),$$
$$r_1^2 = \mathbf{E}\Lambda_1^2, \quad r_2^2 = \mathbf{E}\Lambda_4^2, \quad r_3^2 = \mathbf{E}(\mathbb{U}_2^*)^2, \quad r_4^2 = \mathbf{E}R_W^2, \quad r_5^2 = \mathbf{E}\mathbb{W}_3^2.$$

In the last step we applied the Cauchy–Schwartz inequality. Combining (130) with the identities

$$\mathbf{E}(\mathbb{U}_2^*)^2 = \frac{m(N-m)}{N^3}\gamma_2, \qquad \mathbf{E}\mathbb{W}_3^2 = \frac{\binom{N-m}{3}}{N^5}\zeta_2$$

and invoking the simple bound

$$\mathbf{E}R_W^2 \le \frac{\Delta_4^2}{N^3} \le \frac{D_*}{N^{2+2\nu_1}},$$

we obtain $t^2(r_1 + r_2)(r_3 + r_4 + r_5) \prec \mathcal{R}$. Therefore, (132) implies

$$t\mathbf{E}e^{it\tilde{\mathbb{T}}}\tilde{\Lambda}_1 \sim t\mathbf{E}e^{it(\mathbb{U}_1+\mathbb{W}_2)}\tilde{\Lambda}_1.$$

Let us show that $t\mathbf{E}e^{it(\mathbb{U}_1+\mathbb{W}_2)}\tilde{\Lambda}_1 \sim 0$. Expanding the exponent in powers of $it\mathbb{W}_2$ we get

$$t\mathbf{E}e^{it(\mathbb{U}_1+\mathbb{W}_2)}\tilde{\Lambda}_1 = f_1(t) + f_2(t) + f_3(t) + f_4(t),$$
$$f_1(t) = t\mathbf{E}e^{it\mathbb{U}_1}\tilde{\Lambda}_1, \qquad f_2(t) = it^2\mathbf{E}e^{it\mathbb{U}_1}\Lambda_1\mathbb{W}_2,$$
$$f_3(t) = t^2\mathbf{E}e^{it\mathbb{U}_1}\Lambda_4\mathbb{W}_2\theta_1, \qquad f_4(t) = t^3\mathbf{E}e^{it\mathbb{U}_1}\Lambda_1\mathbb{W}_2^2\theta_2/2,$$

where θ_1, θ_2 are functions of $\mathbb{W}_2$ satisfying $|\theta_i| \le 1$.

Let us show that $f_i \prec \mathcal{R}$, for $i = 1, 2, 3, 4$. Split the set $\Omega_m = \{1, \dots, m\}$ in three (non-intersecting) parts $A_1 \cup A_2 \cup A_3 = \Omega_m$ of (almost) equal size $|A_i| \approx m/3$. The set of pairs $\big\{\{i, j\} \subset \Omega_m\big\}$ splits into six (non-intersecting) parts B_{kr}, $1 \le k \le r \le 3$ (the pair $\{i, j\}$ belongs to B_{kr} if $i \in A_k$ and $j \in A_r$). Write

$$\Lambda_1 = \sum_{1\le k\le r\le 3} \Lambda_1(k, r), \qquad \Lambda_1(k, r) = \sum_{\{i,j\}\in B_{kr}} g_2(X_k, X_l),$$
$$\Lambda_4 = \sum_{1\le k\le 3} \Lambda_4(k), \qquad \Lambda_4(k) = \sum_{i\in A_k} \sum_{m+1\le j<l\le N} g_3(X_i, X_j, X_l).$$

Let us prove $f_4 \prec \mathcal{R}$. We shall show that

$$t^3 \mathbf{E} e^{it\mathbb{U}_1} \Lambda_1(k, r) \mathbb{W}_2^2 \theta_2 \prec \mathcal{R}. \tag{133}$$

Given a pair (k, r) denote $A_i = \Omega_m \setminus (A_k \cup A_r)$ and write $k_i = |A_i|$. Note that $k_i \approx m/3$. We shall assume that $k_i \ge m/4$. Since the random variable $\mathbb{U}_1(A_i) := \sum_{j\in A_i} g_1(X_j)$ and the random variables $\Lambda_1(k, r)$, $\mathbb{W}_2$ are independent, we have

$$\mathbf{E} e^{it\mathbb{U}_1} \Lambda_1(k, r) \mathbb{W}_2^2 \theta_2 = \mathbf{E} e^{it\mathbb{U}_1(A_i)}\, \mathbf{E} \Lambda_1(k, r) \mathbb{W}_2^2 \theta_2.$$

Therefore,

$$|\mathbf{E} e^{it\mathbb{U}_1} \Lambda_1(k, r) \mathbb{W}_2^2 \theta_2| \le |\mathbf{E} e^{it\mathbb{U}_1(A_i)}|\, \mathbf{E} |\Lambda_1(k, r) \mathbb{W}_2^2|. \tag{134}$$

The first factor on the right is bounded from above by $\exp\{-mt^2/16N\}$, for $k_i \ge m/4$, see (165) below. The second factor is bounded from above by r, where

$$r^2 = \mathbf{E} \Lambda_1^2(k, r) \mathbf{E} \mathbb{W}_2^4 \le c_* m^2 N^{-5}.$$

Here we combined the Cauchy–Schwartz inequality and the bounds

$$\mathbf{E} \Lambda_1^2(k, r) \le c_* m^2 N^{-3}, \qquad \mathbf{E} \mathbb{W}_2^4 \le c_* N^{-2}.$$

Finally, (133) follows from (134)

$$\left| t^3 \mathbf{E} e^{it\mathbb{U}_1} \Lambda_1(k, r) \mathbb{W}_2^2 \theta_2 \right| \le c_* |t|^3 e^{-mt^2/16N} m N^{-5/2} \prec \mathcal{R}.$$

The proof of $f_3 \prec \mathcal{R}$ is almost the same as that of $f_4 \prec \mathcal{R}$.

Let us prove $f_2 \prec \mathcal{R}$. Split the set $\Omega' = \{m+1, \dots, N\}$ into three (non-intersecting) parts $B_1 \cup B_2 \cup B_3 = \Omega'$ of (almost) equal sizes $|B_i| \approx (N - m)/3$. Split the set of pairs $\big\{\{i, j\} : m + 1 \le i < j \le N\big\}$ into (non-intersecting) groups $D(k, r)$, for $1 \le k \le r \le 3$. The pair $\{i, j\} \in D(k, r)$ if $i \in B_k$ and $j \in B_r$. Write

$$\mathbb{W}_2 = \sum_{1\le k\le r\le 3} \mathbb{W}_2(k,r), \qquad \mathbb{W}_2(k,r) = \sum_{\{i,j\}\in D(k,r)} g_2(X_i, X_j).$$
$$\Lambda_4 = \sum_{1\le k\le r\le 3} \Lambda_4(k,r), \qquad \Lambda_4(k,r) = \sum_{1\le s\le m} \sum_{\{i,j\}\in D(k,r)} g_3(X_s, X_i, X_j),$$

In order to prove $f_2 \prec \mathcal{R}$ we shall show that

$$t^2 \mathbf{E} e^{it\mathbb{U}_1} \Lambda_1 \mathbb{W}_2(k,r) \prec \mathcal{R}. \tag{135}$$

Write $B_i = \Omega' \setminus (B_k \cup B_r)$ and denote $m_i = |B_i|$. We shall assume that $m_i \ge N/4$. Since the random variable $\mathbb{U}_1(B_i) = \sum_{j\in B_i} g_1(X_j)$ and the random variables Λ_1 and $\mathbb{W}_2(k,r)$ are independent, we have, cf. (134),

$$|\mathbf{E} e^{it\mathbb{U}_1} \Lambda_1 \mathbb{W}_2(k,r)| \le |\mathbf{E} e^{it\mathbb{U}_1(B_i)}|\ \mathbf{E}|\Lambda_1 \mathbb{W}_2(k,r)|. \tag{136}$$

The first factor in the right is the product $|\alpha^{m_i}(t)| \le e^{-m_i t^2/4N}$, see the argument used in the proof of (133) above. The second factor is bounded from above by $\tilde{r}$, where

$$\tilde{r}^2 = \mathbf{E}\Lambda_1^2 \mathbf{E}\mathbb{W}_2^2(k,r) \le c_* m^2 N^{-4}.$$

Finally, we obtain, using the inequality $m_i \ge N/4$,

$$|\mathbf{E} e^{it\mathbb{U}_1}|\ \mathbf{E}|\Lambda_1 \mathbb{W}_2(k,r)| \le c_* \frac{m}{N^2} \exp\{-t^2 \frac{m_i}{4N}\} \le c_* \frac{m}{N^2} \exp\{-\frac{t^2}{16}\}.$$

This in combination with (136) shows (135). We obtain $f_2 \prec \mathcal{R}$.

Let us prove $f_1 \prec \mathcal{R}$. We shall show that $f^* \prec \mathcal{R}$ and $f^\star \prec \mathcal{R}$, where

$$f^\star = t\mathbf{E} e^{it\mathbb{U}_1} \Lambda_1 \qquad \text{and} \qquad f^* = t\mathbf{E} e^{it\mathbb{U}_1} \Lambda_4$$

satisfy $f^* + f^\star = f_1$.

Let us show $f^\star \prec \mathcal{R}$. Denote $\mathbb{U}_1^\star = \sum_{j=m+1}^N g_1(X_j)$. We obtain, by the independence of $\mathbb{U}_1^\star$ and Λ_1 that

$$|\mathbf{E} e^{it\mathbb{U}_1} \Lambda_1| \le |\mathbf{E} e^{it\mathbb{U}_1^\star}|\ \mathbf{E}|\Lambda_1|.$$

Invoking, for $N - m > N/2$, the bound $|\mathbf{E} e^{it\mathbb{U}_1^\star}| \le e^{-t^2/8}$, see (165) below, and the bound $\mathbf{E}|\Lambda_1| \le (\mathbf{E}\Lambda_1^2)^{1/2} \le c_* m N^{-3/2}$ we obtain

$$|f^\star(t)| \le c_* |t| e^{-t^2/8} N^{-3/2} \prec \mathcal{R}.$$

Let us prove $f^* \prec \mathcal{R}$. We shall show that, for $1 \le k \le r \le 3$,

$$t\mathbf{E} e^{it\mathbb{U}_1} \Lambda_4(k,r) \prec \mathcal{R}. \tag{137}$$

Proceeding as in the proof of (135) we obtain the chain of inequalities

$$|\mathbf{E}e^{it\mathbb{U}_1}\Lambda_4(k,r)| \le e^{-t^2/16}\mathbf{E}|\Lambda_4(k,r)| \le c_* e^{-t^2/16} m^{1/2} N^{-3/2}. \quad (138)$$

In the last step we applied Cauchy–Schwartz and the simple bound $\mathbf{E}\Lambda_4^2(k,r) \le c_* m N^{-3}$. Clearly, (138) implies (137).

5.2 Proof of the second relation of (128)

Here we prove the second relation of (128). Firstly, we shall show that

$$\mathbf{E}e^{it\mathbb{T}} \sim \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2 + \mathbb{U}_3)\}, \quad (139)$$

$$\mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2 + \mathbb{U}_3)\} \sim \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2)\} + \binom{N}{3} e^{-t^2/2}(it)^4 w, \quad (140)$$

where $w = \mathbf{E}g_3(X_1, X_2, X_3)g_1(X_1)g_1(X_2)g_1(X_3)$.

Let $m(t)$ be an integer valued function such that

$$m(t) \approx C_1 N t^{-2} \ln(t^2+1), \qquad C_1 \le |t| \le t_1, \quad (141)$$

and put $m(t) \equiv 10$, for $|t| \le C_1$. Here C_1 denotes a large absolute constant (one can take, e.g., $C_1 = 200$). Assume, in addition, that the numbers $m = m(t)$ are even.

5.2.1 Proof of (139)

Given m write

$$\mathbb{T} = \mathbb{U}_1 + \mathbb{U}_2 + \mathbb{U}_3 + \mathbb{H},$$

where

$$\mathbb{H} = \mathbb{H}_1 + \mathbb{H}_2, \qquad \mathbb{H}_1 = \sum_{|A|\ge 4,\ A\cap\Omega_m = \emptyset} T_A, \qquad \mathbb{H}_2 = \sum_{|A|\ge 4,\ A\cap\Omega_m \ne \emptyset} T_A.$$

In order to show (139) we expand the exponent in powers of $it\mathbb{H}$ and $it\mathbb{U}_3$,

$$\mathbf{E}\exp\{it\mathbb{T}\} = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2 + \mathbb{U}_3)\} + \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2)\}it\mathbb{H} + R,$$

where $|R| \le t^2(\mathbf{E}\mathbb{H}^2 + \mathbf{E}|\mathbb{U}_3\mathbb{H}|)$. Invoking the bounds, see (166), (167), (5), (6),

$$\mathbf{E}\mathbb{H}^2 \le N^{-3}\Delta_4^2 \le c_* N^{-2-2\nu_1}, \qquad \mathbf{E}\mathbb{U}_3^2 \le N^{-2}\zeta_2 \le c_* N^{-2} \quad (142)$$

we obtain, by Cauchy–Schwartz, $|R| \le c_* t^2 N^{-2-\nu_1} \prec \mathcal{R}$. We complete the proof of (139) by showing that

$$\mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2)\}it\mathbb{H} \prec \mathcal{R}. \quad (143)$$

Before proving (143) we collect some auxiliary inequalities. For $m = 2k$ write

$$\Omega_m = A_1 \cup A_2, \text{ where } \quad A_1 = \{1, \dots, k\}, \quad A_2 = \{k+1, \dots, 2k\}. \tag{144}$$

Furthermore, split the sum

$$\begin{aligned} &\mathbb{U}_2 = \mathbb{Z}_1 + \mathbb{Z}_2 + \mathbb{Z}_3 + \mathbb{Z}_4, \\ &\mathbb{Z}_1 = \sum_{1 \le i < j \le m} g_2(X_i, X_j), \qquad \mathbb{Z}_2 = \sum_{i \in A_1} \sum_{m < j \le N} g_2(X_i, X_j), \\ &\mathbb{Z}_3 = \sum_{i \in A_2} \sum_{m < j \le N} g_2(X_i, X_j), \qquad \mathbb{Z}_4 = \sum_{m < i < j \le N} g_2(X_i, X_j). \end{aligned} \tag{145}$$

In what follows we shall use the simple bounds, see (5),

$$\begin{aligned} &\mathbf{E}\mathbb{Z}_1^2 \le \frac{m^2}{N^3}\gamma_2 \le c_* \frac{m^2}{N^3}, \qquad \mathbf{E}\mathbb{Z}_4^2 \le \frac{\gamma_2}{N} \le \frac{c_*}{N}, \\ &\mathbf{E}\mathbb{Z}_i^2 \le \frac{m}{N^2}\gamma_2 \le c_* \frac{m}{N^2}, \quad \mathbf{E}\mathbb{Z}_i^4 \le c\frac{m^2}{N^4}\gamma_4 \le c_* \frac{m^2}{N^4}, \quad i = 2, 3. \end{aligned} \tag{146}$$

Let us prove (143). Expand the exponent $\exp\{it(\mathbb{U}_1 + \mathbb{Z}_1 + \cdots + \mathbb{Z}_4)\}$ in powers of $it\mathbb{Z}_1$ to get

$$\mathbf{E} \exp\{it(\mathbb{U}_1 + \mathbb{U}_2)\} it\mathbb{H} = h_1(t) + R,$$

where $h_1(t) = \mathbf{E} \exp\{it(\mathbb{U}_1 + \mathbb{Z}_2 + \cdots + \mathbb{Z}_4)\} it\mathbb{H}$ and where

$$|R| \le t^2 \mathbf{E}|\mathbb{H}\mathbb{Z}_1| \le t^2 (\mathbf{E}\mathbb{H}^2)^{1/2} (\mathbf{E}\mathbb{Z}_1^2)^{1/2} \le c_* t^2 m N^{-(5+2\nu_1)/2}.$$

For $m = m(t)$ satisfying (141) we have $R \prec \mathcal{R}$. Therefore, we obtain

$$\mathbf{E} \exp\{it(\mathbb{U}_1 + \mathbb{U}_2)\} it\mathbb{H} \sim h_1.$$

In order to prove $h_1 \prec \mathcal{R}$ we write $h_1 = h_2 + h_3$ and show that $h_2, h_3 \prec \mathcal{R}$, where

$$\begin{aligned} h_2 &= \mathbf{E} \exp\{it(\mathbb{U}_1 + \mathbb{Z}_2 + \cdots + \mathbb{Z}_4)\} it\mathbb{H}_1, \\ h_3 &= \mathbf{E} \exp\{it(\mathbb{U}_1 + \mathbb{Z}_2 + \cdots + \mathbb{Z}_4)\} it\mathbb{H}_2. \end{aligned}$$

Let us show that $h_2 \prec \mathcal{R}$. Firstly, we prove that

$$h_2 \sim h_{2.1} + h_{2.2} + h_{2.3}, \tag{147}$$

where $h_{2.1}(t) = \mathbf{E} \exp\{it(\mathbb{U}_1 + \mathbb{Z}_4)\} it\mathbb{H}_1$ and, for $j = 2, 3$,

$$h_{2.j}(t) = \mathbf{E} \exp\{it(\mathbb{U}_1 + \mathbb{Z}_4)\} (it)^2 \mathbb{H}_1 \mathbb{Z}_j.$$

Expanding the exponent in powers of $it(\mathbb{Z}_2+\mathbb{Z}_3)$ we obtain

$$h_2 = h_{2.1} + h_{2.2} + h_{2.3} + R,$$

where $|R| \le |t|^3 \mathbf{E}|\mathbb{H}_1|(\mathbb{Z}_2+\mathbb{Z}_3)^2$ is bounded from above by

$$|t|^3(\mathbf{E}\mathbb{H}_1^2)^{1/2}(\mathbf{E}(\mathbb{Z}_2+\mathbb{Z}_3)^4)^{1/2} \le c_*|t|^3 m N^{-3-\nu_1} \prec \mathcal{R}.$$

In the last step we used $\mathbf{E}\mathbb{H}_1^2 \le \mathbf{E}\mathbb{H}^2$ and applied (142) and (146). Therefore, (147) follows.

Let us show $h_{2.i} \prec \mathcal{R}$, for $i = 1, 2, 3$. The random variable $\mathbb{U}_1(A_1)$ does not depend on the observations X_j, $j \in \Omega \setminus A_1$. Therefore, we can write

$$h_{2.3} = \mathbf{E}\exp\{it\mathbb{U}_1(A_1)\}\mathbf{E}\exp\{it(\mathbb{U}_1(\Omega\setminus A_1)+\mathbb{Z}_4)\}(it)^2\mathbb{H}_1\mathbb{Z}_3.$$

Furthermore, using (165) we obtain, for $|A_1| = m/2$,

$$|h_{2.3}| \le t^2|\alpha^{m/2}(t)|\mathbf{E}|\mathbb{H}_1\mathbb{Z}_3| \le c_* t^2 \exp\{-t^2\frac{m}{8N}\}\frac{m^{1/2}}{N^{2+\nu_1}}. \qquad (148)$$

In the last step we combined the bound $\mathbf{E}\mathbb{H}_1^2 \le c_* N^{-2-2\nu_1}$ and (146) to get

$$\mathbf{E}|\mathbb{H}_1\mathbb{Z}_3| \le (\mathbf{E}\mathbb{H}_1^2)^{1/2}(\mathbf{E}\mathbb{Z}_3^2)^{1/2} \le c_* m^{1/2} N^{-2-\nu_1}.$$

Note that choosing of C_1 in (141) sufficiently large implies, for $|t| \ge C_1$,

$$t^2 m/12N \approx (C_1/12)\ln(t^2+1) \ge 10\ln(t^2+1).$$

An application of this bound to the argument of the exponent in (148) shows $h_{2.3} \prec \mathcal{R}$. The proof of $h_{2.i} \prec \mathcal{R}$, for $i = 1, 2$, is almost the same. Therefore, we obtain $h_2 \prec \mathcal{R}$.

Let us prove $h_3 \prec \mathcal{R}$. Firstly we collect some auxiliary inequalities. Write $m = 2k$ (recall that the number m is even) and split $\Omega_m = B \cup D$, where B denotes the set of odd numbers and D denotes the set of even numbers. Split $\mathbb{H}_2 = \mathbb{H}_B + \mathbb{H}_D + \mathbb{H}_C$. Here, for $A \subset \Omega_N$ and $|A| \ge 4$, we denote by $\mathbb{H}_B$ the sum of T_A such that $A \cap B = \emptyset$ and $A \cap D \ne \emptyset$; $\mathbb{H}_D$ denotes the sum of T_A such that $A \cap B \ne \emptyset$ and $A \cap D = \emptyset$; $\mathbb{H}_C$ denotes the sum of T_A such that $A \cap B \ne \emptyset$ and $A \cap D \ne \emptyset$. It follows from the inequalities (177) and (6) that

$$\mathbf{E}\,\mathbb{H}_C^2 \le c_* m^2 N^{-4-2\nu_1}, \qquad \mathbf{E}\,\mathbb{H}_B^2 = \mathbf{E}\,\mathbb{H}_D^2 \le c_* m N^{-3-2\nu_1}. \qquad (149)$$

Using the notation $z = it\exp\{it(\mathbb{U}_1+\mathbb{Z}_2+\mathbb{Z}_3+\mathbb{Z}_4)\}$ write

$$h_3 = \mathbf{E}z\mathbb{H}_2 = h_{3.1} + h_{3.2} + h_{3.3},$$
$$h_{3.1} = \mathbf{E}z\mathbb{H}_B, \quad h_{3.2} = \mathbf{E}z\mathbb{H}_D, \quad h_{3.3} = \mathbf{E}z\mathbb{H}_C.$$

We shall show that $h_{3.i} \prec \mathcal{R}$, for $i = 1, 2, 3$. The relation $h_{3.3} \prec \mathcal{R}$ follows from (149) and (146), and by Cauchy–Schwartz, $|h_{3.3}| \le c_* |t|\, m N^{-2-\nu_1} \prec \mathcal{R}$.

Let us show that $h_{3.2} \prec \mathcal{R}$. Expanding the exponent in powers of $it(\mathbb{Z}_2 + \mathbb{Z}_3)$ we obtain

$$h_{3.2} = h^*_{3.2} + R, \qquad h^*_{3.2} := \mathbf{E} \exp\{it(\mathbb{U}_1 + \mathbb{Z}_4)\} it \mathbb{H}_D,$$

where $|R| \le t^2 \mathbf{E}|\mathbb{H}_D(\mathbb{Z}_2 + \mathbb{Z}_3)|$. Combining the bounds (146) and (149) we obtain, by Cauchy–Schwartz, $|R| \le c_* t^2 m N^{-(5+2\nu_1)/2} \prec \mathcal{R}$. Next we show that $h^*_{3.2} \prec \mathcal{R}$. The random variable $\mathbb{U}_1(D) = \sum_{j \in D} g_1(X_j)$ and the random variable $\mathbb{H}_D$ are independent. Therefore, we can write

$$|h^*_{3.2}| \le |t|\, |\mathbf{E} \exp\{it\mathbb{U}_1(D)\}| \mathbf{E}|\mathbb{H}_D|.$$

Combining (165) and (149) we obtain using Cauchy–Schwartz,

$$|h^*_{3.2}| \le c_* |t|\, e^{-mt^2/8N} m^{1/2} N^{-(3+2\nu_1)/2} \prec \mathcal{R}.$$

The proof of $h_{3.1} \prec \mathcal{R}$ is similar. Therefore, we obtain $h_3 \prec \mathcal{R}$. This together with the relation $h_2 \prec \mathcal{R}$, proved above, implies $h_1 \prec \mathcal{R}$. Thus we arrive at (143) completing the proof of (139).

5.2.2 Proof of (140)

We start with some auxiliary moment inequalities. Split

$$\mathbb{U}_3 = W + Z, \qquad W = \sum_{|A|=3,\ A \cap \Omega_m \ne \emptyset} T_A, \qquad Z = \sum_{|A|=3,\ A \cap \Omega_m = \emptyset} T_A.$$

Using the orthogonality and moment bounds for U-statistics, see, e.g., Dharmadhikari et al. [20], one can show that

$$\mathbf{E} W^2 \le m N^2 \mathbf{E} g_3^2(X_1, X_2, X_3), \qquad \mathbf{E} Z^2 \le N^3 \mathbf{E} g_3^2(X_1, X_2, X_3),$$

and $\mathbf{E}|Z|^s \le c N^{3s/2} \mathbf{E}|g_3(X_1, X_2, X_3)|^s$. Invoking (5) we obtain

$$\mathbf{E} W^2 \le c_* m N^{-3}, \qquad \mathbf{E} Z^2 \le c_* N^{-2}, \qquad \mathbf{E}|Z|^s \le c_* N^{-s}. \tag{150}$$

For the sets $A_1, A_2 \subset \Omega_m$ defined in (144) write

$$\begin{aligned}
\mathcal{D} &= \{A \subset \Omega_N : |A| = 3,\ A \cap \Omega_m \ne \emptyset\},\\
\mathcal{D}_1 &= \{A \in \mathcal{D} :\ A \cap A_1 = \emptyset\},\\
\mathcal{D}_2 &= \{A \in \mathcal{D} :\ A \cap A_2 = \emptyset\},\\
\mathcal{D}_3 &= \{A \in \mathcal{D} :\ A \cap A_1 \ne \emptyset,\ A \cap A_2 \ne \emptyset\}.
\end{aligned}$$

We have $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3$ and $W = \sum_{A\in\mathcal{D}} T_A$. Therefore, we can write $W = W_1 + W_2 + W_3$, where $W_j = \sum_{A\in\mathcal{D}_j} T_A$.

A calculation shows that

$$\mathbf{E}W_1^2 = \mathbf{E}W_2^2 \le kN^2\mathbf{E}g_3^2(X_1, X_2, X_3), \qquad \mathbf{E}W_3^2 \le k^2N\mathbf{E}g_3^2(X_1, X_2, X_3).$$

Therefore, we obtain form (5) that

$$\mathbf{E}W_1^2 = \mathbf{E}W_2^2 \le c_* m N^{-3}, \qquad \mathbf{E}W_3^2 \le c_* m^2 N^{-4}. \tag{151}$$

Let us prove (140). Write $\mathbb{U}_3 = W + Z$. Expanding the exponent in powers of itW we obtain

$$\begin{aligned}&\mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2 + \mathbb{U}_3)\} = h_4 + h_5 + R,\\ &h_4 = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2 + Z)\},\\ &h_5 = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2 + Z)\}itW,\end{aligned}$$

where, by (150), $|R| \le t^2\mathbf{E}W^2 \le c_* t^2 m N^{-3} \prec \mathcal{R}$. This implies

$$\mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2 + \mathbb{U}_3)\} \sim h_4 + h_5.$$

In order to prove (140) we shall show that

$$h_5 \sim \mathbf{E}\exp\{it\mathbb{U}_1\}itW, \tag{152}$$

$$h_4 \sim \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2)\} + \mathbf{E}\exp\{it\mathbb{U}_1\}itZ, \tag{153}$$

$$\mathbf{E}\exp\{it\mathbb{U}_1\}it\mathbb{U}_3 \sim \binom{N}{3} e^{-t^2/2}(it)^4 w. \tag{154}$$

Let us prove (152). Expanding the exponent (in h_5) in powers of itZ we obtain

$$h_5 = h_6 + R, \qquad h_6 = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2)\}itW,$$

where, by (150) and Cauchy–Schwartz,

$$|R| \le t^2\mathbf{E}|WZ| \le c_* t^2 m^{1/2} N^{-5/2} \prec \mathcal{R}.$$

We have, $h_5 \sim h_6$.

It remains to show that $h_6 \sim \mathbf{E}\exp\{it\mathbb{U}_1\}itW$. Split

$$\mathbb{U}_2 = \mathbb{U}_2^* + \mathbb{U}_2^\star, \qquad \mathbb{U}_2^* = \sum_{|A|=2,\ A\cap\Omega_m\neq\emptyset} T_A, \qquad \mathbb{U}_2^\star = \sum_{|A|=2,\ A\cap\Omega_m=\emptyset} T_A. \tag{155}$$

We have, see (146),

$$\mathbf{E}(\mathbb{U}_2^*)^2 \le c_* m N^{-2}, \qquad \mathbf{E}(\mathbb{U}_2^\star)^2 \le c_* N^{-1}. \tag{156}$$

Expanding the exponent (in h_6) in powers of $it\mathbb{U}_2^*$ we obtain

$$h_6 = h_7 + R, \qquad \text{where} \qquad h_7 = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2^\star)\}itW,$$

and where, by (150), (156) and Cauchy–Schwartz,

$$|R| \le t^2\mathbf{E}|W\mathbb{U}_2^*| \le c_* t^2 m N^{-5/2} \prec \mathcal{R}.$$

Therefore, we obtain $h_6 \sim h_7$.

We complete the proof of (152) by showing that $h_7 \sim \mathbf{E}\exp\{it\mathbb{U}_1\}itW$. Use the decomposition $W = W_1 + W_2 + W_3$ and write

$$h_7 = h_{7.1} + h_{7.2} + h_{7.3}, \qquad h_{7.j} = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2^\star)\}itW_j.$$

We shall show that

$$h_{7.j} \sim \mathbf{E}\exp\{it\mathbb{U}_1\}itW_j, \qquad j = 1, 2, 3. \tag{157}$$

Expanding in powers of $it\mathbb{U}_2^\star$ we obtain

$$h_{7.j} = \mathbf{E}\exp\{it\mathbb{U}_1\}itW_j + R_j,$$

where $R_j = (it)^2\mathbf{E}\exp\{it\mathbb{U}_1\}W_j\mathbb{U}_2^\star\theta$ and where θ is a function of $\mathbb{U}_2^\star$ satisfying $|\theta| \le 1$. In order to prove (157) we show that $R_j \prec \mathcal{R}$, for $j = 1, 2, 3$.

Combining (151) and (156) we obtain via Cauchy–Schwartz

$$|R_3| \le c_* t^2 m N^{-5/2} \prec \mathcal{R}.$$

Furthermore, using the fact that the random variable $\mathbb{U}_1(A_2)$ and the random variables $\mathbb{U}_2^\star$ and W_2 are independent, we can write

$$|R_2| \le t^2|\mathbf{E}\exp\{it\mathbb{U}_1(A_2)\}|\mathbf{E}|W_2\mathbb{U}_2^\star| \le c_* t^2 e^{-mt^2/8N} m^{1/2} N^{-2} \prec \mathcal{R}.$$

Here we used (165) and the moment inequalities (151) and (156). The proof of $R_1 \prec \mathcal{R}$ is similar. We arrive at (157) and, thus, complete the proof of (152).

Let us prove (153). We proceed in two steps. Firstly we show

$$\begin{aligned} &h_4 \sim h_8 + h_9, \\ &h_8 = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2)\}, \qquad h_9 = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2)\}itZ. \end{aligned} \tag{158}$$

Secondly, we show

$$h_9 \sim \mathbf{E}\exp\{it\mathbb{U}_1\}itZ. \tag{159}$$

In order to prove (158) we write

$$h_4 = h_8 + h_9 + R, \qquad R = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2)\}\tilde{r}, \qquad \tilde{r} = \exp\{itZ\} - 1 - itZ,$$

and show that $R \prec \mathcal{R}$. In order to bound the remainder R we write $\mathbb{U}_2 = \mathbb{U}_2^* + \mathbb{U}_2^\star$, see (155), and expand the exponent in powers of $it\mathbb{U}_2^*$. We obtain $R = R_1 + R_2$, where

$$R_1 = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2^\star)\}\tilde{r} \quad \text{and} \quad |R_2| \le \mathbf{E}|it\mathbb{U}_2^*\tilde{r}|.$$

Note that, for $2 < s \le 3$, we have $|\tilde{r}| \le c|tZ|^{s/2}$. Combining (150) and (156) we obtain via Cauchy–Schwartz,

$$|R_2| \le |t|^{1+s/2}\mathbf{E}|Z|^{s/2}|\mathbb{U}_2^*| \le c_*|t|^{1+s/2}m^{1/2}N^{-1-s/2} \prec \mathcal{R}.$$

In order to prove $R_1 \prec \mathcal{R}$ we use the fact that the random variable $\mathbb{U}_1(\Omega_m)$ and the random variables $\mathbb{U}_2^\star$ and $\tilde{r}$ are independent. Invoking the inequality $|\tilde{r}| \le t^2Z^2$ we obtain from (165) and (150)

$$|R_1| \le t^2|\alpha^m(t)|\mathbf{E}Z^2 \le c_*t^2e^{-mt^2/4N}N^{-2} \prec \mathcal{R}.$$

We thus arrive at (158).

Let us prove (159). Use the decomposition (145) and expand the exponent (in h_9) in powers of $it\mathbb{Z}_1$ to get $h_9 = h_{10} + R$, where

$$h_{10} = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{Z}_2 + \mathbb{Z}_3 + \mathbb{Z}_4)\}itZ, \quad |R| \le t^2\mathbf{E}|Z\mathbb{Z}_1|.$$

Combining (146) and (150) we obtain via Cauchy–Schwartz

$$|R| \le c_*t^2mN^{-5/2} \prec \mathcal{R}.$$

Therefore, we have

$$h_9 \sim h_{10}.$$

Now we expand the exponent in h_{10} in powers of $it(\mathbb{Z}_2 + \mathbb{Z}_3)$ and obtain $h_{10} = h_{11} + h_{12} + R$, where

$$h_{11} = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{Z}_4)\}itZ, \qquad h_{12} = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{Z}_4)\}(it)^2Z(\mathbb{Z}_2 + \mathbb{Z}_3),$$

and where $|R| \le |t|^3\mathbf{E}|Z|\,|\mathbb{Z}_2 + \mathbb{Z}_3|^2$. Combining (146) and (150) we obtain via Cauchy–Schwartz $|R| \le |t|^3mN^{-3} \prec \mathcal{R}$. Therefore, we have

$$h_{10} \sim h_{11} + h_{12}.$$

We complete the proof of (159) by showing that

$$h_{11} \sim \mathbf{E}\exp\{it\mathbb{U}_1\}itZ \quad \text{and} \quad h_{12} \prec \mathcal{R}. \tag{160}$$

In order to prove the second bound write

$$h_{12} = R_2 + R_3, \qquad \text{where} \qquad R_j = \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{Z}_4)\}(it)^2 Z\mathbb{Z}_j.$$

We shall show that $R_3 \prec \mathcal{R}$. Using the fact that the random variable $\mathbb{U}_1(A_1)$ and the random variables Z, $\mathbb{Z}_3$ and $\mathbb{Z}_4$ are independent we obtain from (165)

$$|R_3| \le t^2|\alpha^{m/2}(t)|\mathbf{E}|Z\mathbb{Z}_3| \le t^2 e^{-mt^2/8} m^{1/2} N^{-2} \prec \mathcal{R}.$$

In the last step we combined (146), (150) and Cauchy–Schwartz. The proof of $R_1 \prec \mathcal{R}$ is similar.

In order to prove the first relation of (160) we expand the exponent in powers of $it\mathbb{Z}_4$ and obtain $h_{11} = \mathbf{E}\exp\{it\mathbb{U}_1\}itZ + R$. Furthermore, combining (165), (146) and (150) we obtain

$$|R| \le t^2|\alpha^m(t)|\mathbf{E}|Z\mathbb{Z}_4| \le c_* t^2 e^{-mt^2/4N} N^{-3/2} \prec \mathcal{R}.$$

Hence the first relation of (160). The proof of (153) is complete.

Let us prove (154). By symmetry and the independence,

$$\mathbf{E}e^{it\mathbb{U}_1} it\mathbb{U}_3 = \binom{N}{3} h_{13}\mathbf{E}e^{it\mathbb{U}_*}, \qquad h_{13} = \mathbf{E}e^{itx_1}e^{itx_2}e^{itx_3}itz. \tag{161}$$

Here we denote $z = g_3(X_1, X_2, X_3)$ and write,

$$\mathbb{U}_1 = x_1 + x_2 + x_3 + \mathbb{U}_*, \qquad \mathbb{U}_* = \sum_{4\le j\le N} g_1(X_j), \qquad x_j = g_1(X_j).$$

Furthermore, write

$$r_j = e^{itx_j} - 1 - itx_j, \qquad v_j = e^{itx_j} - 1.$$

In what follows we expand the exponents in powers of itx_j, $j = 1, 2, 3$ and use the fact that $\mathbf{E}\big(g_3(X_1, X_2, X_3)\big|X_1, X_2\big) = 0$ as well as the obvious symmetry. Thus, we have

$$\begin{aligned}
h_{13} &= h_{14} + R_1, \qquad h_{14} = \mathbf{E}e^{itx_2}e^{itx_3}(it)^2 zx_1, \qquad R_1 = \mathbf{E}e^{itx_2}e^{itx_3}itzr_1,\\
h_{14} &= h_{15} + R_2, \qquad h_{15} = \mathbf{E}e^{itx_3}(it)^3 zx_1x_2, \qquad R_2 = \mathbf{E}e^{itx_3}(it)^2 zx_1r_2\\
h_{15} &= h_{16} + R_3, \qquad h_{16} = \mathbf{E}(it)^4 zx_1x_2x_3, \qquad R_3 = \mathbf{E}(it)^3 zx_1x_2r_3.
\end{aligned}$$

Furthermore, we have

$$R_1 = \mathbf{E}itz_1r_1v_2v_3, \qquad R_2 = \mathbf{E}(it)^2 zx_1r_2v_3.$$

Invoking the bounds $|r_j| \le |tx_j|^2$ and $|v_j| \le |tx_j|$ we obtain

$$h_{13} = h_{16} + R, \tag{162}$$

where $|R| \le c|t|^5 \mathbf{E}|zx_1x_2|\, x_3^2$. The bound, $|R| \le c_*|t|^5 N^{-9/2}$ (which follows, by Cauchy–Schwartz) in combination with (161) and (162) implies

$$\mathbf{E}e^{it\mathbb{U}_1} it\mathbb{U}_3 \sim \binom{N}{3} \mathbf{E}e^{it\mathbb{U}_*}(it)^4 w. \tag{163}$$

Note that $\binom{N}{3}|w| \le c_* N^{-1}$. In order to show (154) we replace $\mathbf{E}e^{it\mathbb{U}_*}$ by $e^{-t^2/2}$. Therefore, (154) follows from (163) and the inequalities

$$\frac{(it)^4}{N}(\mathbf{E}e^{it\mathbb{U}_*} - e^{-t^2\sigma^2(N-3)/2N}) \prec \mathcal{R}, \quad \frac{(it)^4}{N}(e^{-t^2\sigma^2(N-3)/2N} - e^{-t^2/2}) \prec \mathcal{R}.$$

The second inequality is a direct consequence of (169). The proof of the first inequality is routine and here omitted. Thus the proof of (140) is complete.

5.2.3 Completion of the proof of (128)

Here we show that

$$\mathbf{E}\exp\{it\mathbb{U}_1 + \mathbb{U}_2)\} + \binom{N}{3} e^{-t^2/2}(it)^4 w \sim \hat{G}(t). \tag{164}$$

This relation in combination with (139) and (140) implies $\mathbf{E}e^{it\mathbb{T}} \sim \hat{G}(t)$.

Let $G_U(t)$ denote the two term Edgeworth expansion of the U- statistic $\mathbb{U}_1 + \mathbb{U}_2$. That is, $G_U(t)$ is defined by (2), but with κ_4 replaced by κ_4^*, where κ_4^* is obtained from κ_4 after removing the summand $4\mathbf{E}g(X_1)g(X_2)g(X_3)\chi(X_1, X_2, X_3)$. Furthermore, let $\hat{G}_U(t)$ denote the Fourier transform of $G_U(t)$. It easy to show that

$$\hat{G}(t) = \hat{G}_U(t) + \binom{N}{3} e^{-t^2/2}(it)^4 w.$$

Therefore, in order to prove (164) it suffices to show that $\hat{G}_U(t) \sim \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2)\}$. The bound

$$\int_{|t|\le t_1} |\hat{G}_U(t) - \mathbf{E}\exp\{it(\mathbb{U}_1 + \mathbb{U}_2)\}| \frac{dt}{|t|} \le \varepsilon_N N^{-1}$$

where $\varepsilon_N \downarrow 0$, was shown by Callaert, Janssen and Veraverbeke [16] and Bickel, Götze and van Zwet [11]. An inspection of their proofs shows that under the moment conditions (5) one can replace ε_n by $c_* N^{-\nu}$. This completes the proof of (127).

For the reader convenience we formulate in Lemma 3 a known result on upper bounds for characteristic functions.

Lemma 3 *Assume that (16) holds. There exists a constant c_* depending on D_*, M_*, r, s, ν_1 only such that, for $N > c_*$ and $|t| \le N^{1/2}/10^3\beta_3$ and $B \subset \Omega_N$, we have*

$$|\alpha(t)| \le 1 - t^2/4N, \qquad \mathbf{E}\exp\{it\mathbb{U}_1(B)\}| \le |\alpha(t)|^{|B|} \le e^{-|B|t^2/4N}. \tag{165}$$

Here $\alpha(t) = \mathbf{E}\exp\{itg_1(X_1)\}$ *and* $\mathbb{U}_1(B) = \sum_{j\in B} g_1(X_j)$.

Proof Let us prove the first inequality of (165). Expanding the exponent, see (188), we obtain

$$\begin{aligned}|\alpha(t)| &\le \left|1 - 2^{-1}t^2\mathbf{E}g_1^2(X_1)\right| + 6^{-1}|t|^3\mathbf{E}|g_1(X_1)|^3 \\ &= \left|1 - \sigma^2 t^2/2N\right| + \beta_3\sigma^3|t|^3/6N^{3/2}\end{aligned}$$

Invoking the inequality $1 - 10^{-3} \le \sigma^2 \le 1$ which follows from (169) for $N > c_*$, where c_* is sufficiently large, we obtain $|\alpha(t)| \le 1 - t^2/4N$, for $|t| \le N^{1/2}/10^3\beta_3$.

The second inequality of (165) follows from the first one via the inequality $1 + x \le e^x$, for $x \in R$. □

Funding Open Access funding enabled and organized by Projekt DEAL.

6 Appendix 1

In Lemma 4 below we compare the moments Δ_m^2 and $\mathbf{E}R_m^2$, where R_m is the remainder of expansion (14),

$$\mathbb{T} = \mathbf{E}\mathbb{T} + \mathbb{U}_1 + \cdots + \mathbb{U}_{m-1} + R_m, \qquad R_m := \mathbb{U}_m + \cdots + \mathbb{U}_N.$$

For $k = 1, \ldots, N$, write $\Omega_k = \{1, 2, \ldots, k\}$ and denote $\sigma_k^2 := \mathbf{E}g_k^2(X_1, \ldots, X_k) = \mathbf{E}T_{\Omega_k}^2$. It follows from (14), by the orthogonality property (15), that

$$\sigma_{\mathbb{T}}^2 = \sum_{k=1}^N \mathbf{E}\mathbb{U}_k^2, \qquad \mathbf{E}R_m^2 = \sum_{k=m}^N \mathbf{E}\mathbb{U}_k^2, \qquad \mathbf{E}\mathbb{U}_k^2 = \binom{N}{k}\sigma_k^2. \tag{166}$$

Lemma 4 *Assume that* $\mathbf{E}\mathbb{T}^2 < \infty$. *Then*

$$\mathbf{E}R_m^2 \le N^{-(m-1)}\Delta_m^2, \tag{167}$$

$$\Delta_m^2 \le N^{2m-1}\sigma_m^2 + N^{-1}\Delta_{m+1}^2, \tag{168}$$

Assume that (5) *and* (6) *hold, then there exists a constant* $c_* < \infty$ *depending on* D_*, M_*, r, s, ν_1 *such that*

$$0 \le 1 - \sigma^2\sigma_{\mathbb{T}}^{-2} \le c_* N^{-1}. \tag{169}$$

Remark. For $m = 3$, inequality (168) yields $\Delta_3^2 \le \zeta_2 + N^{-1}\Delta_4^2$.

Proof Let us prove (167). The identity

$$D_1 \cdots D_m \mathbb{T} = \sum_{A:\, \Omega_m \subset A \subset \Omega_N} T_A = \sum_{m \le k \le N} \mathbb{U}_{k|m},$$

where $\mathbb{U}_{k|m} = \sum_{|A|=k,\, A \supset \Omega_m} T_A$, implies

$$\mathbf{E}(D_1 \cdots D_m \mathbb{T})^2 = \sum_{m \le k \le N} \mathbf{E}\mathbb{U}_{k|m}^2, \qquad \mathbf{E}\mathbb{U}_{k|m}^2 = \sigma_k^2 \binom{N-m}{k-m}. \tag{170}$$

We have

$$\mathbf{E}(D_1 D_2 \cdots D_m \mathbb{T})^2 = \sum_{m \le k \le N} \sigma_k^2 \binom{N-m}{k-m} = \sum_{m \le k \le N} \sigma_k^2 \binom{N}{k} b_k, \tag{171}$$

where $b_k = [k]_m/[N]_m$ satisfies $b_k \ge b_m \ge m! N^{-m}$. Here we denote $[x]_m = x(x-1)\cdots(x-m+1)$. A comparison of (166) and (171) shows (167)

$$\mathbf{E}R_m^2 \le N^m \mathbf{E}(D_1 \cdots D_m \mathbb{T})^2 = N^{-(m-1)}\Delta_m^2.$$

Let us prove (168). We have

$$\begin{aligned}\mathbf{E}(D_1 \cdots D_m \mathbb{T})^2 &= \sigma_m^2 + \sum_{m < k \le N} \sigma_k^2 \binom{N-m}{k-m} \\ &= \sigma_m^2 + \sum_{m < k \le N} \sigma_k^2 \binom{N-m-1}{k-m-1} \tilde{b}_k,\end{aligned}$$

where $\tilde{b}_k = (N-m)/(k-m) \le N$. We obtain the inequality

$$\mathbf{E}(D_1 \cdots D_m \mathbb{T})^2 \le \sigma_m^2 + N\mathbf{E}(D_1 \cdots D_{m+1}\mathbb{T})^2$$

which implies (168).

Let us prove (169). From (166), (167) we have, for $\sigma^2 = N\sigma_1^2$,

$$0 \le 1 - \frac{\sigma^2}{\sigma_{\mathbb{T}}^2} \le \binom{N}{2}\frac{\sigma_2^2}{\sigma_{\mathbb{T}}^2} + \binom{N}{3}\frac{\sigma_3^2}{\sigma_{\mathbb{T}}^2} + \frac{1}{N^3}\frac{\Delta_4^2}{\sigma_{\mathbb{T}}^2}.$$

Invoking the bounds, which follow from (5),

$$N^3\sigma_2^2 = \mathbf{E}\psi^2(X_1, X_2) \le M_*^{2/r}\sigma_{\mathbb{T}}^2, \qquad N^5\sigma_3^2 = \mathbf{E}\chi^2(X_1, X_2, X_3) \le M_*^{2/s}\sigma_{\mathbb{T}}^2$$

and using (6) we obtain (169). □

In Lemma 5 below we establish moments bounds for various parts of Hoeffding decomposition defined in Sect. 2.

Lemma 5 *Assume that* $\sigma_{\mathbb{T}}^2 = 1$. *For* $3 \le m \le N$ *and* $s > 2$, *we have*

$$\mathbf{E}\Lambda_3^2 \le \frac{m^3}{N^5}\Delta_3^2, \quad \mathbf{E}\Lambda_2^2 \le \frac{m^2}{N^4}\Delta_3^2, \quad \mathbf{E}|\Lambda_1|^3 \le c\frac{m^3}{N^{9/2}}\gamma_3, \tag{172}$$

$$\mathbf{E}|\Lambda_4|^s \le c(s)\, m^{s/2}N^{-3s/2}\zeta_s, \quad \mathbf{E}\eta_i^2 \le N^{-4}\Delta_4^2, \quad \mathbf{E}\Lambda_5^2 \le mN^{-4}\Delta_4^2. \tag{173}$$

Here c *denotes an absolute constant and* $c(s)$ *denotes a constant which depends only on* s.

Proof The inequalities (172) are proved in [4].

Let us prove (173). Split $\Lambda_4 = z_1 + \cdots + z_m$, where

$$z_i = \sum_{|A|=3,\ A\cap\Omega_m = i} T_A.$$

Let $\mathbf{E}'$ denote the conditional expectation given $X_{m+1}, \ldots, X_N$. It follows from Rosenthal's inequality that almost surely

$$\mathbf{E}'|\Lambda_4|^s \le c(s)\sum_{i=1}^m \mathbf{E}'|z_i|^s + c(s)\Big(\sum_{i=1}^m \mathbf{E}'z_i^2\Big)^{s/2}.$$

Invoking Hölder's inequality we obtain, by symmetry,

$$\mathbf{E}|\Lambda_4|^s = \mathbf{E}\mathbf{E}'|\Lambda_4|^s \le c(s)m^{s/2}\mathbf{E}|z_1|^s. \tag{174}$$

Using well known martingale moment inequalities (and their applications to U statistics), see [20], one can show the bound $\mathbf{E}|z_1|^s \le c(s)N^{-3s/2}\zeta_s$. Invoking this bound in (174) we obtain the first bound of (173).

In order to prove the second bound of (173) write

$$\eta_i = \sum_{k=4}^{N-m+1} U_k^*, \qquad U_k^* = \sum_{|A|=k,\ A\cap\Omega_m=\{i\}} T_A.$$

A simple calculation shows $\mathbf{E}(U_k^*)^2 = \binom{N-m}{k-1}\sigma_k^2$. Therefore, by orthogonality,

$$\begin{aligned}\mathbf{E}\eta_i^2 &= \sum_{k=4}^{N-m+1} \binom{N-m}{k-1}\sigma_k^2 = \sum_{k=4}^{N-m+1} \binom{N-4}{k-4} b_k \sigma_k^2 \\ &\leq N^3 \mathbf{E}(D_1 \cdots D_4 \mathbb{T})^2. \end{aligned} \tag{175}$$

In the last step we invoke (170) and use the bound $b_k \leq N^3$, where $b_k = \binom{N-m}{k-1}\binom{N-4}{k-4}^{-1}$. Clearly, (175) implies $\mathbf{E}\eta_i^2 \leq N^{-4}\Delta_4^2$. Finally, using the fact that $\eta_1, \dots, \eta_m$ are uncorrelated we obtain

$$\Lambda_5^2 = \mathbf{E}\eta_1^2 + \cdots + \mathbf{E}\eta_m^2 \leq m\, N^{-4}\Delta_4^2$$

thus completing the proof. □

Before formulating next result we introduce some notation. Given m let $\mathcal{D}$ denote the class of subsets $A \subset \Omega_N$ satisfying $|A| \geq 4$ and $\Omega_m \cap A \neq \emptyset$. Introduce the random variable $\mathbb{H}(m) = \sum_{A \in \mathcal{D}} T_A$. Denote $x_i = 2i - 1$ and $y_i = 2i$. For even integer $m = 2k \leq N$ write

$$\Omega_m = A_k \cup B_k, \qquad A_k = \{x_1, \dots, x_k\}, \qquad B_k = \{y_1, \dots, y_k\}$$

and put $A_0 = B_0 = \emptyset$. Let $\mathcal{A}(k)$ (respectively $\mathcal{B}(k)$) denote the collection of those $A \in \mathcal{D}$ which satisfy $A \cap A_k = \emptyset$ (respectively $A \cap B_k = \emptyset$). Furthermore, let $\mathcal{C}(k)$ denote the collection of $A \in \mathcal{D}$ such that $A \cap A_k \neq \emptyset$ and $A \cap B_k \neq \emptyset$. Write

$$\mathbb{H}_A(k) = \sum_{A \in \mathcal{A}(k)} T_A, \qquad \mathbb{H}_B(k) = \sum_{A \in \mathcal{B}(k)} T_A, \qquad \mathbb{H}_C(k) = \sum_{A \in \mathcal{C}(k)} T_A.$$

Lemma 6 *There exists an absolute constant c such that,*

$$\mathbf{E}\mathbb{H}^2(m) \leq c\frac{m}{N^4}\Delta_4^2, \qquad \textit{for} \qquad m = 4, 5, \dots, N. \tag{176}$$

For even integer $m = 2k < N$ we have

$$\mathbf{E}\mathbb{H}_A^2(k) = \mathbf{E}\mathbb{H}_B^2(k) \leq c\frac{k}{N^4}\Delta_4^2, \qquad \mathbf{E}\mathbb{H}_C^2(k) \leq c\frac{k^2}{N^5}\Delta_4^2. \tag{177}$$

Proof Let us prove the first bound of (176). For $m = 4$ we have

$$\mathbb{H}(4) = H_1 + H_2 + H_3 + H_4, \qquad H_k = \sum_{|A| \geq 4,\ |A \cap \Omega_4| = k} T_A.$$

A calculation shows that, for $k = 1, 2, 3, 4$,

$$\mathbf{E}H_k^2 = \binom{4}{k} \sum_{j=4}^{N} \sigma_j^2 \binom{N-4}{j-k} = \binom{4}{k} \sum_{j=4}^{N} \sigma_j^2 \binom{N-4}{j-4} a_k(j),$$

where the numbers

$$a_k(j) = \frac{\binom{N-4}{j-k}}{\binom{N-4}{j-4}} \le N^{4-k}.$$

Invoking (171) we obtain

$$\mathbf{E}H_k^2 \le c\, N^{4-k} \mathbf{E}(D_1 \cdots D_4 \mathbb{T})^2 = cN^{-3-k} \Delta_4^2. \tag{178}$$

Finally, we obtain (176) for $m = 4$

$$\mathbf{E}\mathbb{H}^2(4) = \mathbf{E}H_1^2 + \cdots + \mathbf{E}H_4^2 \le cN^{-4} \Delta_4^2.$$

In order to prove (176) for $m = 5, 6, \ldots$ we apply a recursive argument. Write

$$\mathbf{E}\mathbb{H}^2(m+1) = \mathbf{E}\mathbb{H}^2(m) + \mathbf{E}d_m^2, \tag{179}$$

where $d_m = \mathbb{H}(m+1) - \mathbb{H}(m)$ is the sum of those T_A with $|A| \ge 4$ satisfying $A \cap \Omega_m = \emptyset$ and $A \cap \Omega_{m+1} \ne \emptyset$. In particular, we have

$$d_m = \sum_{|A| \ge 3,\; A \cap \Omega_{m+1} = \emptyset} T_{A \cup \{m+1\}}.$$

Therefore,

$$\mathbf{E}d_m^2 = \sum_{j=4}^{N} \sigma_j^2 \binom{N-m-1}{j-1} = \sum_{j=4}^{N} \sigma_j^2 \binom{N-4}{j-4} c_j,$$

where the numbers

$$c_j = \frac{\binom{N-m-1}{j-1}}{\binom{N-4}{j-4}} \le N^3.$$

Invoking (171) we obtain $\mathbf{E}d_m^2 \le N^{-4} \Delta_4^2$. This bound together with (179) implies (176).

Let us prove (177). Note that for $m = 2k$ we have $\mathbb{H}(m) = \mathbb{H}_A(k) + \mathbb{H}_B(k) + \mathbb{H}_C(k)$ and the summands are uncorrelated. Therefore, the first bound of (177) follows from (176).

Let us show the second inequality of (177). For $k = 2$ we have $\mathcal{C}(2) \subset \mathcal{C}$, where $\mathcal{C}$ denotes the class of subsets $A \subset \Omega_N$ such that $|A| \geq 4$ and $|A \cap \Omega_4| \geq 2$. Write $\mathbb{H}_C = \sum_{A \in \mathcal{C}} T_A$. We have

$$\mathbf{E}\mathbb{H}_C^2(2) \leq \mathbf{E}\mathbb{H}_C^2 = \mathbf{E}H_2^2 + \mathbf{E}H_3^2 + \mathbf{E}H_4^2 \leq cN^{-5}\Delta_4^2.$$

In the last step we applied (178). We obtain (177), for $k = 2$.

In order to prove the bound (177), for $k = 3, 4, \dots$, we apply a recursive argument similar to that used in the proof of (176). Denote

$$d_{[k]} = \mathbb{H}_C(k+1) - \mathbb{H}_C(k) = \sum_{A \in \mathcal{C}(k+1) \setminus \mathcal{C}(k)} T_A.$$

We shall show that

$$\mathbf{E}d_{[k]}^2 \leq ckN^{-5}\Delta_4^2. \tag{180}$$

This bound in combination with the identity $\mathbf{E}\mathbb{H}_C^2(k+1) = \mathbf{E}\mathbb{H}_C^2(k) + \mathbf{E}d_{[k]}^2$ shows (177) for arbitrary k.

In order to show (180) split the set $\mathcal{C}(k+1) \setminus \mathcal{C}(k)$ into $2k + 1$ non-intersecting parts

$$\mathcal{C}(k+1) \setminus \mathcal{C}(k) = \left(\cup_{i=1}^k \mathcal{C}_{x.i}\right) \cup \left(\cup_{i=1}^k \mathcal{C}_{y.i}\right) \cup \mathcal{C}_{x.y},$$

where we denote

$$\begin{aligned}
\mathcal{C}_{x.y} &= \left\{A = \tilde{A} \cup \{x_{k+1}, y_{k+1}\} : \ \tilde{A} \cap (B_k \cup A_k) = \emptyset, \ |\tilde{A}| \geq 2\right\}, \\
\mathcal{C}_{x.i} &= \left\{A = \tilde{A} \cup \{y_{k+1}, x_i\} : \ \tilde{A} \cap (B_k \cup A_{i-1}) = \emptyset, \ |\tilde{A}| \geq 2\right\}, \\
\mathcal{C}_{y.i} &= \left\{A = \tilde{A} \cup \{x_{k+1}, y_i\} : \ \tilde{A} \cap (B_{i-1} \cup A_k) = \emptyset, \ |\tilde{A}| \geq 2\right\}.
\end{aligned}$$

By the orthogonality property ($\mathbf{E}T_A T_V = 0$ for $A \neq V$), the random variables

$$d_{x.i} = \sum_{A \in \mathcal{C}_{x.i}} T_A, \qquad d_{y.i} = \sum_{A \in \mathcal{C}_{y.i}} T_A, \qquad d_{x.y} = \sum_{A \in \mathcal{C}_{x.y}} T_A$$

are uncorrelated. Therefore, we have

$$\mathbf{E}d_{[k]}^2 = \mathbf{E}d_{x.y}^2 + \sum_{i=1}^k (\mathbf{E}d_{x.i}^2 + \mathbf{E}d_{y.i}^2). \tag{181}$$

A calculation shows that

$$\mathbf{E}d_{x.y}^2 = \sum_{j=4}^N \sigma_j^2 \binom{N-2k-2}{j-2} = \sum_{j=4}^N \sigma_j^2 \binom{N-4}{j-4} v_j,$$

where the coefficients

$$v_j = \frac{\binom{N-2k-2}{j-2}}{\binom{N-4}{j-4}} \le N^2.$$

Invoking (171) we obtain $\mathbf{E}d^2_{x.y} \le N^{-5}\Delta_4^2$. The same argument shows $\mathbf{E}d^2_{x.i} = \mathbf{E}d^2_{y.i} \le N^{-5}\Delta_4^2$. The latter bound in combination with (181) shows (180). The lemma is proved. □

7 Appendix 2

Here we construct bounds for the probability density function (and its derivatives) of random variables $g_k^* = (N/M)^{1/2} g_k$, for $1 \le k \le n-1$, where g_k are defined in (74). Since these random variables are identically distributed it suffices to consider

$$g_1^* = \Big(\frac{N}{M}\Big)^{1/2} g_1 = \frac{1}{\sqrt{M}} \sum_{j=m+1}^{m+M} g(Y_j) + \frac{\xi_1}{R}.$$

Here $R = \sqrt{n\,M\,N}$. Introduce the random variables

$$g_2^* = g_1^* - M^{-1/2} g(Y_{m+1}), \qquad g_3^* = g_1^* - M^{-1/2}\big(g(Y_{m+1}) + g(Y_{m+2})\big).$$

Let $p_i(\cdot)$ denote the probability density function of g_i^*, for $i = 1, 2, 3$. Recall that the integers $n \approx N^{50\nu} \le N^{\nu_2/10}$ and $M \approx N/n \ge N^{9/10}$ are introduced in (29) and the number $\nu > 0$ is defined by (17).

Lemma 7 *Assume that conditions of Theorem 1 are satisfied. There exist positive constants C_*, c_*, c'_* depending only on M_*, D_*, δ, r and ν_1, ν such that, for $i = 1, 2, 3$, we have uniformly in $u \in \mathbb{R}$ and $N > C_*$*

$$|p_i(u)| \le c_*, \quad |p_i'(u)| \le c_*, \quad |p_i''(u)| \le c_*, \quad |p_i'''(u)| \le c_*. \tag{182}$$

Furthermore, given $w > 0$ there exists a constant $C_*(w)$ depending on $M_*, D_*, \delta, r, \nu_1, \nu$ and w such that uniformly in $z_* \in [-2w, 2w]$ and $N > C_*(w)$ we have

$$p_i(z_*) \ge c'_*, \qquad i = 1, 2, 3. \tag{183}$$

Proof We shall prove (182) and (183) for $i = 1$. For $i = 2, 3$, the proof is almost the same. Before starting the proof we introduce some notation and collect auxiliary results.

Denote

$$\theta = \mathbf{E}g_1^* = M^{1/2}\theta_1, \qquad \theta_1 = \mathbf{E}g(Y_{m+1}),$$

$$s^2 = \mathbf{E}(g(Y_{m+1}) - \theta_1)^2, \qquad \tilde{\beta}_3 = s^{-3}\mathbf{E}|g(Y_{m+1}) - \theta_1|^3$$

and recall that $q_N = \mathbf{P}\{A_j\}$, where $A_j = \{\|Z'_j\|_r \le N^{\alpha}\}$. It follows from $\mathbf{E}g(X_{m+1}) = 0$ that

$$\theta_1 = q_N^{-1}\mathbf{E}g(X_{m+1})\mathbb{I}_{A_{m+1}} = -q_N^{-1}\mathbf{E}g(X_{m+1})(1 - \mathbb{I}_{A_{m+1}}).$$

Therefore, by Chebyshev's inequality, for $\alpha = 3/(r+2)$ we have

$$|\theta_1| \le q_N^{-1} N^{-\alpha(r-1)}\mathbf{E}|g(X_{m+1})|\,\|Z'_{m+1}\|_r^{r-1} \le c_* N^{-3/2}. \tag{184}$$

In the last step we invoke the inequalities $\alpha(r-1) \ge 1 + (r-1)/(r+2) \ge 3/2$ and $q_N^{-1} \le c_*$, see (45), and $\mathbf{E}|g(X_{m+1})|\,\|Z'_{m+1}\|_r^{r-1} \le M_*$, where the latter inequality follows from (5) by Hölder inequality.

Similarly, the identities

$$s^2 = q_N^{-1}\mathbf{E}g^2(X_{m+1})\mathbb{I}_{A_{m+1}} - \theta_1^2 = q_N^{-1}\sigma^2 - q_N^{-1}\mathbf{E}g^2(X_{m+1})(1 - \mathbb{I}_{A_{m+1}}) - \theta_1^2$$

in combination with (44) and the inequalities

$$\mathbf{E}g^2(X_{m+1})(1 - \mathbb{I}_{A_{m+1}}) \le N^{-\alpha(r-2)}\mathbf{E}g^2(X_{m+1})\,\|Z'_{m+1}\|_r^{r-2} \le N^{-\alpha(r-2)}M_*$$

and $\alpha(r-2) = 1 + 2(r-4)/(r-2) \ge 1$ yield

$$|s^2 - \sigma^2| \le c_* N^{-1}. \tag{185}$$

Introduce the random variables

$$g_* = S + \frac{\xi_1}{s\,R}, \quad S = w_1 + \cdots + w_M, \quad w_j = \frac{g(Y_{m+j}) - \theta_1}{M^{1/2}s}.$$

We have $g_* = s^{-1}(g_1^* - \theta)$. Let $p(\cdot)$ denote the density function of g_*. Note that $p_1(u) = s^{-1}p\big(s^{-1}(u-\theta)\big)$. Furthermore, we have, by (184), $|\theta| \le c_* N^{-1}$ and, by (185), (169), $|s^2 - 1| \le c_* N^{-1}$. Therefore, it suffices to prove (182) and (183) for $p(\cdot)$ (the latter inequality we verify for every $z_* \in [-3w, 3w]$).

In order to prove (182) and (183) we approximate the characteristic function $\hat{p}(t) = \mathbf{E}e^{itg_*}$ by $e^{-t^2/2}$ and then apply a Fourier inversion formula. Write

$$\hat{p}(t) = \mathbf{E}e^{itg_*} = \gamma^M(t)\tau\Big(\frac{t}{sR}\Big), \qquad \gamma(t) := \mathbf{E}e^{itw_1}, \quad \tau(t) := \mathbf{E}e^{it\xi_1}.$$

The fact that $\tau(t) = 0$, for $|t| \ge 1$, implies $\hat{p}(t) = 0$, for $|t| > s\,R$. Therefore, we obtain from the Fourier inversion formula,

$$p(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx}\hat{p}(t)dt = \frac{1}{2\pi}\int_{-s\,R}^{s\,R} e^{-itx}\hat{p}(t)dt.$$

Write $\hat{p}(t) - e^{-t^2/2} = r_1(t) + r_2(t)$, where

$$r_1(t) = (\gamma^M(t) - e^{-t^2/2})\tau(t/sR), \qquad r_2(t) = e^{-t^2/2}\big(\tau(t/sR) - 1\big).$$

We shall show below that

$$\int_{|t|\le sR} |r_i(t)|dt \le c_* M^{-1/2}, \qquad i = 1, 2. \tag{186}$$

These bounds in combination with the simple inequality

$$\int_{|t|\ge sR} e^{-t^2/2} dt \le c_* M^{-1/2}$$

show that

$$|p(x) - \varphi(x)| \le c_* M^{-1/2}, \qquad x \in \mathbb{R}. \tag{187}$$

Here φ denotes the standard normal density function

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx} e^{-t^2/2} dt.$$

It follows from (187) that

$$|p(x)| \le c_*, \qquad x \in \mathbb{R}.$$

Furthermore, given w we have uniformly in $|z_*| \le 3w$

$$|p(z_*)| \ge \varphi(3w) - c_* M^{-1/2} \ge c'_* > 0,$$

for sufficiently large M (for $N > C_*(w)$).

In order to prove an upper bound for the $k-$th derivative, $|p^{(k)}(x)| \le c_*$, write

$$p^{(k)}(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} (-it)^k \exp\{-itx\}\hat{p}(t)dt, \qquad k = 1, 2, 3,$$

and replace $\hat{p}(t)$ by $e^{-t^2/2}$ as in the proof of (187). We obtain

$$p^{(k)}(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} (-it)^k \exp\{-itx\} e^{-t^2/2}(t)dt + r, \qquad |r| \le c_* M^{-1/2}.$$

This implies $|p^{(k)}(x) - \varphi^{(k)}(x)| \le c_* M^{-1/2}$. We arrive at the desired bound $|p^{(k)}(x)| \le c_*$, for $k = 1, 2, 3$.

In the remaining part of the proof we verify (186). For $i = 2$ this bound follows from $|\tau(t/sR) - 1| \le ct^2/(sR)^2$. The latter inequality is a consequence of the short expansion

$$\big|\mathbf{E}\exp\{it\xi_1/sR\} - 1 - \mathbf{E}it\xi_1/sR\big| \le \mathbf{E}(t\xi_1)^2/2(sR)^2$$

and $\mathbf{E}\xi_1 = 0$ and $\mathbf{E}\xi_1^2 \le c$, for some absolute constant c.

Let us prove (186) for $i = 1$. Introduce the sequence of i.i.d. centered Gaussian random variables $\eta_1, \eta_2, \dots$ with variances $\mathbf{E}\eta_i^2 = M^{-1}$. Denote

$$f(t) = \mathbf{E}e^{it\eta_1} = e^{-t^2/(2M)} \qquad \text{and} \qquad \delta(t) = \gamma(t) - f(t).$$

We are going to apply the well known inequality

$$\left|e^{iv} - \left(1 + \frac{iv}{1!} + \frac{(iv)^2}{2!} + \dots + \frac{(iv)^{k-1}}{(k-1)!}\right)\right| \le \frac{|v|^k}{k!}. \tag{188}$$

It follows from (188) and identities $\mathbf{E}\eta_1^i = \mathbf{E}w_1^i$, $i = 1, 2$, that

$$|\delta(t)| \le \frac{|t|^3}{3!}\big(\mathbf{E}|w_1|^3 + \mathbf{E}|\eta_1|^3\big) \le c|t|^3\mathbf{E}|w_1|^3. \tag{189}$$

Here we use the inequality $\mathbf{E}|\eta_1|^3 \le c\mathbf{E}|w_1|^3$, which follows from $\mathbf{E}\eta_1^2 = \mathbf{E}w_1^2$.

Combining (189) and the simple identity

$$\gamma^M(t) - f^M(t) = \delta(t)\sum_{k=1}^{M}\gamma^{M-k}(t)f^{k-1}(t)$$

we obtain

$$|\gamma^M(t) - f^M(t)| \le c|t|^3 Z(t)\, M^{-1/2}\tilde{\beta}_3. \tag{190}$$

Here we denote

$$Z(t) = \max_{r+v=M-1} |f^r(t)\gamma^v(t)|.$$

We shall show below that

$$Z(t) \le \exp\left\{-\frac{t^2}{3}\frac{M-1}{M}\right\} + \exp\{-\delta''(M-1)/2\}, \qquad 0 \le |t| \le sR, \tag{191}$$

where $\delta'' > 0$ depends on δ, A_*, D_*, M_*, ν_1 and it is given in (36). This inequality in combination with (190) proves (186).

Let us prove (191). Clearly, $Z \le |f^{M-1}(t)| + |\gamma^{M-1}(t)|$. Furthermore, $f^M(t) = e^{-t^2/2}$. In order to prove (191) we shall show

$$|\gamma^M(t)| \le e^{-t^2/3}, \qquad 0 \le |t| \le M^{1/2}/\tilde{\beta}_3, \tag{192}$$

$$|\gamma(t)| \le e^{-\delta''/2}, \qquad M^{1/2}/\tilde{\beta}_3 \le |t| \le sR. \tag{193}$$

To show (192) we expand e^{itw_1} using (188),

$$\begin{aligned}
|\gamma(t)| = \left|\mathbf{E}e^{itw_1}\right| &\le \left|1 - \frac{t^2}{2}\mathbf{E}w_1^2\right| + \frac{|t|^3}{3!}\mathbf{E}|w_1|^3 \\
&= \left|1 - \frac{t^2}{2M}\right| + \frac{|t|^3}{3!}\frac{\tilde{\beta}_3}{M^{3/2}} \\
&= 1 - \frac{t^2}{2M}\left(1 - \frac{|t|}{3}\frac{\tilde{\beta}_3}{\sqrt{M}}\right) \\
&\le 1 - \frac{t^2}{3M}.
\end{aligned}$$

Here we used the identity $|1 - t^2/2M| = 1 - t^2/2M$, which holds for $|t| < M^{1/2}/\tilde{\beta}_3$, since $\tilde{\beta}_3 \ge 1$. Finally, an application of the inequality $1 - x \le e^{-x}$ to $x = t^2/3M > 0$ completes the proof of (192).

Let us prove (193). For δ'' defined by (36) we shall show $\delta'' \le 2\tilde{\delta}$, where

$$\begin{aligned}
\tilde{\delta} &= 1 - \sup\{|\gamma(t)| : \ M^{1/2}\tilde{\beta}_3^{-1} \le |t| \le sR\} \\
&= 1 - \sup\{|\mathbf{E}\exp\{iu\sigma^{-1}g(Y_{m+1})| : \ \sigma/s\,\tilde{\beta}_3 \le |u| \le \sigma\sqrt{n\,N}\}.
\end{aligned}$$

We are going to replace $g(Y_{m+1}),\ \tilde{\beta}_3,\ s^2$ by $g(X_{m+1}),\ \beta_3,\ \sigma^2$ respectively. Write

$$\begin{aligned}
&\mathbf{E}e^{ivg(Y_{m+1})} = q_N^{-1}\mathbf{E}e^{ivg(X_{m+1})}\mathbb{I}_{A_{m+1}} = \mathbf{E}e^{ivg(X_{m+1})} + r_1 + r_2, \\
&r_1 = q_N^{-1}\mathbf{E}e^{ivg(X_{m+1})}\big(\mathbb{I}_{A_{m+1}} - 1\big), \qquad r_2 = (q_N^{-1} - 1)\mathbf{E}e^{ivg(X_{m+1})}.
\end{aligned}$$

It follows from (44), (45) that, for every $v \in \mathbb{R}$,

$$\begin{aligned}
&|r_1| \le q_N^{-1}\mathbf{E}|\mathbb{I}_{A_{m+1}} - 1| = q_N^{-1}(1 - \mathbf{P}\{A_{m+1}\}) = q_N^{-1} - 1 \le c_* N^{-2}, \\
&|r_2| \le q_N^{-1} - 1 \le c_* N^{-2}.
\end{aligned}$$

These bounds imply

$$|\mathbf{E}e^{ivg(Y_{m+1})} - \mathbf{E}e^{ivg(X_{m+1})}| \le c_* N^{-2}, \qquad \text{for every} \qquad v \in \mathbb{R}. \tag{194}$$

One can show that, for sufficiently large N (i.e., for $N > C_*$), we have

$$|\tilde{\beta}_3/\beta_3 - 1| < 1/5, \qquad |s^2/\sigma^2 - 1| < 1/5, \qquad |s^2 - 1| \le 1/5. \tag{195}$$

Using (194), (195) we get, for $N > C_*$,

$$\tilde{\delta} \ge 1 - \sup\big\{|\mathbf{E}e^{iu\sigma^{-1}g(Y_{m+1})}| : \ (2\beta_3)^{-1} \le |u| \le N^{(1+50\nu)/2}\big\}$$

$$\geq 1 - \sup\bigl\{|\mathbf{E}e^{iu\sigma^{-1}g(X_{m+1})}| : (2\beta_3)^{-1} \leq |u| \leq N^{(\nu_2+1)/2}\bigr\} - c_* N^{-2}$$
$$\geq \delta''/2.$$

We obtain $|\gamma(t)| \leq 1 - \tilde{\delta} \leq 1 - \delta''/2$ and, therefore, $|\gamma(t)| \leq e^{-\delta''/2}$. The lemma is proved. □

8 Appendix 3

The main results of this section are moment inequalities of Lemma 9 and corresponding inequalities for conditional moments of Lemma 10. Lemma 8 provides an auxiliary inequality.

We start with some notation. We call $v = v(\cdot), u = u(\cdot) \in L^r$ orthogonal if $\langle u, v\rangle = 0$, where

$$\langle u, v\rangle = \int_{\mathcal{X}} u(x)v(x)P_X(dx) = \mathbf{E}u(X_1)v(X_1).$$

Given $f \in L^2(P_X)$ we have for the kernel ψ^{**} defined in (41)

$$\mathbf{E}\psi^{**}(X_1, X_2)\bigl(f(X_1)g(X_2) + f(X_2)g(X_1)\bigr) = 0$$

and almost surely

$$\mathbf{E}(\psi^{**}(X_1, X_2)|X_1) = 0, \tag{196}$$
$$\mathbf{E}\bigl(\psi^{**}(X_1, X_2)g(X_1)|X_2\bigr) = 0. \tag{197}$$

The latter identity says that almost all values of the L^r-valued random variable $\psi^{**}(\cdot, X_2)$ are orthogonal to the vector $g(\cdot) \in L^r$.

Let $p_g : L^r \to L^r$ denote the projection on the subspace of elements $u \in L^r$ which are orthogonal to $g = g(\cdot)$. For $v \in L^r$, write $v^* = p_g(v)$. It follows from (197) that

$$\psi^*(\cdot, Y_j)\Bigl(= p_g\bigl(\psi(\cdot, Y_j)\bigr)\Bigr) = \psi^{**}(\cdot, Y_j) + g(Y_j)b^*(\cdot), \tag{198}$$

where $b^*(\cdot) = p_g(b(\cdot)) = \sigma^{-2}p_g\bigl(\mathbf{E}(\psi(\cdot, X_1)g(X_1)\bigr)$. Denote

$$U_k^*\bigl(= p_g(U_k)\bigr) = \frac{1}{\sqrt{N}}\sum_{j\in O_k}\psi^*(\cdot, Y_j), \qquad U_k^{**} = N^{-1/2}\sum_{j\in O_k}\psi^{**}(\cdot, Y_j),$$

where the L^r-valued random variables U_k are introduced in (91). For the random variables g_k and L_k introduced in (72) and (74), we have

$$U_k^* = U_k^{**} + L_k\, b^*(\cdot) = U_k^{**} + (g_k - \frac{1}{\sqrt{n}}\frac{\xi_k}{N})b^*(\cdot). \tag{199}$$

Denote $K = \mathbf{E}|\psi(X_1, X_2)|^r$ and $K_s = \mathbf{E}|\psi^{**}(X_1, X_2)|^s$, $s \le r$.

Lemma 8 *Let* $4 < r \le 5$. *For* $s \le r$, *we have*

$$K_s^{r/s} \le K_r \le c\,K\Big(1 + \frac{\mathbf{E}|g(X_1)|^r}{\sigma^r}\Big)^2. \tag{200}$$

Proof The first inequality of (200) is a consequence of Lyapunov's inequality. Let us prove the second inequality. The inequality $|a + b + c|^r \le 3^r(|a|^r + |b|^r + |c|^r)$ implies

$$K_r = \mathbf{E}|\psi^{**}(X_1, X_2)|^r \le 3^r\big(K + 2\mathbf{E}|b(X_1)|^r\mathbf{E}|g(X_2)|^r\big).$$

Therefore, (200) is a consequence of the inequalities

$$\mathbf{E}|b(X_1)|^r \le \frac{2^r}{\sigma^r}K + \frac{|\kappa|^r}{\sigma^{4r}}\mathbf{E}|g(X_1)|^r,$$
$$\kappa^2 \le \sigma^4\mathbf{E}\psi^2(X_1, X_2) \le \sigma^4 K^{2/r}.$$

Here $\kappa = \mathbf{E}\psi(X_1, X_2)g(X_1)g(X_2)$. To prove the first inequality use $|a + b|^r \le 2^r(|a|^r + |b|^r)$ to get

$$\mathbf{E}|b(X_1)|^r \le \frac{2^r}{\sigma^{2r}}\mathbf{E}\big|\mathbf{E}\big(\psi(X_1, X_2)g(X_2)\big|X_1\big)\big|^r + \frac{\kappa^r}{\sigma^{4r}}\mathbf{E}|g(X_1)|^r.$$

Furthermore, by Cauchy–Schwartz,

$$\big|\mathbf{E}\big(\psi(X_1, X_2)g(X_2)\,|\,X_1\big)\big| \le \big(\mathbf{E}(\psi^2(X_1, X_2)\,|\,X_1)\big)^{1/2}\sigma.$$

Finally, Lyapunov's inequality implies

$$\big(\mathbf{E}(\psi^2(X_1, X_2)\,|\,X_1)\big)^{r/2} \le \mathbf{E}(|\psi(X_1, X_2)|^r\,|\,X_1).$$

We obtain $\mathbf{E}\big|\mathbf{E}\big(\psi(X_1, X_2)g(X_2)\,|\,X_1\big)\big|^r \le K\sigma^r$ thus completing the proof. □

Lemma 9 *Let* $1 \le k \le n - 1$. *For* $\overline{U}_k^*$, *an independent copy of* U_k^*, *we have*

$$2\delta_3^2 - \frac{c_*}{N^{\alpha(r-4)}} \le \frac{N}{M}\mathbf{E}\|U_k^* - \overline{U}_k^*\|_2^2 \le 2\delta_3^2 + c_*, \tag{201}$$
$$\mathbf{E}\|U_k - \overline{U}_k\|_r^r \le c_*\Big(\frac{M}{N}\Big)^{r/2}. \tag{202}$$

Recall that $\delta_3^2 = \mathbf{E}|\psi^{**}(X_1, X_2)|^2$.

Proof Let us prove (201). By symmetry, we have, for $i, j \in O_1$,

$$\mathbf{E}\|U_1^* - \overline{U}_1^*\|_2^2 = 2\frac{M}{N}H_1 - 2\frac{M}{N}H_2,$$
$$H_1 := \mathbf{E}\|\psi^*(\cdot, Y_j)\|_2^2, \qquad H_2 := \mathbf{E}\langle\psi^*(\cdot, Y_j), \psi^*(\cdot, Y_i)\rangle, \quad i \neq j.$$

The inequality (201) follows from the inequalities

$$\delta_3^2 - c_* N^{-\alpha(r-4)} \le H_1 \le \delta_3^2 + c_*, \tag{203}$$
$$H_2 \le c_* N^{-\alpha(r-2)}. \tag{204}$$

Let us prove (203). From (198) we have $H_1 = V_1 + V_2 + 2V_3$, where

$$V_1 = \mathbf{E}\|\psi^{**}(\cdot, Y_j)\|_2^2, \quad V_2 = \|b^*(\cdot)\|_2^2 \mathbf{E} g^2(Y_j), \quad V_3 = \mathbf{E} g(Y_j)\langle\psi^{**}(\cdot, Y_j), b^*(\cdot)\rangle.$$

Let us show that

$$\delta_3^2 - c_* N^{-\alpha(r-2)} \le V_1 \le \delta_3^2 + c_* N^{-\alpha r}. \tag{205}$$

This inequality follows from (44), (45) and the identity

$$V_1 = q_N^{-1}\mathbf{E}|\psi^{**}(X_1, X_j)|^2 \mathbb{I}_{A_j} = q_N^{-1}\mathbf{E}|\psi^{**}(X_1, X_j)|^2 - q_N^{-1} V_1',$$

where $V_1' = \mathbf{E}|\psi^{**}(X_1, X_j)|^2(1 - \mathbb{I}_{A_j})$ satisfies, by (43),

$$0 \le V_1' \le N^{-\alpha(r-2)}\mathbf{E}|\psi^{**}(X_1, X_j)|^2 \|Z_j'\|_r^{r-2} \le c_* N^{-\alpha(r-2)}. \tag{206}$$

In the last step we applied Hölder's inequality and Lemma 8 to get

$$\mathbf{E}|\psi^{**}(X_1, X_j)|^2 \|Z_j'\|_r^{r-2} \le K_r^{2/r}(\mathbf{E}\|Z_j'\|_r^r)^{(r-2)/r} \le K_r^{2/r} K^{(r-2)/r} \le c_*.$$

Let us show that

$$0 \le V_2 \le c_*. \tag{207}$$

For $\tilde{b}(\cdot) := \mathbf{E}\psi(\cdot, X_1) g(X_1)$ we have, by Cauchy–Schwartz,

$$\|\tilde{b}(\cdot)\|_2^2 = \mathbf{E}\big(\mathbf{E}(\psi(X, X_1) g(X_1) \big| X)\big)^2 \le \mathbf{E}\psi^2(X, X_1)\,\sigma^2 \le c_*\sigma^2.$$

Now the identity $b^* = \sigma^{-2} p_g(\tilde{b})$ implies

$$\|b^*\|_2^2 \le \sigma^{-4}\|\tilde{b}\|_2^2 \le \sigma^{-2} c_*. \tag{208}$$

Invoking the bound $\mathbf{E} g^2(Y_j) \le c_*\sigma^2$, see (47), we obtain (207).

Finally, write

$$V_3 = q_N^{-1}\mathbf{E}\tilde{V}\mathbb{I}_{A_j}, \qquad \tilde{V} = g(X_j)\psi^{**}(X_1, X_j) b^*(X_1).$$

Identity (197) implies $\mathbf{E}\tilde{V} = 0$. Therefore $V_3 = q_N^{-1}\mathbf{E}\tilde{V}(\mathbb{I}_{A_j} - 1)$. Invoking (43) and using $q_N^{-1} \le c_*$, see (45), we obtain

$$|V_3| \le c_* N^{-\alpha(r-4)}\mathbf{E}|\tilde{V}|\|Z_j'\|_r^{r-4} \le c_* N^{-\alpha(r-4)}. \tag{209}$$

In the last step we used the bound $\mathbf{E}|\tilde{V}|\|Z_j'\|_r^{r-4} \le c_*$. In order to prove this bound we invoke the inequalities

$$|abc| \le (ab)^2 + c^2 \le a^4 + b^4 + c^2$$

to show that

$$|\tilde{V}| \le |g(X_j)|^4 + |\psi^{**}(X_1, X_j)|^4 + |b^*(X_1)|^2.$$

Furthermore, by Hölder's inequality and (200),

$$\mathbf{E}|g(X_j)|^4\|Z_j'\|_r^{r-4} \le c_*, \qquad \mathbf{E}|\psi^{**}(X_1, X_j)|^4\|Z_j'\|_r^{r-4} \le c_*.$$

By the independence and (208),

$$\mathbf{E}|b^*(X_1)|^2\|Z_j'\|_r^{r-4} = \|b^*\|_2^2\mathbf{E}\|Z_j'\|_r^{r-2} \le c_*.$$

Thus we arrive at (209). Combining (205), (207) and (209) we obtain (203).

Let us prove (204). Using (198) write $H_2 = Q_1 + Q_2 + 2Q_3$, where

$$Q_1 = \mathbf{E}\psi^{**}(X_1, Y_j)\psi^{**}(X_1, Y_i), \qquad Q_2 = \|b^*\|_2^2\mathbf{E}g(Y_j)g(Y_i),$$
$$Q_3 = \mathbf{E}\psi^{**}(X_1, Y_j)g(Y_i)b^*(X_1).$$

It follows from the identity (196) that

$$Q_1 = q_N^{-2}\mathbf{E}\psi^{**}(X_1, X_j)\psi^{**}(X_1, X_i)(\mathbb{I}_{A_j} - 1)(\mathbb{I}_{A_i} - 1).$$

The simple inequality $|\psi^{**}(X_1, X_j)\psi^{**}(X_1, X_i)| \le |\psi^{**}(X_1, X_j)|^2 + |\psi^{**}(X_1, X_i)|^2$ yields, by symmetry,

$$|Q_1| \le 2q_N^{-2}\mathbf{E}|\psi^{**}(X_1, X_j)|^2(1 - \mathbb{I}_{A_j}) \le c_* N^{-\alpha(r-2)}. \tag{210}$$

In the last step we applied (206) and $q_N^{-1} \le c_*$, see (44).

Furthermore, using the identity $\mathbf{E}g(X_i) = 0$ we obtain from (43)

$$\begin{aligned} |\mathbf{E}g(Y_i)| &= q_N^{-1}|\mathbf{E}g(X_i)(\mathbb{I}_{A_i} - 1)| \\ &\le q_N^{-1}N^{-\alpha(r-1)}\mathbf{E}|g(X_i)|\|Z_i\|_r^{r-1} \le c_* N^{-\alpha(r-1)}. \end{aligned} \tag{211}$$

In the last step we applied Hölder's inequality to show $\mathbf{E}|g(X_i)|\|Z_i\|_r^{r-1} \le c_*$.

The bounds (211), (44) and (208) together imply

$$|Q_k| \le c_* N^{-\alpha(r-1)}, \qquad k = 2, 3. \tag{212}$$

The bound (204) follows from (210) and (212).

Let us prove (202). For this purpose we shall show that

$$\mathbf{E}\Big\| \sum_{j \in O_k} \frac{V_j}{\sqrt{M}} \Big\|_r^r \le c_*, \qquad \text{where} \qquad V_j = \psi(\cdot, Y_j) - \psi(\cdot, \overline{Y}_j), \tag{213}$$

and where $\overline{Y}_j$ denote independent copies of Y_j, $j \in O_k$. Using

$$\mathbf{E}\|\psi(\cdot, X_j)\|_r^r = \mathbf{E}|\psi(X_1, X_j)|^r \le c_*$$

we obtain, by symmetry and (47),

$$\mathbf{E}\|V_j\|_r^r \le 2^r \mathbf{E}\|\psi(\cdot, Y_j)\|_r^r \le c_* \mathbf{E}\|\psi(\cdot, X_j)\|_r^r \le c_*.$$

Now (213) follows from the well known inequality

$$\|\xi_1 + \cdots + \xi_k\|_r^r \le c(r) \sum_{i=1}^{k} \mathbf{E}\|\xi_i\|_r^r + c(r) \left(\sum_{i=1}^{k} \mathbf{E}\|\xi_i\|_r^2 \right)^{r/2}, \quad k = 1, 2, \dots \tag{214}$$

which is valid for independent centered random elements ξ_i with values in L^r. One can derive this inequality from Hoffmann–Jorgensen's inequality (see e.g., Proposition 6.8 in Ledoux and Talagrand [32]) using the type 2 property of the Banach space L^r and the symmetrization lemma (see formula (9.8) and Lemma 6.3 ibidem). The proof of the lemma is complete. □

Before formulating and proving Lemma 10 we introduce some more notation. Let $\mathcal{B}(L^r)$ denote the class of Borel sets of L^r. Consider the regular conditional probability $P_k : \mathbb{R} \times \mathcal{B}(L^r) \to [0, 1]$, defined, for $z_k \in \mathbb{R}$ and $B \in \mathcal{B}(L^r)$,

$$P_k(z_k; B) := \mathbf{P}\big(U_k \in B \,\big|\, g_k = z_k\big) = \mathbf{E}(\mathbb{I}_{U_k \in B} | g_k = z_k).$$

Recall, see (82), that ψ_k denotes a L^r valued random variable with the distribution $\mathbf{P}\{\psi_k \in B\} = P_k(z_k; B)$. Note that the L^r valued random variable $\psi_k^* = p_g(\psi_k)$ has distribution

$$\begin{aligned} \mathbf{P}\{\psi_k^* \in B\} &= \mathbf{P}\{p_g(\psi_k) \in B\} = \mathbf{P}\{\psi_k \in p_g^{-1}(B)\} \\ &= \mathbf{P}\big(U_k \in p_g^{-1}(B) \,\big|\, g_k = z_k\big) = \mathbf{P}\big(U_k^* \in B \,\big|\, g_k = z_k\big). \end{aligned} \tag{215}$$

Furthermore, using (199) we write (215) in the form

$$\mathbf{P}\{\psi_k^* \in B\} = \mathbf{P}\left(U_k^{**} + \left(z_k - \frac{1}{N} \frac{\xi_k}{n^{1/2}} \right) b^* \in B \,\Big|\, g_k = z_k \right).$$

Let $\overline{\psi}_k$ respectively $\overline{\psi}_k^*$ denote an independent copy of ψ_k respectively ψ_k^*. Denote

$$\tau_N = M^{-(r-4)/2} + N^{-\alpha(r-2)} M.$$

Lemma 10 *Let $k = 1, \ldots, n-1$. Let $|z_k| \le w\, n^{-1/2}$. There exist positive constants $c_*^{(i)}$, $i = 0, 1, 2, 3$, which depend on $w, r, \nu_1, \nu_2, \delta, A_*, D_*, M_*$ only such that for*

$$\tau_N \le c_*^{(0)} \delta_3^2, \tag{216}$$

we have

$$c_*^{(1)} \delta_3^2 \le n\mathbf{E}\|\psi_k^* - \overline{\psi}_k^*\|_2^2 \le c_*^{(2)} \delta_3^2 \tag{217}$$

$$\mathbf{E}\|\psi_k - \overline{\psi}_k\|_r^r \le c_*^{(3)} n^{-r/2}. \tag{218}$$

Condition (216) requires N to be large enough. A simple calculation shows $\tau_N \le N^{-75\nu}$, for ν satisfying (17). Therefore, (87) implies $\tau_N \le N^{-65\nu}\delta_3^2$. In particular, under (87) the inequality (216) is satisfied provided that $N > c_*$, where c_* does not depend on δ_3^2.

Proof By $\tilde{c}_*, \tilde{c}'_*$ we denote positive constants which depend only on $w, r, \nu_1, \nu_2, \delta, A_*, D_*, M_*$. These constants can be different in different places of the text. Given $i, j \in O_k$, $i \ne j$, introduce random variables

$$g_* = \eta + \zeta, \qquad \eta = \frac{\xi_k}{R}, \quad \zeta = \frac{1}{\sqrt{M}} \sum_{j \in O_k} g(Y_j),$$

$$\zeta_i = \zeta - \frac{g(Y_i)}{\sqrt{M}}, \qquad \zeta_{ij} = \zeta - \frac{g(Y_i)}{\sqrt{M}} - \frac{g(Y_j)}{\sqrt{M}}.$$

Here $R = \sqrt{n\,M\,N}$ satisfies $N/2 \le R \le N$, by the choice of n and M. Let p, p_0, p_1, and p_2 denote the densities of random variables η, $\zeta + \eta$, $\zeta_i + \eta$, and $\zeta_{ij} + \eta$ respectively.

Note that $g_* = \sqrt{N/M} g_k$. Therefore, the condition $g_k = z_k$ is equivalent to $g_* = z_*$, where $z_* = \sqrt{N/M} z_k$. Furthermore, $|z_k| \le w\, n^{-1/2} \Leftrightarrow |z_*| \le w_*$, where $w_* = w\sqrt{N/Mn} \le 2w$.

Given a random variable Y, we denote the conditional expectation $\mathbf{E}(Y|g_* = z_*) = \mathbf{E}(Y|g_k = z_k)$ by $\mathbf{E}_* Y$. For an event A, we have $P(A|g_k = z_k) = P(A|g_* = z_*)$.

Proof of (217). For the L^r valued random variable $\hat{\psi}^* = \psi_k^* - z_k b^*$ we have

$$\mathbf{P}\{\hat{\psi}^* \in B\} = P\Big(U_k^{**} - \frac{1}{N}\frac{\xi_k}{n^{1/2}} b^* \in B \,\Big|\, g_* = z_*\Big). \tag{219}$$

Note that for an independent copy $\overline{\psi}_k^*$ of ψ_k^* the distributions of $\psi_k^* - \overline{\psi}_k^*$ and $\hat{\psi}^* - \hat{\psi}_c^*$ are the same. Here $\hat{\psi}_c^*$ denotes an independent copy of $\hat{\psi}^*$. Therefore,

$$\mathbf{E}\|\psi_k^* - \overline{\psi}_k^*\|_2^2 = \mathbf{E}\|\hat{\psi}^* - \hat{\psi}_c^*\|_2^2 = 2\mathbf{E}\|\hat{\psi}^*\|_2^2 - 2\|\mathbf{E}\hat{\psi}^*\|_2^2. \tag{220}$$

In order to prove (217) we show that

$$\|\mathbf{E}\hat{\psi}^*\|_2^2 \le \tilde{c}_* N^{-1} \tag{221}$$

and, for $\tau_N \le c_*^{(0)} \delta_3^2$ (i.e., for sufficiently large N),

$$\tilde{c}_* \delta_3^2 \le n \mathbf{E} \|\hat{\psi}^*\|_2^2 \le \tilde{c}'_* \delta_3^2. \tag{222}$$

Since $N^{-1} n < \tau_N$, we can choose $c_*^{(0)}$ small enough such that the inequalities (220), (221) and (222) together imply (217)

Proof of (221). Recall that an element $m = m(\cdot) \in L^2(P_X)$ is called mean of an $L^2(P_X)$ valued random variable $\hat{\psi}^* = \hat{\psi}^*(\cdot)$ if for every $f = f(\cdot) \in L^2(P_X)$

$$\langle f, m \rangle = \mathbf{E} \Big\langle f, \hat{\psi}^* \Big\rangle.$$

We shall show below that $\mathbf{E}\|\hat{\psi}^*\|_2^2 < \infty$. Then, by Fubini,

$$\mathbf{E} \Big\langle f, \hat{\psi}^* \Big\rangle = \int f(x) \mathbf{E}\, \hat{\psi}^*(x) P_X(dx).$$

Therefore, $m(x) = \mathbf{E}\hat{\psi}^*(x)$, for P_X almost all x.

For $f \in L^2(P_X)$ it follows from (219) that

$$\begin{aligned} \mathbf{E} \Big\langle f, \hat{\psi}^* \Big\rangle &= \mathbf{E}_* \Big\langle f, U_k^{**} - \frac{1}{N} \frac{\xi_k}{n^{1/2}} b^* \Big\rangle \\ &= \mathbf{E}_* \big\langle f, U_k^{**} \big\rangle - \frac{\sqrt{M}}{\sqrt{N}} \big\langle f, b^* \big\rangle \mathbf{E}_* \eta. \end{aligned} \tag{223}$$

Fix $i \in O_k$. By symmetry,

$$\mathbf{E}_* \big\langle f, U_k^{**} \big\rangle = \frac{M}{\sqrt{N}} \mathbf{E}_* \big\langle f, \psi^{**}(\cdot, Y_i) \big\rangle. \tag{224}$$

An application of (252) yields

$$\begin{aligned} \mathbf{E}_* \big\langle f, \psi^{**}(\cdot, Y_i) \big\rangle &= \frac{1}{p_0(z_*)} \mathbf{E} \big\langle f, \psi^{**}(\cdot, Y_i) \big\rangle\, p_1\big(z_* - \frac{g(Y_i)}{\sqrt{M}}\big) \\ &= \big\langle f, a_{z_*} \big\rangle, \end{aligned} \tag{225}$$

where

$$a_{z_*}(\cdot) = \frac{b_{z_*}(\cdot)}{p_0(z_*)}, \qquad b_{z_*}(\cdot) = \mathbf{E}\psi^{**}(\cdot, Y_i) p_1\big(z_* - \frac{g(Y_i)}{\sqrt{M}}\big)$$

are non-random elements of L^r. It follows from (223), (224), (225) that

$$m(\cdot) = \frac{M}{\sqrt{N}} a_{z_*}(\cdot) - \frac{\sqrt{M}}{\sqrt{N}} b^*(\cdot)\, \mathbf{E}_* \eta.$$

In order to prove (221) we show that, for $|z_*| \le w_*$,

$$\|b_{z_*}\|_2 \le c_* M^{-1}, \tag{226}$$
$$|\mathbf{E}_* \eta| \le \tilde{c}_* M^{-1/2} + \tilde{c}_* R^{-1/2}, \tag{227}$$
$$p_i(z_*) \ge \tilde{c}_*, \qquad i = 0, 1, 2, \tag{228}$$

and apply (208). Note that, by Lemma 7, there exist positive constants $\tilde{c}_*$, $\tilde{c}'_*$ such that, for $M, N > \tilde{c}'_*$, the inequality (228) holds.

Let us prove (226). In Lemma 7 we show, for $i = 1, 2$, that p_i and its derivatives are bounded functions. That is,

$$|p_i| \le c_*, \qquad |p'_i| \le c_*, \qquad |p''_i| \le c_*, \qquad |p'''_i| \le c_*, \quad i = 1, 2. \tag{229}$$

Expanding in powers of $M^{-1/2} g(Y_i)$ we obtain

$$p_1\Big(z_* - \frac{g(Y_i)}{\sqrt{M}}\Big) = p_1(z_*) - \frac{g(Y_i)}{\sqrt{M}} p'_1(z_*) + \frac{g^2(Y_i)}{M} \frac{p''_1(\theta)}{2}. \tag{230}$$

It follows from the identities (196) and (197) that for P_X almost all x

$$\begin{aligned}
\mathbf{E}\psi^{**}(x, Y_i) &= q_N^{-1} \mathbf{E}\psi^{**}(x, X_i) \mathbb{I}_{A_i} \\
&= q_N^{-1} \mathbf{E}\psi^{**}(x, X_i)(\mathbb{I}_{A_i} - 1) \\
&=: q_N^{-1} a_0(x), \\
\mathbf{E}\psi^{**}(x, Y_i) g(Y_i) &= q_N^{-1} \mathbf{E}\psi^{**}(x, X_i) g(X_i) \mathbb{I}_{A_i} \\
&= q_N^{-1} \mathbf{E}\psi^{**}(x, X_i) g(X_i)(\mathbb{I}_{A_i} - 1) \\
&=: q_N^{-1} a_1(x).
\end{aligned}$$

Using (229) and the inequality $q_N^{-1} \le c_*$, see (44), we obtain from (230)

$$\|b_{z_*}(\cdot)\|_2 \le c_* \|a_0(\cdot)\|_2 + \frac{c_*}{\sqrt{M}} \|a_1(\cdot)\|_2 + \frac{c_*}{M} \|a_2(\cdot)\|_2,$$

where we denote $a_2(\cdot) = \mathbf{E}\psi^{**}(\cdot, Y_i) g^2(Y_i)$. In order to prove (226) we show that

$$\|a_0(\cdot)\|_2 \le \frac{c_*}{N^{\alpha(r-1)}}, \qquad \|a_1(\cdot)\|_2 \le \frac{c_*}{N^{\alpha(r-2)}}, \qquad \|a_2(\cdot)\|_2 \le c_*. \tag{231}$$

Let us prove (231). Invoking (43) we obtain, by Hölder's inequality,

$$|a_0(x)| \le \mathbf{E}|\psi^{**}(x, X_i)| \frac{\|Z_i'\|_r^{r-1}}{N^{\alpha(r-1)}} \le w^{1/r}(x) \frac{K^{(r-1)/r}}{N^{\alpha(r-1)}}, \tag{232}$$

where we denote $w(x) = \mathbf{E}|\psi^{**}(x, X_i)|^r$. Furthermore, by Lyapunov's inequality,

$$\|w^{1/r}(\cdot)\|_2^2 = \mathbf{E}w^{2/r}(X) \le \big(\mathbf{E}w(X)\big)^{2/r} = K_r^{2/r}. \tag{233}$$

Clearly, the first bound of (231) follows from (232), (233) and (200). A similar argument shows the second bound of (231). We have

$$|a_1(x)| \le \mathbf{E}|\psi^{**}(x, X_i)g(X_i)| \frac{\|Z_i'\|_r^{r-2}}{N^{\alpha(r-2)}} \le w^{1/r}(x) \frac{V^{(r-1)/r}}{N^{\alpha(r-2)}}, \tag{234}$$

where we denote $V = \mathbf{E}\big(\|Z_i'\|_r^{r-2}|g(X_i)|\big)^{r/(r-1)}$. By Hölder's inequality,

$$V \le \big(\mathbf{E}|g(X_i)|^r\big)^{1/(r-1)} \big(\mathbf{E}\|Z_i'\|_r^r\big)^{(r-2)/(r-1)} \le c_*. \tag{235}$$

Clearly, (233), (234) and (235) imply the second bound of (231). The last bound of (231) follows from (47), by Cauchy-Shwartz. Indeed, we have

$$|a_2(x)| \le c_*\mathbf{E}|\psi^{**}(x, X_i)|g^2(X_i) \le c_*\big(\mathbf{E}|\psi^{**}(x, X_i)|^2\mathbf{E}g^4(X_i)\big)^{1/2}.$$

Therefore, $\|a_2(\cdot)\|_2^2 \le c_* K_2 \mathbf{E}g^4(X_i) \le c_*$, by (200).

Let us prove (227). We have, by (251),

$$\mathbf{E}_*\eta = p_0^{-1}(z_*)\mathbf{E}(z_* - \zeta)p(z_* - \zeta).$$

In order to prove (227) it suffices to show in view of (228) that

$$|\mathbf{E}(z_* - \zeta)p(z_* - \zeta)| \le c_* R^{-1/2} + c_* M^{-1/2}. \tag{236}$$

Let $\tilde{p}$ denote the density function of ξ_k. Then $p(u) = R\,\tilde{p}(R\,u)$. We have

$$\mathbf{E}(z_* - \zeta)p(z_* - \zeta) = 6c_\xi \mathbf{E}\frac{\sin^6(R(z_* - \zeta)/6)}{\big(R\,(z_* - \zeta)/6\big)^5}.$$

Therefore, denoting $H(z_*) = 1 + |R\,(z_* - \zeta)|^5$, we obtain

$$\mathbf{E}|(z_* - \zeta)p(z_* - \zeta)| \le c\mathbf{E}H^{-1}(z_*). \tag{237}$$

On the event $|\zeta - z_*| \ge R^{-1/2}$ we have $H^{-1}(z_*) \le R^{-5/2}$. Furthermore, a bound for the probability of the complementary event

$$\mathbf{P}\{|\zeta - z_*| \le R^{-1/2}\} \le c_* R^{-1/2} + c_* M^{-1/2},$$

follows by the Berry–Esseen bound applied to the sum ζ. Therefore, $\mathbf{E}H^{-1}(z_*)$ is bounded by the right side of (236). Now (236) follows from (237).

Proof of (222). Write

$$U_k^{**} - \frac{1}{N}\frac{\xi_k}{\sqrt{n}}b^* = \frac{\sqrt{M}}{\sqrt{N}}(T_1 - T_2),$$
$$T_1 := \frac{1}{\sqrt{M}}\sum_{j\in O_k}\psi^{**}(\cdot, Y_j), \qquad T_2 := \eta b^*.$$

It follows from (219), by the inequality $\|u+v\|_2^2 \geq \|u\|_2^2/2 - \|v\|_2^2$, for $u, v \in L^2(P_X)$, that

$$\mathbf{E}\|\hat{\psi}^*\|_2^2 = \frac{M}{N}\mathbf{E}_*\|T_1 - T_2\|_2^2 \geq \frac{M}{2N}\mathbf{E}_*\|T_1\|_2^2 - \frac{M}{N}\mathbf{E}_*\|T_2\|_2^2.$$

We shall show that

$$\mathbf{E}_*\|T_2\|_2^2 \leq p_0^{-1}(z_*)\big(c_* R^{-1}M^{-1/2} + c_* R^{-3/2}\big), \tag{238}$$
$$\mathbf{E}_*\|T_1\|_2^2 \geq p_0^{-1}(z_*)\big(p_1(z_*)\delta_3^2 - c_*\tau_N\big), \tag{239}$$
$$\mathbf{E}_*\|T_1\|_2^2 \leq p_0^{-1}(z_*)\big(p_1(z_*)\delta_3^2 + c_*\tau_N\big). \tag{240}$$

The inequalities (238) and (239) imply the lower bound in (222). Indeed, by (228), we have, for small $c_*^{(0)}$,

$$\frac{c_*}{M^{1/2}R} + \frac{c_*}{R^{3/2}} \leq c_*\tau_N \leq c_* c_*^{(0)}\delta_3^2 \leq p_1(z_*)\delta_3^2/4.$$

Similarly, the inequalities (238) and (240) imply the upper bound in (222).

Proof of (238). We have, by (251),

$$\mathbf{E}_*\eta^2 = p_0^{-1}(z_*)W, \qquad W := \mathbf{E}(z_* - \zeta)^2 p(z_* - \zeta).$$

Proceeding as in proof of (236), we obtain

$$W = \frac{36}{R}c_\xi \mathbf{E}\frac{\sin^6(R(z_* - \zeta)/6)}{(R(z_* - \zeta)/6)^4} \leq \frac{c}{R}\mathbf{E}\tilde{H}^{-1}(z_*),$$

where $\tilde{H}(z_*) = 1 + |R(z_* - \zeta)|^4$ satisfies

$$\mathbf{E}\tilde{H}^{-1}(z_*) \leq c_* R^{-1/2} + c_* M^{-1/2}.$$

Therefore, $W \leq c_* R^{-3/2} + c_* R^{-1}M^{-1/2}$. This inequality in combination with (208) implies (238).

Proof of (239). Fix $i, j \in O_k, i \neq j$. By symmetry,

$$\mathbf{E}_* \|T_1\|_2^2 = \mathbf{E}_* T_{11} + (M-1)\mathbf{E}_* T_{12},$$
$$T_{11} = \|\psi^{**}(\cdot, Y_i)\|_2^2, \qquad T_{12} = \big\langle \psi^{**}(\cdot, Y_i),\ \psi^{**}(\cdot, Y_j)\big\rangle. \tag{241}$$

We have, by (252),

$$\mathbf{E}_* T_{11} = p_0^{-1}(z_*) H_1, \qquad \mathbf{E}_* T_{12} = p_0^{-1}(z_*) H_2,$$
$$H_1 = \mathbf{E} T_{11} p_1\big(z_* - \frac{g(Y_i)}{\sqrt{M}}\big), \qquad H_2 = \mathbf{E} T_{12} p_2\big(z_* - \frac{g(Y_i) + g(Y_j)}{\sqrt{M}}\big).$$

The inequality (239) follows from (241) and the bounds

$$H_1 \geq p_1(z_*)\delta_3^2 - c_* M^{-1/2}, \tag{242}$$
$$|H_2| \leq c_* N^{-\alpha(r-2)} + c_* M^{-(r-2)/2}. \tag{243}$$

Let us prove (242). It follows from (229), by the mean value theorem, that

$$H_1 = p_1(z_*)\mathbf{E} T_{11} + Q, \qquad |Q| \leq c_* \mathbf{E} T_{11} \frac{|g(Y_i)|}{\sqrt{M}}, \tag{244}$$

where $|Q| \leq c_* M^{-1/2}$. Indeed, by (47) and Cauchy–Schwartz,

$$\mathbf{E} T_{11} |g(Y_i)| \leq \mathbf{E} \|\psi^{**}(\cdot, X_i)\|_2^2 |g(X_i)| \leq K_4^{1/2} \sigma \leq c_*.$$

In the last step we applied (200). Furthermore, the identity

$$\mathbf{E} T_{11} = q_N^{-1} \mathbf{E} |\psi^{**}(X, X_i)|^2 \mathbb{I}_{A_i} = \mathbf{E} |\psi^{**}(X, X_i)|^2 - b_1 - b_2,$$
$$b_1 = (1 - q_N^{-1}) \mathbf{E} |\psi^{**}(X, X_i)|^2, \qquad b_2 = q_N^{-1} \mathbf{E} |\psi^{**}(X, X_i)|^2 (1 - \mathbb{I}_{A_i})$$

combined with (43), (44) and (45) yields $\mathbf{E} T_{11} \geq \delta_3^2 - c_* M^{-1/2}$. This bound together with (244) shows (242).

Let us prove (243). Write $y_i = g(Y_i)$ and expand

$$p_2\big(z_* - \frac{y_i + y_j}{\sqrt{M}}\big) = p_2(z_*) - p_2'(z_*) \frac{y_i + y_j}{\sqrt{M}} + \frac{p_2''(z_*)}{2} \frac{(y_i + y_j)^2}{M} + \tilde{Q},$$

where $\tilde{Q}$ denotes the remainder term. From (229) it follows, for $2 < r - 2 \leq 3$ that

$$|\tilde{Q}| \leq c_* |y_i + y_j|^{r-2} / M^{(r-2)/2}.$$

Furthermore, denote

$$h_1 = \mathbf{E}T_{12}, \qquad h_2 = \mathbf{E}T_{12}g(Y_i),$$
$$h_3 = \mathbf{E}T_{12}g^2(Y_i), \qquad h_4 = \mathbf{E}T_{12}g(Y_i)g(Y_j).$$

We obtain, by symmetry,

$$H_2 = p_2(z_*)h_1 - 2\frac{p_2'(z_*)}{\sqrt{M}}h_2 + \frac{p_2''(z_*)}{M}(h_3 + h_4) + \mathbf{E}T_{12}\tilde{Q},$$
$$\mathbf{E}|T_{12}\tilde{Q}| \le c_* M^{-(r-2)/2}\mathbf{E}|g(Y_i)|^{r-2}|T_{12}|. \tag{245}$$

Denote

$$\tilde{T}_{12} = q_N^{-2}\psi^{**}(X, X_i)\psi^{**}(X, X_j).$$

It follows from (45), by Hölders inequality and (200), that

$$\mathbf{E}|g(Y_i)|^{r-2}|T_{12}| \le \mathbf{E}|g(X_i)|^{r-2}|\tilde{T}_{12}| \le c_*.$$

Therefore,

$$\mathbf{E}|T_{12}\tilde{Q}| \le c_* M^{-(r-2)/2}. \tag{246}$$

Furthermore, (196) and (197) imply

$$h_1 = \mathbf{E}\tilde{T}_{12}(\mathbb{I}_{A_i} - 1)(\mathbb{I}_{A_j} - 1), \qquad h_2 = \mathbf{E}\tilde{T}_{12}g(X_i)(\mathbb{I}_{A_i} - 1)(\mathbb{I}_{A_j} - 1),$$
$$h_3 = \mathbf{E}\tilde{T}_{12}g^2(X_i)\mathbb{I}_{A_i}(\mathbb{I}_{A_j} - 1), \qquad h_4 = \mathbf{E}\tilde{T}_{12}g(X_i)g(X_j)(\mathbb{I}_{A_i} - 1)(\mathbb{I}_{A_j} - 1).$$

Invoking the inequalities $q_N^{-2} \le c_*$, see (44), and $1 - \mathbb{I}_{A_i} \le V_i^s$, $s > 0$, where $V_i := \|Z_i'\|_r / N^\alpha$, see (43), we obtain, by Hölder's inequality,

$$|h_1| \le \mathbf{E}|\tilde{T}_{12}|(V_i V_j)^{(r-2)/2} \le c_* N^{-\alpha(r-2)},$$
$$|h_2| \le \mathbf{E}|\tilde{T}_{12}g(X_i)|V_i^{(r-4)/2}V_j^{(r-2)/2} \le c_* N^{-\alpha(r-3)},$$
$$|h_3| \le \mathbf{E}|\tilde{T}_{12}|g^2(X_i)V_j^{r-4} \le c_* N^{-\alpha(r-4)},$$
$$|h_4| \le \mathbf{E}|\tilde{T}_{12}g(X_i)g(X_j)|(V_i V_j)^{(r-4)/2} \le c_* N^{-\alpha(r-4)}. \tag{247}$$

Combining (245), (247), (246) and using the simple inequalities

$$\frac{1}{N^{\alpha(r-3)}M^{1/2}} \le \frac{1}{\tilde{N}}, \qquad \frac{1}{N^{\alpha(r-4)}M} \le \frac{1}{\tilde{N}}, \qquad \tilde{N} = \min\{N^{\alpha(r-2)}, M^{(r-2)/2}\}$$

and the inequalities (229), we obtain (243).

Proof of (240). The inequality follows from (241), (243) and the inequality

$$H_1 \le p_1(z_*)\mathbf{E}T_{11} + c_* M^{-1/2} \le p_1(z_*)\delta_3^2 + c_* N^{-\alpha r} + c_* M^{-1/2}$$
$$\le p_1(z_*)\delta_3^2 + c_* M^{-1/2},$$

which is obtained in the same way as (242) above.

Proof of (218). In order to prove (218) we shall show that

$$\mathbf{E}\|\psi_k\|_r^r \le \tilde{c}_* n^{-r/2}. \tag{248}$$

Split $O_k = B \cup D$, where $B \cap D = \emptyset$ and $|B| = [M/2]$ and write

$$U_k = \frac{\sqrt{M}}{\sqrt{N}}(U_B + U_D), \qquad U_B = \frac{1}{\sqrt{M}} \sum_{j \in B} \psi(\cdot, Y_j),$$
$$\zeta = \zeta_B + \zeta_D, \qquad \zeta_B = \frac{1}{\sqrt{M}} \sum_{j \in B} g(Y_j).$$

In particular, we have $g_* = \eta + \zeta_B + \zeta_D$.

The inequality

$$\mathbf{E}\|\psi_k\|_r^r = \mathbf{E}_*\|U_k\|_r^r \le 2^r \Big(\frac{M}{N}\Big)^{r/2} \big(\mathbf{E}_*\|U_B\|_r^r + \mathbf{E}_*\|U_D\|_r^r\big)$$

combined with the bounds

$$\mathbf{E}_*\|U_B\|_r^r \le c_*, \qquad \mathbf{E}_*\|U_D\|_r^r \le c_* \tag{249}$$

imply (248). Let us prove the first bound of (249). By (252), we have

$$\mathbf{E}_*\|U_B\|_r^r = p_0^{-1}(z_*)\mathbf{E}\|U_B\|_r^r p_3(z_* - \zeta_B),$$

where p_3 denotes the density of $\eta + \zeta_D$. Furthermore, invoking the bound $\sup_{x\in\mathbb{R}} |p_3(x)| \le c_*$, (which is obtained using the same argument as in the proof of Lemma 7) and the inequality (228), we obtain $\mathbf{E}_*\|U_B\|_r^r \le \tilde{c}_*\mathbf{E}\|U_B\|_r^r$. Finally, invoking the bound

$$\mathbf{E}\|U_B\|_r^r \le c_*\mathbf{E}\Big\|\frac{1}{\sqrt{M}} \sum_{j \in B} \psi(\cdot, X_j)\Big\|_r^r \le c_*, \tag{250}$$

see (47) and (214), we obtain the first bound of (249). The second bound is obtained in the same way. This completes the proof of the lemma. □

We collect some facts about conditional moments in a separate lemma.

Lemma 11 *Let η and ζ be independent random variables. Assume that η is real valued and has a density, say $x \to p(x)$.*
(i) *Assume that ζ is real valued. Then the function*

$$x \to \mathbf{E}p(x-\zeta), \qquad x \in \mathbb{R},$$

is a density of the distribution $P_{\eta+\zeta}$ of $\eta+\zeta$. Let $w : \mathbb{R} \to \mathbb{R}$ be a measurable function such that $\mathbf{E}|w(\eta)| < \infty$. For $P_{\eta+\zeta}$ almost all $x \in \mathbb{R}$, we have

$$\mathbf{E}\big(w(\eta)\,\big|\,\eta+\zeta=x\big) = \frac{\mathbf{E}w(x-\zeta)p(x-\zeta)}{\mathbf{E}p(x-\zeta)}. \tag{251}$$

(ii) *Assume that ζ takes values in a measurable space, say $\mathcal{Y}$. Assume that $u, v : \mathcal{Y} \to \mathbb{R}$ are measurable functions and denote $P_{\eta+u(\zeta)}$ the distribution of $\eta+u(\zeta)$. If $\mathbf{E}|v(\zeta)| < \infty$, then for $P_{\eta+u(\zeta)}$ almost all $x \in \mathbb{R}$,*

$$\mathbf{E}\big(v(\zeta)\,\big|\,\eta+u(\zeta)=x\big) = \frac{\mathbf{E}v(\zeta)p\big(x-u(\zeta)\big)}{\mathbf{E}p\big(x-u(\zeta)\big)}. \tag{252}$$

9 Appendix 4

In the next lemma we consider independent and identically distributed random vectors $(\xi,\ \eta)$ and $(\xi',\ \eta')$ with values in $\mathbb{R}^2$ and the symmetrization (ξ_s, η_s) where $\xi_s = \xi - \xi'$ and $\eta_s = \eta - \eta'$. Note that in the main text we apply this lemma to $\xi = g(X_1)$ and $\eta = N^{-1/2}\sum_{j=m+1}^{N}\psi(X_1, Y_j)$.

Lemma 12 *Let $0 < \nu < 1/2$ and $r > 2$. Assume that $\mathbf{E}|\xi|^r + \mathbf{E}|\eta|^r < \infty$. The following statements hold.*
(a) *For $c_r = (7/12)2^{-r}$ the conditions*

$$|t|^{r-2}\mathbf{E}|\xi_s|^r \le c_r\mathbf{E}\xi_s^2, \qquad \mathbf{E}\xi_s\eta_s = 0, \qquad \mathbf{E}|\eta_s|^r \le c_r\mathbf{E}\eta_s^2$$

imply $1 - |\mathbf{E}\exp\{i(t\xi+\eta)\}|^2 \ge 6^{-1}(t^2\mathbf{E}\xi_s^2 + \mathbf{E}\eta_s^2)$.
(b) *Assume that for some $\tilde{c}_1, \tilde{c}_2 > 0$ we have*

$$\mathbf{E}\xi_s^2/12 - N^{-1}\mathbf{E}\eta_s^2 > \tilde{c}_1^2 \quad \text{and} \quad c_r\mathbf{E}\xi_s^2/\mathbf{E}|\xi_s|^r \ge \tilde{c}_2^{r-2}. \tag{253}$$

Let $\varepsilon > 0$ be such that

$$\varepsilon < 1/6\tilde{c}_3, \qquad \varepsilon^{(r-2)/2} < \sigma_z^2/4A, \qquad \varepsilon^{r-2} < \sigma_z^2/4B, \tag{254}$$

where $\tilde{c}_3 = 2 + (5/\tilde{c}_1)^2\sigma_z^2$ and where the numbers A, B are defined in (265). Here $\sigma_z^2 = \mathbf{E}(\xi_s + N^{-1/2}\eta_s)^2$. Assume that for some $0 < \delta < \tilde{c}_2$ and $\delta' > 10\varepsilon^2$,

$$\sup_{\delta<|t|<N^{-\nu+1/2}} |\mathbf{E}e^{it\xi_s}| \le 1-\delta' \qquad \text{and} \qquad \mathbf{E}|\eta_s| \le \delta' N^{\nu}/2. \tag{255}$$

Then for every T^*, satisfying $N^{1/2-\nu} \le |T^*| \le N^{\nu+1/2}$, the set

$$I^* = \{T^* \le t \le T^* + N^{1/2-\nu} : |\mathbf{E}e^{it(\xi+N^{-1/2}\eta)}|^2 \ge 1-\varepsilon^2\}$$

is an interval of size at most $5\tilde{c}_1^{-1}\varepsilon$.

Proof Proof of (**a**). Invoking the inequality $1-\cos x \ge x^2/2 - x^2/24 - |x|^r$ and using the simple inequality $|a+b|^r \le 2^{r-1}(|a|^r + |b|^r)$ we obtain

$$\begin{aligned}1 - |\mathbf{E}\exp\{i(t\xi+\eta)\}|^2 &= 1 - \mathbf{E}\cos(t\xi_s + \eta_s)\\ &\ge \frac{11}{24}\mathbf{E}(t\xi_s+\eta_s)^2 - 2^{r-1}(\mathbf{E}|t\xi_s|^r + \mathbf{E}|\eta_s|^r)\\ &\ge 6^{-1}(t^2\mathbf{E}\xi_s^2 + \mathbf{E}\eta_s^2).\end{aligned}$$

In the last step we use the conditions **a)**.

Proof of (**b**). Introduce the function $t \to \tau_t^* = 1 - |\mathbf{E}e^{it(\xi+N^{-1/2}\eta)}|^2$. Assume that the set I^* is non-empty and choose $s, t \in I^*$, i.e., we have $\tau_t^*, \tau_s^* \le \varepsilon^2$. Firstly we show that $|s-t| \le 5\tilde{c}_1^{-1}\varepsilon$, thus proving the bound for the size of the set I^*.

The inequality $1-\cos(x+y) \ge (1-\cos x)/2 - (1-\cos y)$ implies

$$1 - |\mathbf{E}e^{i(X+Y)}|^2 \ge 2^{-1}(1 - |\mathbf{E}e^{iX}|^2) - (1 - |\mathbf{E}e^{iY}|^2), \tag{256}$$

for arbitrary random variables X, Y. Choosing $\tilde{Y} = t(\xi + N^{-1/2}\eta)$ and $\tilde{X} = (s-t)(\xi + N^{-1/2}\eta)$ shows

$$\tau_s^* \ge (1 - |\mathbf{E}e^{i\tilde{X}}|^2)/2 - \tau_t^*. \tag{257}$$

Now we show that the inequality $|t-s| > 5\tilde{c}_1^{-1}\varepsilon$ implies $1 - |\mathbf{E}e^{i\tilde{X}}|^2 > 5\varepsilon^2$, thus, contradicting to our choice $\tau_s^*, \tau_t^* < \varepsilon^2$ and (257). In what follows the cases of "large" and "small" values of $|t-s|$ are treated separately.

For $5\tilde{c}_1^{-1}\varepsilon < |t-s| \le \delta$ we shall apply (256) to $\tilde{X} = X + Y$, where $X = (s-t)\xi$ and $Y = (s-t)N^{-1/2}\eta$. Note that the statement **a)** implies

$$1 - |\mathbf{E}e^{iX}|^2 \ge (t-s)^2\mathbf{E}\xi_s^2/6. \tag{258}$$

Indeed, in view of the second inequality of (253), the conditions of **a)** are satisfied for $|t-s| \le \delta \le \tilde{c}_2$. Furthermore, we have

$$0 \le 1 - |\mathbf{E}e^{iY}|^2 = 1 - \cos\big(N^{-1/2}(s-t)\eta_s\big) \le (s-t)^2N^{-1}\mathbf{E}\eta_s^2. \tag{259}$$

Invoking the bounds (258) and (259) in (256) we obtain

$$1 - |\mathbf{E}e^{i\tilde{X}}|^2 \ge (s-t)^2\mathbf{E}\xi_s^2/12 - (s-t)^2N^{-1}\mathbf{E}\eta_s^2 \ge \tilde{c}_1^2(s-t)^2 \ge 25\varepsilon^2.$$

In the last step we used (253).

For $\delta < |t-s| \le N^{-\nu+1/2}$ we expand in powers of $a = i(s-t)N^{-1/2}\eta_s$ to get

$$\begin{aligned}1 - |\mathbf{E}e^{i\tilde{X}}|^2 &= 1 - \mathbf{E}\exp\{i(s-t)\xi_s + a\} \ge 1 - \mathbf{E}\exp\{i(s-t)\xi_s\} - \mathbf{E}|a| \\ &\ge \delta' - \mathbf{E}|(t-s)N^{-1/2}\eta_s| \ge \delta' - N^{-\nu}\mathbf{E}|\eta_s| \\ &\ge \delta'/2 \ge 5\varepsilon^2.\end{aligned}$$

In the last step we applied (255).

Let us prove that I^* is indeed an interval. Assume the contrary, i.e. there exist $s < u < t$ such that $s, t \in I^*$ and $u \notin I^*$. In particular, $\tau_t^* \le \varepsilon^2 < \tau_u^*$. Clearly, we can choose u to be a local maximum (stationary) point of the function $t \to \tau_t^*$. Denote

$$z = \xi_s + N^{-1/2}\eta_s, \qquad \sigma_z^2 = \mathbf{E}z^2.$$

An application of (256) to $Y' = (t-u)(\xi + N^{-1/2}\eta)$ and $X' = u(\xi + N^{-1/2}\eta)$ gives

$$\tau_t^* \ge \tau_u^*/2 - \big(1 - \mathbf{E}e^{i(t-u)z}\big) = \tau_u^*/2 - \big(1 - \mathbf{E}\cos(t-u)z\big).$$

Invoking the inequalities $\tau_t^* \le \varepsilon^2$ and $1 - \cos(t-u)z \le (t-u)^2 z^2/2$ we obtain

$$\tau_u^* \le 2\varepsilon^2 + (t-u)^2\sigma_z^2 \le \varepsilon^2\tilde{c}_3, \qquad \tilde{c}_3 = 2 + (5/\tilde{c}_1)^2\sigma_z^2. \tag{260}$$

Here we used the bound $|t-u| \le |t-s| \le 5\varepsilon/\tilde{c}_1$ proved above.

Denoting $y = (t-u)z$ we have $\tau_t^* = 1 - \mathbf{E}e^{iuz}e^{iy}$. Invoking the expansion $e^{iy} = 1 + iy + (iy)^2/2 + R'$, where $|R'| \le y^2/6 + |y|^r$, we obtain

$$\tau_t^* = \tau_u^* - i\mathbf{E}ye^{iuz} + 2^{-1}\mathbf{E}y^2e^{iuz} + R, \qquad |R| \le \mathbf{E}y^2/6 + \mathbf{E}|y|^r =: R_0. \tag{261}$$

For a stationary point u we have $0 = \frac{\partial}{\partial t}\tau_t^*\big|_{t=u} = -i\mathbf{E}ze^{iuz}$. Therefore, $\mathbf{E}ye^{iuz} = 0$ and (261) implies

$$\tau_t^* \ge \tau_u^* + 2^{-1}(t-u)^2\mathbf{E}z^2e^{iuz} - R_0.$$

Write the right hand side in the form $\tau_u^* + 2^{-1}(t-u)^2R_1$, where

$$R_1 = \mathbf{E}z^2e^{iuz} - 3^{-1}\sigma_z^2 - 2\mathbf{E}|z|^r|t-u|^{r-2}.$$

Note that the inequality $R_1 > 0$ contradicts to our assumption $\tau_t^* < \tau_u^*$. We complete the proof by showing that $R_1 > 0$.

Since the random variable z is symmetric we have $\mathbf{E}z^2\sin uz = 0$. Therefore,

$$\mathbf{E}z^2e^{iuz} = \mathbf{E}z^2\cos uz = \sigma_z^2 - \mathbf{E}z^2(1 - \cos uz). \tag{262}$$

Given $\lambda > 0$ split

$$\mathbf{E}z^2(1-\cos uz) = \mathbf{E}z^2(1-\cos uz)\Big(\mathbb{I}_{\{z^2<\lambda^2\}} + \mathbb{I}_{\{z^2\geq\lambda^2\}}\Big)$$
$$\leq \lambda^2\mathbf{E}(1-\cos uz) + 2\mathbf{E}|z|^r\lambda^{2-r}. \tag{263}$$

In the last step we used Chebyshev's inequality. Furthermore, invoking the inequality $\mathbf{E}(1-\cos uz) = \tau_u^* \leq \tilde{c}_3\varepsilon^2$, see (260), we obtain from (262) and (263) for $\lambda^2 = \varepsilon^{-1}\sigma_z^2$

$$\mathbf{E}z^2e^{iuz} \geq \sigma_z^2 - \varepsilon\tilde{c}_3\sigma_z^2 - \varepsilon^{(r-2)/2}2\mathbf{E}|z|^r\sigma_z^{2-r}. \tag{264}$$

Finally, invoking the inequality $|t-u| \leq |t-s| \leq 5\tilde{c}_1^{-1}\varepsilon$ we obtain from (264)

$$R_1 \geq \sigma_z^2(1-3^{-1}-\varepsilon\tilde{c}_3) - \varepsilon^{(r-2)/2}A - \varepsilon^{r-2}B,$$

where for random variable $z = \xi_s + N^{-1/2}\eta_s$ we write

$$A = 2\mathbf{E}|z|^r\sigma_z^{2-r} \qquad \text{and} \qquad B = 2\mathbf{E}|z|^r(5/\tilde{c}_1)^{r-2}. \tag{265}$$

Thus, for ε satisfying (254) we have $R_1 > 0$. □

10 Appendix 5

Let $Z_1,\ldots,Z_N$ be independent copies of the L^r valued random element $Z = \{x \to \psi(x,Y)\}$. Recall that almost surely $\|Z\| \leq N^\alpha$. Here $\|\cdot\|$ denotes the norm of the Banach space L^r, where $r > 4$ and $1/2 > \alpha > 0$. Write $M_p = \mathbf{E}|\psi(X_1,X_2)|^p$.

Lemma 13 (i) *Assume that $\|\mathbf{E}Z\|^2 \leq \mathbf{E}\|Z\|^2/N$. Then there exists a constant $c(r) > 0$ such that for $k \leq N$ and $x > c(r)$ we have*

$$\mathbf{P}\{\|Z_1+\cdots+Z_k\| > k^{1/2}u\,x\} \leq \exp\{-2^{-5}x^2(1+xN^\alpha/k^{1/2}u)^{-1}\}. \tag{266}$$

Here $u^2 = \mathbf{E}\|Z\|^2$.

(ii) *The following inequalities hold*

$$\|\mathbf{E}Z\| \leq M_r/q_NN^{(r-1)\alpha}, \tag{267}$$
$$q_N^{-1}(M_2 - M_rN^{-(r-2)\alpha}) \leq \mathbf{E}\|Z\|^2 \leq q_N^{-1}(M_r^{2/r} + M_rN^{-(r-2)\alpha}). \tag{268}$$

Remark. Assume that

$$M_2 \geq 2M_rN^{-(r-2)\alpha}, \qquad M_r^2 \leq (q_N/2)M_2N^{\varkappa}, \qquad \text{where} \qquad \varkappa = 2(r-1)\alpha - 1.$$

Then (267) and (268) imply the inequality $\|\mathbf{E}Z\|^2 \leq \mathbf{E}\|Z\|^2/N$. Note that $r\alpha > 2$ implies $\varkappa > 2$. Furthermore, by (44), the probability $q_N > 1 - M_rN^{-r\alpha}$.

Proof We derive (i) from Yurinskii's [36] inequality. Denote $\zeta_k = Z_1 + \cdots + Z_k$. Using the type$-2$ inequality for an L^r valued random variable $\zeta_k - \mathbf{E}\zeta_k$,

$$\mathbf{E}\|\zeta_k - \mathbf{E}\zeta_k\|^2 \leq k\tilde{c}(r)\mathbf{E}\|Z_1 - \mathbf{E}Z_1\|^2,$$

and the inequality $\|Z_1 - \mathbf{E}Z_1\|^2 \leq 2\|Z_1\|^2 + 2\|\mathbf{E}Z_1\|^2$, we obtain

$$\mathbf{E}\|\zeta_k - \mathbf{E}\zeta_k\| \leq \big(\mathbf{E}\|\zeta_k - \mathbf{E}\zeta_k\|^2\big)^{1/2} \leq k^{1/2}c'(r)(u + \|\mathbf{E}Z_1\|).$$

We have

$$\begin{aligned}\mathbf{E}\|\zeta_k\| &\leq \mathbf{E}\|\zeta_k - \mathbf{E}\zeta_k\| + k\|\mathbf{E}Z_1\| \\ &\leq c'(r)k^{1/2}u + k(1 + c'(r)k^{-1/2})\|\mathbf{E}Z_1\| =: \beta_k.\end{aligned}$$

It follows from the inequality $\|Z_1\| \leq N^\alpha$ that

$$\mathbf{E}\|Z_1\|^L \leq 2^{-1}L!u^2N^{\alpha(L-2)}, \qquad L = 2, 3, \ldots.$$

Write $B_k^2 = ku^2$. Theorem 2.1 of Yurinskii [36] shows

$$\mathbf{P}\{\|\zeta_k\| \geq xB_k\} \leq \exp\{-B\}, \quad B = \frac{\overline{x}^2}{8}(1 + (\overline{x}N^\alpha/2B_k))^{-1}, \tag{269}$$

provided that $\overline{x} = x - \beta_k/B_k > 0$.

Since $\beta_k/B_k \leq 1 + c'(r)(1 + k^{-1/2})$ we have, for $x > c(r) := 4c'(r) + 2$,

$$x > 2\beta_k/B_k \qquad \text{and} \qquad x > \overline{x} > x/2.$$

The latter inequality implies

$$B \geq B' := \frac{(x/2)^2}{8}(1 + (xN^\alpha/B_k))^{-1}.$$

Finally, replacing B by B' in (269) we obtain (266).

Let us prove (ii). The mean value $\mathbf{E}Z = \{x \to \mathbf{E}\psi(x, Y)\}$ is an element of L^r. For P_X almost all $x \in \mathcal{X}$ we have $\mathbf{E}\psi(x, X) = 0$. Therefore,

$$\mathbf{E}Z = q_N^{-1}\mathbf{E}\psi(x, X)\mathbb{I}_A = q_N^{-1}\mathbf{E}\psi(x, X)(\mathbb{I}_A - 1).$$

Invoking (43) and using Chebyshev and Hölder inequalities, we obtain, for P_X almost all x,

$$|\mathbf{E}Z| \leq \frac{1}{q_N N^{\alpha(r-1)}}\mathbf{E}\|Z'\|_r^{r-1}|\psi(x, X)| \leq \frac{1}{q_N N^{\alpha(r-1)}}(\mathbf{E}\|Z'\|_r^r)^{(r-1)/r}a(x),$$

where $a(x) = (\mathbf{E}|\psi(x, X)|^r)^{1/r}$. Note that $\mathbf{E}\|Z'\|_r^r = M_r$ and $\|a\|^r = M_r$. Finally,

$$\|\mathbf{E}Z\| \le \|a\| M_r^{(r-1)/r} / q_N N^{\alpha(r-1)} = M_r / q_N N^{\alpha(r-1)}.$$

Let us prove (268). Denote $b_p(x) = (\mathbf{E}_{X_1}|\psi(X_1, x)|^p)^{1/p}$. Here $\mathbf{E}_{X_1}$ denotes the conditional expectation given all the random variables, but X_1. We have

$$\mathbf{E}\|Z\|^2 = q_N^{-1}\mathbf{E}\mathbb{I}_A b_r^2(X) = q_N^{-1}\mathbf{E}b_r^2(X) + q_N^{-1}R, \qquad R = \mathbf{E}(\mathbb{I}_A - 1)b_r^2(X). \quad (270)$$

By Hölder's inequality $b_r(x) \ge b_2(x)$, for P_X almost all x. Therefore,

$$M_2 = \mathbf{E}b_2^2(X) \le \mathbf{E}b_r^2(X) \le M_r^{2/r}. \quad (271)$$

Combining (271) and (270) and the bound $|R| \le M_r N^{-(r-2)\alpha}$ we obtain (268). In order to bound $|R|$ we use (43), $|R| \le N^{-(r-2)\alpha}\mathbf{E}\|Z'\|_r^{r-2} b_r^2(X)$, and apply Hölder's inequality,

$$\mathbf{E}\|Z'\|_r^{r-2} b_r^2(X) \le (\mathbf{E}\|Z'\|_r^r)^{(r-2)/r}(\mathbf{E}b_r^r(X))^{2/r} = M_r.$$

The lemma is proved. □

References

1. Angst, J., Poly, G.: A weak Cramér condition and application to Edgeworth expansions. Electron. J. Probab. **22**(59), 1–24 (2017)
2. Babu, G.J., Bai, Z.D.: Edgeworth expansions of a function of sample means under minimal moment conditions and partial Cramer's condition. Sankhya Ser. A **55**, 244–258 (1993)
3. Bai, Z.D., Rao, C.R.: Edgeworth expansion of a function of sample means. Ann. Stat. **19**, 1295–1315 (1991)
4. Bentkus, V., Götze, F., van Zwet, W.R.: An Edgeworth expansion for symmetric statistics. Ann. Stat. **25**, 851–896 (1997)
5. Bentkus, V., Götze, F.: Lattice point problems and distribution of values of quadratic forms. Ann. Math. (2) **150**(3), 977–1027 (1999)
6. Bentkus, V., Götze, F.: Optimal bounds in non-Gaussian limit theorems for U-statistics. Ann. Probab. **27**, 454–521 (1999)
7. Bhattacharya, R.N., Rao, R.R.: Normal Approximation and Asymptotic Expansions. Robert E. Krieger Publishing Company, Inc., Malabar (1986)
8. Bhattacharya, R.N., Ghosh, J.K.: On the validity of the formal Edgeworth expansion. Ann. Stat. **6**, 434–451 (1978)
9. Bhattacharya, R.N., Ghosh, J.K.: Correction to: On the validity of the formal Edgeworth expansion. Ann. Stat. **8**, 1399 (1980)
10. Bickel, P.J.: Edgeworth expansions in nonparametric statistics. Ann. Stat. **2**, 1–20 (1974)
11. Bickel, P.J., Götze, F., van Zwet, W.R.: The Edgeworth expansion for U-statistics of degree two. Ann. Stat. **14**, 1463–1484 (1986)
12. Bickel, P., et al.: Willem van Zwet's research. Ann. Stat. **49**, 2439–2447 (2021)
13. Bickel, P.J., Robinson, J.: Edgeworth expansions and smoothness. Ann. Probab. **10**, 500–503 (1982)
14. Bobkov, S.G.: Khinchine's theorem and Edgeworth approximations for weighted sums. Ann. Stat. **47**, 1616–1633 (2019)
15. Bollobás, B.: Combinatorics. Set Systems, Hypergraphs, Families of Vectors and Combinatorial Probability. Cambridge University Press, Cambridge (1986)

16. Callaert, H., Janssen, P., Veraverbeke, N.: An Edgeworth expansion for U-statistics. Ann. Stat. **8**, 299–312 (1980)
17. Chibisov, D.M.: Asymptotic expansion for the distribution of a statistic admitting a stochastic expansion. I. Teor. Veroyatn. Primen. **25**, 745–757 (1980)
18. Chung, K.-L.: The approximate distribution of Student's statistic. Ann. Math. Stat. **17**, 447–465 (1946)
19. Cramér, H.: Random variables and probability distributions. In: Cambridge Tracts in Mathematics and Mathematical Physics, No. 36, 3rd edn (1970). Cambridge University Press (1937)
20. Dharmadhikari, S.W., Fabian, V., Jogdeo, K.: Bounds on the moments of martingales. Ann. Math. Stat. **39**, 1719–1723 (1968)
21. Efron, B., Stein, C.: The jackknife estimate of variance. Ann. Stat. **9**, 586–596 (1981)
22. Esseen, C.G.: Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law. Acta Math. **77**, 1–125 (1945)
23. Götze, F.: Asymptotic expansions for bivariate von Mises functionals. Z. Wahrsch. Verw. Gebiete **50**, 333–355 (1979)
24. Götze, F.: Lattice point problems and values of quadratic forms. Invent. Math. **157**, 195–226 (2004)
25. Götze, F., van Zwet W.R.: Edgeworth expansions for asymptotically linear statistics. Manuscript 1–45 (1992)
26. Götze, F., van Zwet, W.R.: An Expansion for a Discrete Non-lattice Distribution. Frontiers in Statistics, pp. 257–274. Imperial College, London (2006)
27. Götze, F., Zaitsev, A.: Explicit rates of approximation in the CLT for quadratic forms. Ann. Probab. **42**, 354–397 (2014)
28. Hall, P.: Edgeworth expansion for Student's t statistic under minimal moment conditions. Ann. Probab. **15**, 920–931 (1987)
29. Helmers, R.: Edgeworth Expansions for Linear Combinations of Order Statistics. Mathematical Centre Tracts, vol. 105. CWI, Amsterdam (1982)
30. Hodges, J.L., Jr., Lehmann, E.L.: Deficiency. Ann. Math. Stat. **41**, 783–801 (1970)
31. Hoeffding, W.: A class of statistics with asymptotically normal distribution. Ann. Math. Stat. **19**, 293–325 (1948)
32. Ledoux, M., Talagrand, M.: Probability in Banach Spaces. Isoperimetry and Processes. Springer, Berlin (1991)
33. Petrov, V.V.: Sums of Independent Random Variables. Springer, New York (1975)
34. Pfanzagl, J.: Asymptotic expansions for general statistical models. With the assistance of W. Wefelmeyer. In: Lecture Notes in Statistics, vol. 31. Springer, Berlin (1985)
35. Serfling, R.J.: Approximation Theorems of Mathematical Statistics. Wiley, New York (1980)
36. Yurinskii, V.V.: Exponential inequalities for sums of random vectors. J. Multivar. Anal. **6**, 473–499 (1976)
37. van Zwet, W.R.: A Berry–Esseen bound for symmetric statistics. Z. Wahrsch. Verw. Gebiete **66**, 425–440 (1984)