Two sample rank tests with adaptive score functions using kernel density estimation

Dissertation

zur Erlangung des Doktorgrades
an den Naturwissenschaftlichen Fachbereichen
(Mathematik)
der Justus-Liebig-Universität Gießen

vorgelegt von

Brandon Greene

betreut von

Prof. Dr. Winfried Stute

Marburg, Mai 2017

Acknowledgments

My sincerest gratitude and appreciation goes above all to my *Doktorvater* Professor Winfried Stute for all of his excellent advice and guidance in this and other projects, his unquestioning generosity with his time, his patience, and his willingness to give me a chance to begin studying mathematics in Gießen many years ago. I would also like to thank all of my academic teachers from my time in Gießen, particularly Professor Erich Häusler and Dr. Gerrit Eichner from the stochastics research group, for providing an invaluable educational experience that gave me the knowledge and confidence needed to begin this undertaking.

My thanks also go to my family, especially to my wife Beate for her patience and encouragement during times of frustration, and to my daughters for providing a much needed balance to work and study and for asking the questions that continue to show me how much more there always is to learn.

Contents

CHAPTER 1

Introduction

In the two-sample testing problem in its most general form we are interested in deciding between the null hypothesis that two distributions F and G are equal $\mathsf{H}_0 : F = G$ and the alternative that they are different in some way $\mathsf{H}_1 : F \neq G$ on the basis of two independently identically distributed samples $X_i \sim F$, $1 \leq i \leq m$ and $Y_k \sim G$, $1 \leq k \leq n$ from F and G respectively. If the testing problem is simplified such that H_1 contains only a single fixed alternative (i.e. (F, G) are a pair of known distribution functions such that $F \neq G$ and we may write $\mathsf{H}_1 : (F, G)$) which is to be compared against the null hypothesis $\mathsf{H}_0 : (F_0, F_0)$ that both of the samples are taken from the same distribution F_0, and F, G and F_0 possess densities $\frac{dF}{d\mu}$, $\frac{dG}{d\mu}$ and $\frac{dF_0}{d\mu}$ with respect to some σ-finite measure μ, then the well-known classical Neyman-Pearson lemma shows that the most powerful α-level test for comparing H_0 and H_1 may be found quite easily by using the likelihood ratio

$$\frac{\prod_{i=1}^{m} \frac{dF}{d\mu}(X_i) \times \prod_{k=1}^{n} \frac{dG}{d\mu}(Y_k)}{\prod_{i=1}^{m} \frac{dF_0}{d\mu}(X_i) \times \prod_{k=1}^{n} \frac{dF_0}{d\mu}(Y_k)}$$

as a test statistic and setting the critical value as needed to ensure the level α is not exceeded.

In most practical applications, however, we are not willing to make such a strong assumption and specify F, G and F_0 completely. In the case of *parametric* tests we are willing to make assumptions about the form of F and G, such as in the simple t-test, where it is assumed that F and G are normal with equal variances, possibly differing in expectation (i.e. $X_i \sim N(\mu_1, \sigma^2)$ and $Y_k \sim N(\mu_2, \sigma^2)$). In this case, the testing problem becomes one of comparing hypotheses regarding whether certain parameters of the chosen distributional family are equal or not in the case of F and G, while often some nuisance parameters, such as the unknown common variance σ^2 in the example of the t-test, must still be estimated from the data.

In some applications it is not feasible or possible to make any kind of assumption regarding the form of the distributions F and G, beyond perhaps some degree of smoothness or symmetry. This leads us to the use of *nonparametric* methods which comprise large classes of tests including *permutation tests* and the *rank tests*, that we will be concerned with here.

By rank tests we mean tests which operate only on the basis of the ranks $R_{11}, R_{12}, \ldots, R_{1m}$ and $R_{21}, R_{22}, \ldots, R_{2n}$ of the $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ respectively in the pooled sample. Thus, test statistics of rank tests can be written as a function of the R_{1i} and R_{2k} alone, which brings many advantages, since the distribution of the vector of ranks $(R_{11}, R_{12}, \ldots, R_{1m}, R_{21}, R_{22}, \ldots, R_{2n})$ is known to be uniform under H_0 *regardless of the form of the underlying distribution* F, meaning that any of the $(m+n)!$ possible rank vectors in the combined sample is equally probable. This allows the distribution of the test statistic under H_0 to be determined exactly, independent of F.

There is, of course, a price to be paid for the ability to construct tests which require virtually no assumptions regarding the form of the underlying distributions to be made in order to be valid, which is put succinctly by Hájek and Šidák (1967) in their seminal work *Theory of Rank Tests*.

> We have tried to organize the multitude of rank tests into a compact system. However, we need to have *some knowledge of the form of the unknown density* in order to make a rational selection from this system.

That is, although in a given testing situation all rank tests are identically distributed under H_0 independent of F, their efficiency in terms of power under the alternatives will indeed depend on the form of the true underlying distributions.

Hájek and Šidák (1967) show, for example, that in a simple shift model where $G(x) = F(x-\theta)$ the optimal - in the sense of locally most powerful - choice of rank tests in the case of normal F is given by the statistic

$$S_N = \sum_{i=1}^{m} \Phi^{-1}\left(\frac{R_{1i}}{m+n+1}\right)$$

while the well-known Wilcoxon rank-sum test (Wilcoxon (1945)), which simply sums the ranks of the first sample

$$S_N = \sum_{i=1}^{m} R_{1i}$$

is optimal for logistic F.

In the following we will re-visit an idea presented by K. Behnen and G. Neuhaus in a series of publications (Behnen (1972); Behnen and Neuhaus (1983); Behnen et al. (1983); Behnen and Hušková (1984); Neuhaus (1987); Behnen and Neuhaus (1989)) in which tests based on statistics of the form

$$S_N(b_N) = m^{-1} \sum_{i=1}^{m} b_N\left(\frac{R_{1i}}{N}\right) \tag{1.1}$$

are proposed, where

$$H_N = \frac{m}{N}\,F + \frac{n}{N}\,G$$

with $N = m+n$ is the *pooled* distribution function and

$$b_N = f_N - g_N$$

where f_N and g_N are the Lebesgue-densities of the $H_N(X_i)$ and $H_N(Y_k)$ respectively.

In the works cited above the authors consider the broader class of nonparametric alternatives of the form $\mathsf{H}_1 : F \neq G$ rather than the simpler more restrictive shift model alternatives $\mathsf{H}_1 : G(x) = F(x-\theta)$, $\theta \neq 0$. In this context statistics of the form (1.1) can be motivated - among other ways - by considering the case of testing a simple fixed alternative (i.e. $X_i \sim F$ and $Y_k \sim G$ for a known pair (F, G) of distribution functions with $F \neq G$) against the simple hypothesis $\mathsf{H}_0 : X_i \sim H_N$, $Y_k \sim H_N$ (i.e. both X_i and Y_k come from the pooled distribution H_N). Under the assumption that F and G are absolutely continuous with Lebesgue-densities $\frac{dF}{d\mu}$ and $\frac{dG}{d\mu}$ then $H_N = \frac{m}{N}F + \frac{n}{N}G$ is absolutely continuous as well with Lebesgue-density $\frac{dH_N}{d\mu}$ and $F = H_N = G$ under H_0 so that the optimal test is given according to the Neyman-Pearson lemma by the likelihood - or equivalent log-likelihood - statistic:

$$\log\left[\frac{\prod_{i=1}^{m} \frac{dF}{d\mu}(X_i) \times \prod_{k=1}^{n} \frac{dG}{d\mu}(Y_k)}{\prod_{i=1}^{m} \frac{dH_N}{d\mu}(X_i) \times \prod_{k=1}^{n} \frac{dH_N}{d\mu}(Y_k)}\right]. \tag{1.2}$$

H_N obviously dominates F and G so there exist Radon-Nikodym derivatives $\frac{dF}{dH_N}$ and $\frac{dG}{dH_N}$ and we may write (1.2) as

$$\begin{aligned}
&\log\left[\prod_{i=1}^{m} \frac{\frac{dF}{d\mu}(X_i)}{\frac{dH_N}{d\mu}(X_i)} \times \prod_{k=1}^{n} \frac{\frac{dG}{d\mu}(Y_k)}{\frac{dH_N}{d\mu}(Y_k)}\right] \\
&= \sum_{i=1}^{m} \log\left(\frac{dF}{dH_N}(X_i)\right) + \sum_{k=1}^{n} \log\left(\frac{dG}{dH_N}(Y_k)\right) \\
&= \sum_{i=1}^{m} \log(f_N \circ H_N(X_i)) + \sum_{k=1}^{n} \log(g_N \circ H_N(Y_k)) \\
&= \sum_{i=1}^{m} \log\left[1 + nN^{-1}\, b_N \circ H_N(X_i)\right] + \sum_{k=1}^{n} \log\left[1 - mN^{-1}\, b_N \circ H_N(Y_k)\right]
\end{aligned}$$

by using the fact that

$$f_N = \frac{dF}{dH_N} \circ H_N^{-1} \qquad \text{and} \qquad g_N = \frac{dG}{dH_N} \circ H_N^{-1}$$

(see e.g. Behnen and Neuhaus (1989)) and

$$\frac{m}{N}\, f_N + \frac{n}{N}\, g_N = 1$$

(see proof of lemma A.1). Replacing $H_N(X_i)$ and $H_N(Y_k)$ by the natural empirical estimators $\hat{H}_N(X_i) = N^{-1}R_{1i}$ and $\hat{H}_N(Y_k) = N^{-1}R_{2k}$ leads to the rank statistic

$$\sum_{i=1}^{m} \log\left[1 + nN^{-1}\, b_N(N^{-1}R_{1i})\right] + \sum_{k=1}^{n} \log\left[1 - mN^{-1}\, b_N(N^{-1}R_{2k})\right]$$

which can be approximated in local situations where $\|b_N\| \to 0$ (see Behnen and Neuhaus (1983)) by

$$\begin{aligned}
&\sum_{i=1}^{m} nN^{-1}\, b_N(N^{-1}R_{1i}) - \sum_{k=1}^{n} mN^{-1}\, b_N(N^{-1}R_{2k}) \\
&\quad = \sum_{i=1}^{m} nN^{-1}\, b_N(N^{-1}R_{1i}) + mN^{-1} \sum_{i=1}^{m} b_N(N^{-1}R_{1i}) \\
&\qquad - mN^{-1} \sum_{i=1}^{m} b_N(N^{-1}R_{1i}) - \sum_{k=1}^{n} mN^{-1}\, b_N(N^{-1}R_{2k}) \\
&\quad = (nN^{-1} + mN^{-1}) \sum_{i=1}^{m} b_N(N^{-1}R_{1i}) - mN^{-1} \sum_{i=1}^{N} b_N(N^{-1}i) \\
&\quad = m(S_N(b_N) - \int_0^1 b_N(u)\, du + o(1)) \\
&\quad = m(S_N(b_N) + o(1))
\end{aligned}$$

since

$$\int_0^1 b_N(u)\, du = \int_0^1 f_N(u)\, du - \int_0^1 g_N(u)\, du = 0.$$

In practical applications the problem remains, however, of how to estimate $b_N = f_N - g_N$ from the data. Behnen and Neuhaus (1989) propose - among other approaches - to use kernel density estimators of the form

$$\hat{\hat{f}}_N(t) = m^{-1}\sum_{i=1}^{m} K_N\left(t,\ N^{-1}\left(R_{1i}-\frac{1}{2}\right)\right) \qquad \text{and} \qquad \hat{\hat{g}}_N(t) = n^{-1}\sum_{k=1}^{n} K_N\left(t,\ N^{-1}\left(R_{2k}-\frac{1}{2}\right)\right)$$

where

$$K_N(t,s) = a_N^{-1}\left[K\left(\frac{t-s}{a_N}\right) + K\left(\frac{t+s}{a_N}\right) + K\left(\frac{t-2+s}{a_N}\right)\right]$$

which are essentially kernel density estimators using the shifted and scaled original ranks of the first and second samples

$$\frac{R_{1i}-\frac{1}{2}}{N},\ 1 \le i \le m \qquad \text{and} \qquad \frac{R_{2k}-\frac{1}{2}}{N},\ 1 \le k \le n,$$

each augmented by the artificial samples created by reflecting the $N^{-1}(R_{1i}-\frac{1}{2})$ and $N^{-1}(R_{2k}-\frac{1}{2})$ about the points 0 and 1 respectively. This has the effect of making certain that $\hat{\hat{f}}_N$ and $\hat{\hat{g}}_N$ are, as the true f_N and g_N, probability densities on $[0,1]$ with $\int_0^1 \hat{\hat{f}}_N(u)\,du = \int_0^1 \hat{\hat{g}}_N(u)\,du = 1$ for all N. For this reason we will refer to $\hat{\hat{f}}_N$ and $\hat{\hat{g}}_N$ as the *restricted* kernel density estimators that lead to the non-linear adaptive rank statistic

$$S_N(\hat{\hat{b}}_N) = m^{-1}\sum_{i=1}^{m}\hat{\hat{b}}_N\left(\frac{R_{1i}-\frac{1}{2}}{N}\right). \tag{1.3}$$

Behnen and Hušková (1984) claim asymptotic normality of (1.3) under $\mathsf{H}_0 : F = G$ after proper centering and scaling

$$m a_N^{\frac{1}{2}}\, S_N(\hat{\hat{b}}_N) \xrightarrow[N]{\mathcal{L}} N(0,1)$$

for $K : [0,1] \to \mathbb{R}_0$ suitably smooth and $\frac{1}{2} > a_N \to 0$ such that $N a_N^6 \to \infty$, so that it appears asymptotic theory could be used to get critical values and p-values for $S_N(\hat{\hat{b}}_N)$ for N suitably large. However, extensive simulations showed that even for very large sample sizes ($N = 2000$) the resulting distribution is neither centered, nor standardized, nor normal (see chapter 3).

In the present work we approach the estimation problem again using simple, non-restricted kernel density estimators

$$\hat{f}_N(t) = m^{-1}a_N^{-1}\sum_{i=1}^{m} K\left(\frac{t-N^{-1}R_{1i}}{a_N}\right) \qquad \text{and} \qquad \hat{g}_N(t) = n^{-1}a_N^{-1}\sum_{k=1}^{n} K\left(\frac{t-N^{-1}R_{2k}}{a_N}\right).$$

As it will turn out, these will admit a linearization of $S_N(\hat{b}_N)$ as a simple i.i.d. sum and negligible rest terms for bandwidth sequences a_N converging even more quickly to 0 ($N a_N^5 \to \infty$) from which we can derive asymptotic normality under H_0 as $N \to \infty$. Monte-carlo simulations in chapter 3 show that there are still problems with centering and scaling under H_0 which can be corrected by introducing appropriate modifications to $\hat{f}_N$, improved variance estimates for $\mathsf{Var}[S_N(\hat{b}_N)]$, and K other than the typical bell-shaped kernels. However, further simulations in chapter 4 show that although $a_N \to 0$ more quickly, in most cases there is a price to be paid when using the non-restricted kernel estimators $\hat{f}_N$ and $\hat{g}_N$ as far as reduced power under H_1.

CHAPTER 2

Main results

2.1. Definitions and notation

In order to work with the general two-sample testing problem of comparing distribution functions F and G against stochastic alternatives

$$\mathsf{H}_0 \,:\, F = G \qquad\qquad \text{versus} \qquad\qquad \mathsf{H}_1 \,:\, F \le G,\; F \ge G,\; F \ne G$$

using independent samples $X_1,\, X_2,\, \ldots,\, X_m$ i.i.d. from F and $Y_1,\, Y_2,\, \ldots,\, Y_n$ i.i.d. from G, we will use the following definitions, notation and assumptions throughout.

Let

$$X_i \sim F, \quad 1 \le i \le m, \qquad\qquad \text{and} \qquad\qquad Y_k \sim G, \quad 1 \le k \le n \tag{2.1}$$

be independent, real-valued random variables with continuous distribution functions F and G, and let

$$R_{11},\, R_{12},\, \ldots,\, R_{1m} \qquad\qquad \text{and} \qquad\qquad R_{21},\, R_{22},\, \ldots,\, R_{2n} \tag{2.2}$$

be the ranks of $X_1,\, X_2,\, \ldots,\, X_m$ and $Y_1,\, Y_2,\, \ldots,\, Y_n$ in the *pooled* sample respectively.

Further, let

$$N = m + n \qquad\qquad \text{and} \qquad\qquad \lambda_N = \frac{m}{N} \tag{2.3}$$

be the size of the pooled sample and the fraction of the pooled sample made up of the first sample, and

$$H_N = \frac{m}{N}\, F + \frac{n}{N}\, G \tag{2.4}$$

be the continuous distribution function defined by the mixture of F and G with respect to the fractions of the sample sizes.

In the sequel we will often work with the random variables $H_N(X_i)$ and $H_N(Y_k)$. These can be shown to have distribution functions $F \circ H_N^{-1}$ and $G \circ H_N^{-1}$ respectively (see Lemma A.1). Since $F \circ H_N^{-1}$ and $G \circ H_N^{-1}$ are dominated by the Lebesgue measure μ on the interval $(0, 1)$ (see Behnen and Neuhaus (1989), Chapter 1.3), we can define f_N and g_N to be the Lebesgue-densities of the random variables $H_N(X_i)$ and $H_N(Y_k)$:

$$f_N = \frac{d(F \circ H_N^{-1})}{d\mu} \qquad\qquad \text{and} \qquad\qquad g_N = \frac{d(G \circ H_N^{-1})}{d\mu}.$$

Later in our development of the test statistic, we will use kernel estimators of the densities f_N and g_N. For this reason we will require a bandwidth sequence a_N and a kernel K with the following properties:

$$a_N < \frac{1}{2} \;\forall\; N, \tag{2.5}$$

$$a_N \to 0 \text{ as } N \to \infty, \tag{2.6}$$

$$Na_N^5 \to \infty \text{ as } N \to \infty, \tag{2.7}$$

$$K \text{ is symmetric,} \tag{2.8}$$

$$K \text{ is zero outside of } (-1, 1), \tag{2.9}$$

$$K \text{ is twice continuously differentiable,} \tag{2.10}$$

$$\int_{-1}^{1} K(v)\, dv = 1. \tag{2.11}$$

Now, we introduce the kernel estimators $\hat{f}_N$ and $\hat{g}_N$

$$\hat{f}_N(t) = m^{-1} a_N^{-1} \sum_{i=1}^{m} K\left(\frac{t - N^{-1} R_{1i}}{a_N}\right), \tag{2.12}$$

$$\hat{g}_N(t) = n^{-1} a_N^{-1} \sum_{k=1}^{n} K\left(\frac{t - N^{-1} R_{2k}}{a_N}\right). \tag{2.13}$$

Since $\hat{f}_N$ and $\hat{g}_N$ are rank-based estimators and F and G are continuous, we may assume no ties without loss of generality and the kernel estimators may be written as

$$\hat{f}_N(t) = m^{-1} a_N^{-1} \sum_{i=1}^{m} K\left(\frac{t - \hat{H}_N(X_i)}{a_N}\right), \tag{2.14}$$

$$\hat{g}_N(t) = n^{-1} a_N^{-1} \sum_{k=1}^{n} K\left(\frac{t - \hat{H}_N(Y_k)}{a_N}\right), \tag{2.15}$$

where $\hat{H}_N$ is the empirical distribution function of the pooled sample

$$\hat{H}_N = \frac{m}{N}\, \hat{F}_m + \frac{n}{N}\, \hat{G}_n. \tag{2.16}$$

At this point, we also define functions $\bar{f}_N$ and $\bar{g}_N$

$$\bar{f}_N(t) = a_N^{-1} \int K\left(\frac{t - H_N(y)}{a_N}\right) F(dy), \quad 0 \le t \le 1, \tag{2.17}$$

$$\bar{g}_N(t) = a_N^{-1} \int K\left(\frac{t - H_N(y)}{a_N}\right) G(dy), \quad 0 \le t \le 1, \tag{2.18}$$

theoretical analogs to the empirical (2.12) and (2.13) which we will use frequently to center certain random variables involving the kernel estimators $\hat{f}_N$ and $\hat{g}_N$.

Lastly, define b_N, $\hat{b}_N$ and $\bar{b}_N$ as differences

$$b_N = f_N - g_N, \qquad \hat{b}_N = \hat{f}_N - \hat{g}_N \qquad \text{and} \qquad \bar{b}_N = \bar{f}_N - \bar{g}_N. \tag{2.19}$$

In addition, all asymptotic results will be under the standard assumption that the ratio of the two sample sizes converges to some constant, i.e.

$$\lambda_N \to \lambda \in (0, 1) \text{ as } N \to \infty. \tag{2.20}$$

2.2. Results

In this chapter I will present the main results of my work with the test statistic $S_N(\hat{b}_N)$ proposed below, showing first a representation of $S_N(\hat{b}_N)$ as a centered i.i.d sum, a negligible term, and a deterministic term that vanishes under $\mathsf{H}_0 : F = G$ but is responsible for the power of the test under H_1. In a second theorem I will show asymptotic normality of $S_N(\hat{b}_N)$ under H_0 after proper scaling and present a simple representation of the asymptotic null variance, so that critical values and p-values for the asymptotic test can be calculated quickly and easily from the standard normal distribution.

THEOREM 2.1. *Define the kernel estimators* $\hat{f}_N$ *and* $\hat{g}_N$ *as* (2.12) *and* (2.13) *and set* $\hat{b}_N = \hat{f}_N - \hat{g}_N$. *Define the test statistic* $S_N = S_N(\hat{b}_N)$ *as*

$$S_N(\hat{b}_N) = m^{-1} \sum_{i=1}^{m} \hat{b}_N(N^{-1} R_{1i})$$

and let the functions $\bar{f}_N$, $\bar{g}_N$ *be defined as in* (2.17) *and* (2.18).

Then under the assumptions (2.5) *through* (2.7) *on the bandwidth sequence* a_N *as well as the assumptions* (2.8) *through* (2.11) *and* (2.20) *on the kernel function* K, *we have for any continuous distribution functions* F *and* G

$$\begin{aligned}
S_N(\hat{b}_N) = & \int \Big[\bar{f}_N - \bar{g}_N\Big] \circ H_N(x) \Big[\hat{F}_m(dx) - F(dx)\Big] && (2.21)\\
& + \int \Big[\bar{f}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big] \, F(dx) && (2.22)\\
& + \int \bar{f}_N \circ H_N(x) \Big[\hat{F}_m(dx) - F(dx)\Big] && (2.23)\\
& - \int \bar{f}_N \circ H_N(x) \Big[\hat{G}_n(dx) - G(dx)\Big] && (2.24)\\
& + \int \bar{f}'_N \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big] \, F(dx) && (2.25)\\
& - \int \bar{f}'_N \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big] \, G(dx) && (2.26)\\
& + \int \Big[\bar{f}_N - \bar{g}_N\Big] \circ H_N(x) \, F(dx) && (2.27)\\
& + O_P(N^{-1} a_N^{-2}). && (2.28)
\end{aligned}$$

We note that (2.21) through (2.26) are simple centered i.i.d. sums, while (2.27) is the non-random term responsible for power under the alternative. The following is a result of Theorem 2.1 giving the asymptotic null distribution of $S_N(\hat{b}_N)$.

THEOREM 2.2. *Under the assumptions of Theorem 2.1, we have under the null hypothesis* $\mathsf{H}_0 : F = G$

$$N^{\frac{1}{2}} a_N^{-\frac{1}{2}} \cdot S_N(\hat{b}_N) \xrightarrow[N]{\mathcal{L}} N\big(0,\, \sigma^2_{K,\lambda}\big) \tag{2.29}$$

with

$$\sigma^2_{K,\lambda} = 2 \left[\lambda^{-1} + (1-\lambda)^{-1}\right] \int_{-1}^{0} \left[\int_{-1}^{x} K(v)\, dv\right]^2 dx. \tag{2.30}$$

In order to prove Theorem 2.1, we will proceed by first deriving an integral representation of S_N, which can then be decomposed into terms which are either asymptotically negligible, responsible for the asymptotic distribution or responsible for power.

Proof of Theorem 2.1.

$$\begin{aligned}
S_N(\hat{b}_N) &= m^{-1}\sum_{i=1}^{m}\hat{b}_N(N^{-1}R_{1i})\\
&= m^{-1}\sum_{i=1}^{m}\Big[\hat{f}_N(N^{-1}R_{1i}) - \hat{g}_N(N^{-1}R_{1i})\Big]\\
&= m^{-1}\sum_{i=1}^{m}\Big[\hat{f}_N \circ \hat{H}_N(X_i) - \hat{g}_N \circ \hat{H}_N(X_i)\Big]\\
&= \int \Big[\hat{f}_N - \hat{g}_N\Big] \circ \hat{H}_N(x)\, \hat{F}_m(dx).
\end{aligned}$$

Next, we expand the integral representation by centering with functions $\bar{f}_N$ and $\bar{g}_N$. This gives us

$$S_N = \int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big] \circ \hat{H}_N(x) \Big[\hat{F}_m(dx) - F(dx)\Big] \tag{2.31}$$

$$+ \int \Big[\bar{f}_N - \bar{g}_N\Big] \circ \hat{H}_N(x) \Big[\hat{F}_m(dx) - F(dx)\Big] \tag{2.32}$$

$$+ \int \Big[\bar{f}_N - \bar{g}_N\Big] \circ \hat{H}_N(x)\, F(dx) \tag{2.33}$$

$$+ \int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big] \circ \hat{H}_N(x)\, F(dx) \tag{2.34}$$

We further expand this by applying Taylor (see Remark 1) to (2.31), (2.32), (2.33) and (2.34) respectively, which yields

$$S_N = \int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big] \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big] \tag{2.35}$$

$$+ \int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\Big[\hat{F}_m(dx) - F(dx)\Big] \tag{2.36}$$

$$+ \int \int_{H_N(x)}^{\hat{H}_N(x)} \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big]''(t) \cdot \big(\hat{H}_N(x) - t\big)\, dt\, \Big[\hat{F}_m(dx) - F(dx)\Big] \tag{2.37}$$

$$+ \int \Big[\bar{f}_N - \bar{g}_N\Big] \circ H_N(x) \Big[\hat{F}_m(dx) - F(dx)\Big] \tag{2.38}$$

$$+ \int \Big[\bar{f}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big] \Big[\hat{F}_m(dx) - F(dx)\Big] \tag{2.39}$$

$$+ \int \int_{H_N(x)}^{\hat{H}_N(x)} \Big[\bar{f}_N - \bar{g}_N\Big]''(t) \cdot \big(\hat{H}_N(x) - t\big)\, dt\, \Big[\hat{F}_m(dx) - F(dx)\Big] \tag{2.40}$$

$$+ \int [\bar{f}_N - \bar{g}_N] \circ H_N(x)\, F(dx) \tag{2.41}$$

$$+ \int \Big[\bar{f}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx) \tag{2.42}$$

$$+ \int \int_{H_N(x)}^{\hat{H}_N(x)} \Big[\bar{f}_N - \bar{g}_N\Big]''(t) \cdot \big(\hat{H}_N(x) - t\big)\, dt\, F(dx) \tag{2.43}$$

$$+ \int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big] \circ H_N(x)\, F(dx) \tag{2.44}$$

$$+ \int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx) \tag{2.45}$$

$$+ \int \int_{H_N(x)}^{\hat{H}_N(x)} \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big]''(t) \cdot \big(\hat{H}_N(x) - t\big)\, dt\, F(dx). \tag{2.46}$$

Lemmas 5.17, 5.24, 5.27 and 5.32 show that terms (2.35), (2.36), (2.39) and (2.45) are of the orders $O_{\mathbb{P}}(N^{-1}a_N^{-2})$, $O_{\mathbb{P}}(N^{-1}a_N^{-\frac{3}{2}})$, $O_{\mathbb{P}}(N^{-1}a_N^{-2})$, and $O_{\mathbb{P}}(N^{-1}a_N^{-2})$ respectively and the combination of the four Taylor rest terms (2.37), (2.40), (2.43) and (2.46) is shown in Lemma 5.10 to be asymptotically negligible of the order $O_{\mathbb{P}}(N^{-1}a_N^{-2})$ as well. Altogether this yields

$$\begin{aligned}
S_N = &\int \Big[\bar{f}_N - \bar{g}_N\Big] \circ H_N(x) \Big[\hat{F}_m(dx) - F(dx)\Big] \\
&+ \int \big[\bar{f}_N - \bar{g}_N\big] \circ H_N(x)\, F(dx) \\
&+ \int \Big[\bar{f}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx) \\
&+ \int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big] \circ H_N(x)\, F(dx) \\
&+ O_{\mathbb{P}}(N^{-1}a_N^{-2}).
\end{aligned}$$

Use Lemma 5.9 to write the last integral as the sum of four simple integrals and a negligible term and rearrange terms to get the desired representation of S_N:

$$\begin{aligned}
S_N(\hat{b}_N) = &\int \Big[\bar{f}_N - \bar{g}_N\Big] \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big] \\
&+ \int \Big[\bar{f}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx) \\
&+ \int \bar{f}_N \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big] \\
&- \int \bar{f}_N \circ H_N(x)\Big[\hat{G}_n(dx) - G(dx)\Big] \\
&+ \int \bar{f}_N' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx) \\
&- \int \bar{f}_N' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, G(dx) \\
&+ \int \Big[\bar{f}_N - \bar{g}_N\Big] \circ H_N(x)\, F(dx) \\
&+ O_{\mathbb{P}}(N^{-1}a_N^{-2}).
\end{aligned}$$

□

REMARK 1. Here – and later in further expansions of the leading terms (2.35), (2.36), (2.38), (2.39), (2.41), (2.42), (2.44) and (2.45) as well – we will often use the integral form of the Taylor remainder (see Chapter 14 of Königsberger (2004)) rather than the Lagrange form, which will help us to more easily achieve a sharper upper bound for the respective rest terms.

PROOF OF THEOREM 2.2. Recall again the representation of S_N shown in Theorem 2.1 to be valid under $\mathsf{H}_0 : F = G$ as well as under the alternative $\mathsf{H}_1 : F \neq G$:

$$
\begin{aligned}
S_N(\hat{b}_N) = &\int \Big[\bar{f}_N - \bar{g}_N\Big] \circ H_N(x) \Big[\hat{F}_m(dx) - F(dx)\Big] && (\ 2.21) \\
&+ \int \Big[\bar{f}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx) && (\ 2.22) \\
&+ \int \bar{f}_N \circ H_N(x) \Big[\hat{F}_m(dx) - F(dx)\Big] && (\ 2.23) \\
&- \int \bar{f}_N \circ H_N(x) \Big[\hat{G}_n(dx) - G(dx)\Big] && (\ 2.24) \\
&+ \int \bar{f}'_N \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx) && (\ 2.25) \\
&- \int \bar{f}'_N \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, G(dx) && (\ 2.26) \\
&+ \int \Big[\bar{f}_N - \bar{g}_N\Big] \circ H_N(x)\, F(dx) && (\ 2.27) \\
&+ O_{\mathbb{P}}(N^{-1}a_N^{-2}).
\end{aligned}
$$

If we restrict ourselves to H_0, then the terms (2.21), (2.22), (2.25), (2.26) and (2.27) vanish, since in this case $\bar{f}_N = \bar{g}_N$, so that under H_0 we have

$$
\begin{aligned}
S_N(\hat{b}_N) &= \int \bar{f}_N \circ H_N(x) \Big[\hat{F}_m(dx) - F(dx)\Big] \\
&\quad - \int \bar{f}_N \circ H_N(x) \Big[\hat{G}_n(dx) - G(dx)\Big] \\
&\quad + O_{\mathbb{P}}(N^{-1}a_N^{-2}) \\
&= \sum_{i=1}^{m} m^{-1} a_N^{-1} \Big[\int K\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \\
&\qquad - \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\, F(dx) \Big] && (2.47) \\
&\quad - \sum_{k=1}^{n} n^{-1} a_N^{-1} \Big[\int K\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) \\
&\qquad - \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\, F(dx) \Big] && (2.48) \\
&\quad + O_{\mathbb{P}}(N^{-1}a_N^{-2})
\end{aligned}
$$

From this follows that the asymptotic null distribution of $S_N(\hat{b}_N)$ will be completely determined by the fairly simple i.i.d. sums (2.47) and (2.48) after proper scaling.

If we define w_N as

$$
w_N(s) = a_N^{-1} \int K\big(a_N^{-1}(H_N(x) - H_N(s))\big)\, F(dx)
$$

then we may write the sums (2.47) and (2.48) as

$$
T_N = \sum_{i=1}^{m} m^{-1} \big[w_N(X_i) - \mathbb{E}\big[w_N(X_1)\big]\big] - \sum_{k=1}^{n} n^{-1} \big[w_N(Y_k) - \mathbb{E}\big[w_N(Y_1)\big]\big].
$$

Now, we see immediately that the sequence of sums T_N is formed by summing across rows of a triangular array with centered, mutually independent summands $m^{-1}\big[w_N(X_i) - \mathbb{E}\big[w_N(X_1)\big]\big]$, $1 \le i \le m$, and $-n^{-1}\big[w_N(Y_k) - \mathbb{E}\big[w_N(Y_1)\big]\big]$, $1 \le k \le n$. Let

$$\sigma_N^2 = \mathrm{Var}\big(T_N\big).$$

Then

$$\begin{aligned}
\sigma_N^2 &= \sum_{i=1}^m \mathrm{Var}\big(m^{-1}\big[w_N(X_i) - \mathbb{E}\big[w_N(X_1)\big]\big]\big) + \sum_{k=1}^n \mathrm{Var}\big(n^{-1}\big[w_N(Y_k) - \mathbb{E}\big[w_N(Y_1)\big]\big]\big) \\
&= \sum_{i=1}^m m^{-2}\,\mathbb{E}\big[w_N(X_1) - \mathbb{E}\big[w_N(X_1)\big]\big]^2 + \sum_{k=1}^n n^{-2}\,\mathbb{E}\big[w_N(Y_1) - \mathbb{E}\big[w_N(Y_1)\big]\big]^2 \\
&= m^{-1}\,\mathbb{E}\big[w_N(X_1) - \mathbb{E}\big[w_N(X_1)\big]\big]^2 + n^{-1}\,\mathbb{E}\big[w_N(Y_1) - \mathbb{E}\big[w_N(Y_1)\big]\big]^2.
\end{aligned}$$

But under H_0 we have $X_1 \sim Y_1$, so this simplifies to

$$\sigma_N^2 = (m^{-1} + n^{-1})\,\mathbb{E}\big[w_N(X_1) - \mathbb{E}\big[w_N(X_1)\big]\big]^2.$$

Thus, using Lemmas 5.33 and 5.34 we may write σ_N^2 as

$$\begin{aligned}
\sigma_N^2 &= (m^{-1} + n^{-1})\,\mathbb{E}\big[w_N(X_1) - \mathbb{E}\big[w_N(X_1)\big]\big]^2 \\
&= (m^{-1} + n^{-1})\Bigg[\mathbb{E}\left[a_N^{-1}\int K\left(a_N^{-1}(H_N(x) - H_N(X_1))\right)\,F(dx)\right]^2 \\
&\qquad - \left[a_N^{-1}\iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\,F(dx)\,F(dy)\right]^2\Bigg] \\
&= (m^{-1} + n^{-1})\Bigg[1 + 2\,a_N\int_{-1}^0\left[\int_{-1}^x K(v)\,dv\right]^2 dx - 4\,a_N\int_0^1 vK(v)\,dv \\
&\qquad - \left[1 - 4\,a_N\int_0^1 vK(v)\,dv + 4\,a_N^2\left[\int_0^1 vK(v)\,dv\right]^2\right]\Bigg] \\
&= (m^{-1} + n^{-1})\left[2\,a_N\int_{-1}^0\left[\int_{-1}^x K(v)\,dv\right]^2 dx - 4\,a_N^2\left[\int_0^1 vK(v)\,dv\right]^2\right] \\
&= N^{-1}a_N\big[\lambda_N^{-1} + (1-\lambda_N)^{-1}\big]\cdot\left[2\int_{-1}^0\left[\int_{-1}^x K(v)\,dv\right]^2 dx - 4\,a_N\left[\int_0^1 vK(v)\,dv\right]^2\right].
\end{aligned}$$

From this representation we see that

$$\lim_N\,(Na_N^{-1})\cdot\sigma_N^2 = \sigma_{K,\lambda}^2,$$

and thus that

$$\sigma_N^2 = O(N^{-1}a_N).$$

Now, the sequence $\sigma_N^{-1}\,T_N$ is a triangular array with centered, mutually independent summands

$$\sigma_N^{-1}m^{-1}\big[w_N(X_i) - \mathbb{E}\big[w_N(X_1)\big]\big],\ 1 \le i \le m,$$

and

$$-\sigma_N^{-1}n^{-1}\big[w_N(Y_k) - \mathbb{E}\big[w_N(Y_1)\big]\big],\ 1 \le k \le n,$$

such that

$$\sum_{i=1}^{m} \mathrm{Var}\big(\sigma_N^{-1} m^{-1}\big[w_N(X_i) - \mathbb{E}\big[w_N(X_1)\big]\big]\big) + \sum_{k=1}^{n} \mathrm{Var}\big(\sigma_N^{-1} n^{-1}\big[w_N(Y_k) - \mathbb{E}\big[w_N(Y_1)\big]\big]\big) = 1.$$

Also, due to (A.2) in Lemma A.1, we can bound w_N with a convergent sequence:

$$\big\|w_N\big\| \le 2\,\big\|K\big\|(1 + nm^{-1}),$$

so that

$$\begin{aligned}\mathrm{Var}(w_N(X_1)) &= \mathbb{E}\big[w_N(X_1) - \mathbb{E}\big[w_N(X_1)\big]\big]^2 \\ &\le 8\,\big\|K\big\|^2 (1 + nm^{-1})^2.\end{aligned}$$

This will allow us to easily show that the Lindeberg condition is satisfied since $\forall\ \epsilon > 0$

$$\begin{aligned}&\sum_{i=1}^{m} \mathbb{E}\big[\big[\sigma_N^{-1} m^{-1}\big[w_N(X_i) - \mathbb{E}\big[w_N(X_1)\big]\big]\big]^2 \cdot 1_{\{|\sigma_N^{-1} m^{-1}[w_N(X_i) - \mathbb{E}[w_N(X_1)]]| > \epsilon\}}\big] \\ &\quad + \sum_{k=1}^{n} \mathbb{E}\big[\big[\sigma_N^{-1} n^{-1}\big[w_N(Y_k) - \mathbb{E}\big[w_N(Y_1)\big]\big]\big]^2 \cdot 1_{\{|\sigma_N^{-1} n^{-1}[w_N(Y_k) - \mathbb{E}[w_N(Y_1)]]| > \epsilon\}}\big] \\ &\le \sum_{i=1}^{m} 16\,\big\|K\big\|^2 (1 + nm^{-1})^2\, \sigma_N^{-2} m^{-2} \cdot \mathbb{E}\big[1_{\{|\sigma_N^{-1} m^{-1}[w_N(X_i) - \mathbb{E}[w_N(X_1)]]| > \epsilon\}}\big] \\ &\quad + \sum_{k=1}^{n} 16\,\big\|K\big\|^2 (1 + nm^{-1})^2\, \sigma_N^{-2} n^{-2} \cdot \mathbb{E}\big[1_{\{|\sigma_N^{-1} n^{-1}[w_N(Y_k) - \mathbb{E}[w_N(Y_1)]]| > \epsilon\}}\big] \\ &\le 16\,\big\|K\big\|^2 (1 + nm^{-1})^2\, \sigma_N^{-2} m^{-1} \cdot P\big(\big|w_N(X_1) - \mathbb{E}\big[w_N(X_1)\big]\big| > \epsilon \cdot \sigma_N\, m\big) \\ &\quad + 16\,\big\|K\big\|^2 (1 + nm^{-1})^2\, \sigma_N^{-2} n^{-1} \cdot P\big(\big|w_N(Y_1) - \mathbb{E}\big[w_N(Y_1)\big]\big| > \epsilon \cdot \sigma_N\, n\big) \\ &\le 16\,\big\|K\big\|^2 (1 + nm^{-1})^2\, \sigma_N^{-2} m^{-1} \cdot \mathrm{Var}\big(w_N(X_1)\big) \cdot \big(\epsilon \cdot \sigma_N\, m\big)^{-2} \\ &\quad + 16\,\big\|K\big\|^2 (1 + nm^{-1})^2\, \sigma_N^{-2} n^{-1} \cdot \mathrm{Var}\big(w_N(Y_1)\big) \cdot \big(\epsilon \cdot \sigma_N\, n\big)^{-2} \\ &\le 128\,\big\|K\big\|^4 (1 + nm^{-1})^4\, \sigma_N^{-4} m^{-3} \epsilon^{-2} + 128\,\big\|K\big\|^4 (1 + nm^{-1})^4\, \sigma_N^{-4} n^{-3} \epsilon^{-2} \\ &= O(N^2\, a_N^{-2}) \cdot O(N^{-3}) \\ &= O(N^{-1} a_N^{-2}),\end{aligned}$$

and $N^{-1} a_N^{-2} \longrightarrow 0$, since we require that our bandwidth sequence a_N converge to zero slowly enough that $N a_N^5 \longrightarrow \infty$.

Thus, we have (2.29) immediately by Slutsky's Theorem and the Central Limit Theorem for triangular arrays, since

$$\begin{aligned}N^{\frac{1}{2}} a_N^{-\frac{1}{2}} \cdot S_N(\hat{b}_N) &= N^{\frac{1}{2}} a_N^{-\frac{1}{2}} \cdot (T_N + O_{\mathbb{P}}(N^{-1} a_N^{-2})) \\ &= N^{\frac{1}{2}} a_N^{-\frac{1}{2}} \cdot T_N + O_{\mathbb{P}}(N^{-\frac{1}{2}} a_N^{-\frac{5}{2}})\end{aligned}$$

and

$$N^{\frac{1}{2}} a_N^{-\frac{1}{2}} \cdot T_N = (N^{\frac{1}{2}} a_N^{-\frac{1}{2}} \sigma_N) \cdot (\sigma_N^{-1}\, T_N) \xrightarrow[N]{\mathcal{L}} N\big(0,\, \sigma_{K,\lambda}^2\big).$$

$\square$

CHAPTER 3

A modified test statistic

In this chapter, we will first use simulation results to highlight some problems with adaptive rank statistics $S_N(\hat{b}_N)$ of the form described in chapter 1 which lead to unexpected behavior under the null hypothesis of equal distributions $F = G$. Then in further simulations we will use a heuristic approach to try to isolate the source of these problems and propose some simple changes to $\hat{b}_N$ and some modified variance estimators that lead to improved behavior of the statistic under H_0.

We begin by looking at simulations of $S_N(\hat{\hat{b}}_N)$ for $\hat{\hat{b}}_N = \hat{\hat{f}}_N - \hat{\hat{g}}_N$ with kernel estimators

$$\hat{\hat{f}}_N(t) = m^{-1} \sum_{i=1}^{m} a_N^{-1} \left[K\left(\frac{t + \frac{R_{1i} - \frac{1}{2}}{N}}{a_N} \right) + K\left(\frac{t - \frac{R_{1i} - \frac{1}{2}}{N}}{a_N} \right) + K\left(\frac{t - 2 + \frac{R_{1i} - \frac{1}{2}}{N}}{a_N} \right) \right], \tag{3.1}$$

$$\hat{\hat{g}}_N(t) = n^{-1} \sum_{k=1}^{n} a_N^{-1} \left[K\left(\frac{t + \frac{R_{2k} - \frac{1}{2}}{N}}{a_N} \right) + K\left(\frac{t - \frac{R_{2k} - \frac{1}{2}}{N}}{a_N} \right) + K\left(\frac{t - 2 + \frac{R_{2k} - \frac{1}{2}}{N}}{a_N} \right) \right] \tag{3.2}$$

as proposed in Behnen et al. (1983) and Behnen and Hušková (1984) and using the centering and scaling from Thereom 2.2 Behnen and Hušková (1984), where it is claimed that for $S_N(\hat{\hat{b}}_N)$ with the kernel estimators described above

$$m\, a_N^{\frac{1}{2}} \left[2 \int K^2(x)\, dx \right]^{-\frac{1}{2}} \cdot \left[S_N(\hat{\hat{b}}_N) - m^{-1} a_N^{-1} K(0) \right] \xrightarrow[N]{\mathcal{L}} N(0,1) \tag{3.3}$$

under H_0 for a kernel K fulfilling (2.8) through (2.11) and a bandwidth sequence a_N such that $\frac{1}{2} > a_N \to 0$ and $N\, a_N^6 \to \infty$.

Each histogram in figure 2 shows the results of 10,000 monte-carlo simulations of the test statistic using the centering and scaling shown above for increasing sample sizes of $m = n =$ 10, 20, 30, 50, 70 and 100 with the true density function of the standard normal distribution $N(0,1)$ superimposed for comparison. The upper set of simulations used a decreasing bandwidth sequence of $a_N = 0.625\, N^{-\frac{1}{7}}$, while the lower set used a constant bandwidth of $a_N = 0.40$ as recommended in Behnen and Neuhaus (1989). The upper right corner of each histogram includes the empirical mean and standard deviation of the simulated samples.

For a kernel K fulfilling the assumptions (2.8) through (2.11) we choose the typically bell-shaped Parzen-2 kernel:

$$K(x) = \begin{cases} \frac{4}{3} - 8x^2 + 8|x|^3 & \text{if } |x| \le \frac{1}{2}, \\ \frac{8}{3}(1 - |x|)^3 & \text{if } \frac{1}{2} < |x| \le 1, \\ 0 & \text{if } |x| > 1. \end{cases}$$

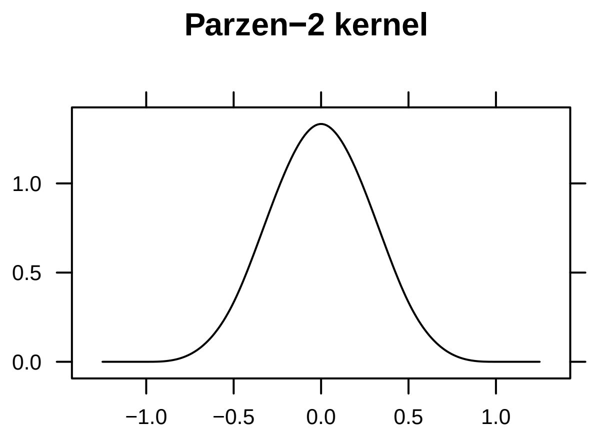

FIGURE 1. The Parzen-2 kernel.

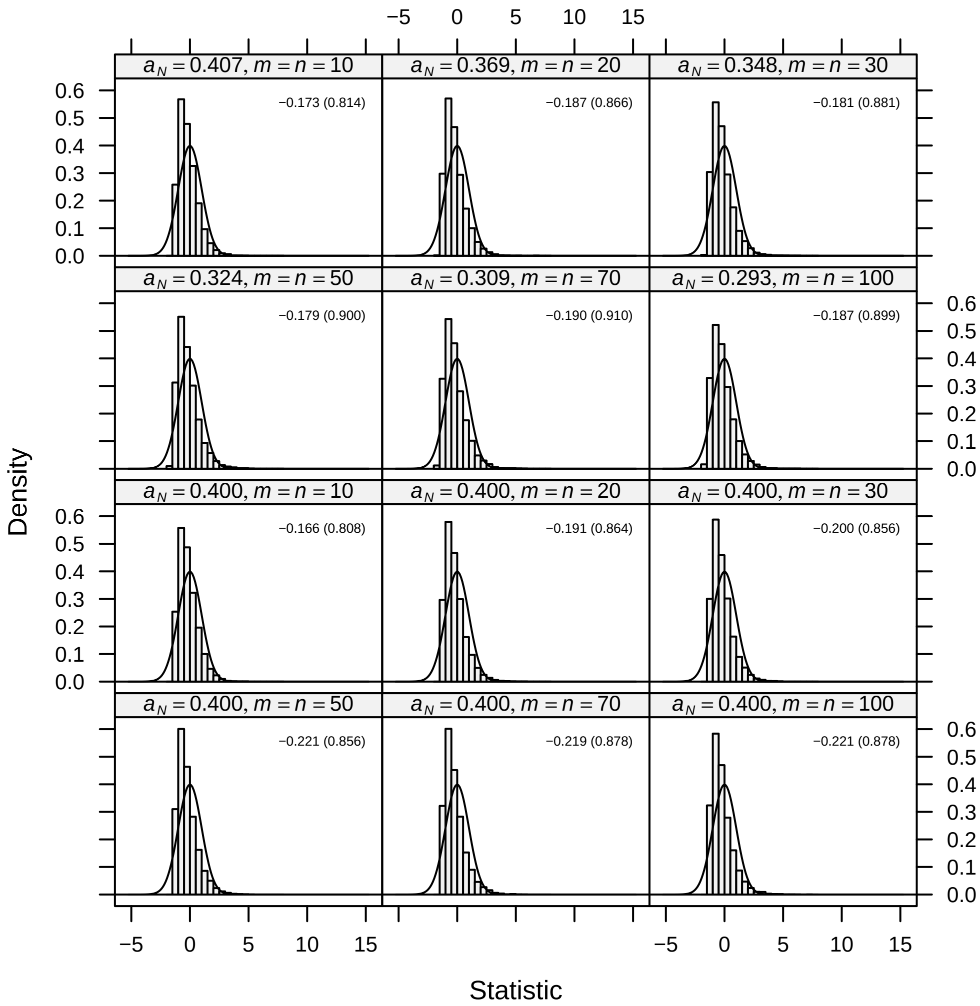

FIGURE 2. Histograms using 10,000 monte-carlo samples each of $S_N(\hat{\hat{b}}_N)$ under H_0 : $F = G$ after centering and scaling as in (3.3) with kernel density estimators (3.1) and (3.2) using the Parzen-2 kernel for sample sizes $m = n =$ 10, 20, 30, 50, 70 and 100 and bandwidth sequences $a_N = 0.625\, N^{-\frac{1}{7}}$ (upper set of graphs) and $a_N = 0.40$ (lower set of graphs). Empirical mean and standard deviation (*mean* (*sd*)) of each set of samples are given in the upper right corner, and the true standard normal density function has been superimposed for comparison.

The simulations in figure 2 clearly show problems with centering and scaling in this case, as even for quite large sample sizes of $m = n = 100$ the distribution still appears to be shifted too far to the left by the centering term $m^{-1}a_N^{-1}K(0)$, and the mean does not appear to be approaching 0 as N increases. In addition the distribution is obviously skewed to the right and the scaling factor of $m\,a_N^{\frac{1}{2}}(2\int K^2(x)\,dx)^{-\frac{1}{2}}$ seems to be overestimating the variance, although the standardized variance does appear to be moving toward 1 for very large N.

Next, we look at a similar set of simulations of the test statistic $S_N(\hat{b}_N)$ as we have proposed in chapter 2 using the simple non-restricted kernels $\hat{f}_N$ and $\hat{g}_N$ defined in (2.12) and (2.13). In this case, we have from Theorem 2.2 that

$$N^{\frac{1}{2}}a_N^{-\frac{1}{2}}\sigma_{K,\lambda}^{-1}\cdot S_N(\hat{b}_N) \xrightarrow[N]{\mathcal{L}} N(0,1)$$

under H_0 for a kernel K fulfilling (2.8) through (2.11) and a bandwidth sequence a_N fulfilling (2.5) through (2.7), and where

$$\sigma_{K,\lambda}^2 = 2\left[\lambda^{-1} + (1-\lambda)^{-1}\right]\int_{-1}^{0}\left[\int_{-1}^{x} K(v)\,dv\right]^2 dx$$

so that we use no centering terms and a standardization factor of $N^{\frac{1}{2}}a_N^{-\frac{1}{2}}\sigma_{K,\lambda}^{-1}$.

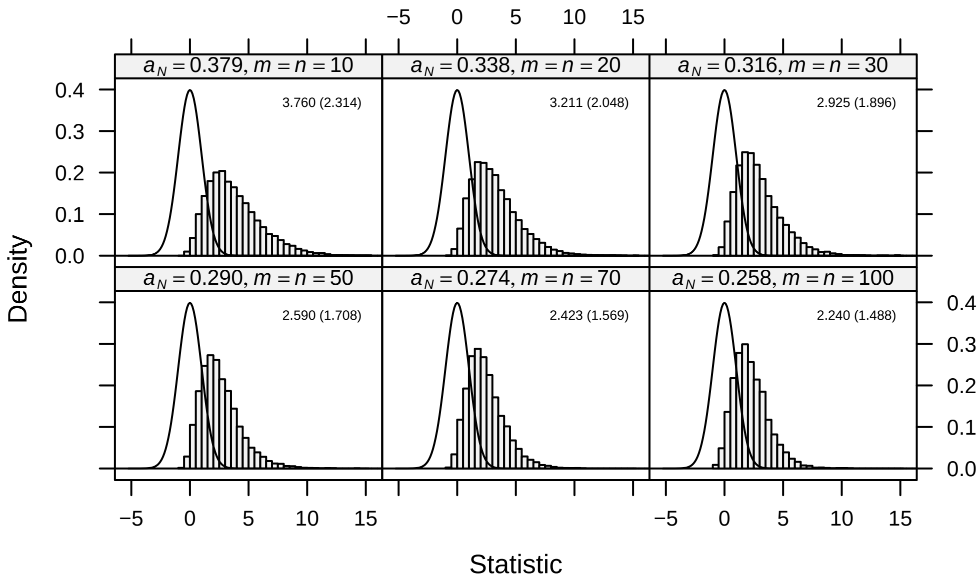

FIGURE 3. Histograms using 10,000 monte-carlo samples each of $S_N(\hat{b}_N)$ under $\mathsf{H}_0 : F = G$ after scaling with $N^{\frac{1}{2}}a_N^{-\frac{1}{2}}\sigma_{K,\lambda}^{-1}$ as in Theorem 2.2 with non-restricted kernel density estimators (2.12) and (2.13) using the Parzen-2 kernel for sample sizes $m = n = 10$, 20, 30, 50, 70 and 100 and bandwidth sequence $a_N = 0.625\,N^{-\frac{1}{6}}$. Empirical mean and standard deviation (*mean* (*sd*)) of each set of samples are given in the upper right corner, and the true standard normal density function has been superimposed for comparison.

From figure 3 we see that although the distribution of the test statistic as we have proposed in Theorem 2.2 doesn't seem to have as much of a problem with skew as the version proposed in Behnen and Hušková (1984), there are problems with centering and scaling as it is shifted much too far to the right and the scaling factor seems to be underestimating the variance for finite N. In contrast to the simulations in figure 2 there is notable improvement as N gets larger, but even for sample sizes as large as $m = n = 100$ the simulations indicate that the standard normal distribution obviously cannot be used to determine critical values or get valid p-values even for large finite N.

The centering problem is due to the construction of the sum in

$$S_N(\hat{b}_N) = m^{-1} \sum_{i=1}^{m} \Big[\hat{f}_N - \hat{g}_N\Big] \circ \hat{H}_N(X_i)$$

which requires that $\hat{f}_N$ be evaluated at each of the $\hat{H}_N(X_i)$, $1 \le i \le m$. Since $\hat{f}_N(t) = m^{-1}a_N^{-1}\sum_{j=1}^{m} K(a_N^{-1}(t - \hat{H}_N(X_j)))$ the result is a double sum over *all* $1 \le i \le m$ and $1 \le j \le m$ combinations, forcing the inclusion of the positive constant term $m^{-1}a_N^{-1}K(0)$ when $i = j$ – in total m times – leading to a positive shift in $S_N(\hat{b}_N)$ of $m^{-1}a_N^{-1}K(0)$.

This is basically a nuisance constant independent of F and G which is present under H_0 as well as H_1, doesn't contribute to the power of the test and disappears asymptotically - even after scaling - as $N \to \infty$. In this case the centering problem can be solved quickly by replacing $\hat{f}_N$ by

$$\hat{f}_N^0(t) = m(m-1)^{-1}\hat{f}_N(t) - (m-1)^{-1}a_N^{-1}K(0).$$

This drops the $i = j$ terms from the double sum mentioned above and eliminates the shift in $S_N(\hat{b}_N)$.

Using $\hat{f}_N^0$ in place of $\hat{f}_N$ in $S_N(\hat{b}_N)$, we can define

$$S_N(\hat{b}_N^0) = m^{-1} \sum_{i=1}^{m} \Big[\hat{f}_N^0 - \hat{g}_N\Big] \circ (N^{-1}R_{1i})$$

which is centered under H_0 for all finite N, since for $F = G$ we have

$$\begin{aligned}
\mathsf{E}\Big[S_N(\hat{b}_N^0)\Big] &= \mathsf{E}\Big[m^{-1}\sum_{i=1}^{m}\Big[\hat{f}_N^0(N^{-1}R_{1i}) - \hat{g}_N(N^{-1}R_{1i})\Big]\Big] \\
&= \mathsf{E}\Big[m^{-1}\sum_{i=1}^{m}\Big[m(m-1)^{-1}\hat{f}_N(N^{-1}R_{1i}) - (m-1)^{-1}a_N^{-1}K(0)\Big] - m^{-1}\sum_{i=1}^{m}\hat{g}_N(N^{-1}R_{1i})\Big] \\
&= \mathsf{E}\Big[m^{-1}\sum_{i=1}^{m}\Big[m(m-1)^{-1}m^{-1}a_N^{-1}\sum_{j=1}^{m}K\big(a_N^{-1}(\hat{H}_N(X_i) - \hat{H}_N(X_j))\big) \\
&\qquad - (m-1)^{-1}a_N^{-1}K(0)\Big] - m^{-1}\sum_{i=1}^{m}n^{-1}a_N^{-1}\sum_{k=1}^{n}K\big(a_N^{-1}(\hat{H}_N(X_i) - \hat{H}_N(Y_k))\big)\Big] \\
&= \mathsf{E}\Big[m^{-1}(m-1)^{-1}a_N^{-1}\sum_{1\le i\ne j\le m}K\big(a_N^{-1}(\hat{H}_N(X_i) - \hat{H}_N(X_j))\big) \\
&\qquad - m^{-1}n^{-1}a_N^{-1}\sum_{i=1}^{m}\sum_{k=1}^{n}K\big(a_N^{-1}(\hat{H}_N(X_i) - \hat{H}_N(Y_k))\big)\Big] \\
&= m^{-1}(m-1)^{-1}a_N^{-1}\sum_{1\le i\ne j\le m}\mathsf{E}\Big[K\big(a_N^{-1}(\hat{H}_N(X_i) - \hat{H}_N(X_j))\big)\Big] \\
&\qquad - m^{-1}n^{-1}a_N^{-1}\sum_{i=1}^{m}\sum_{k=1}^{n}\mathsf{E}\Big[K\big(a_N^{-1}(\hat{H}_N(X_i) - \hat{H}_N(Y_k))\big)\Big] \\
&= a_N^{-1}\,\mathsf{E}\Big[K\big(a_N^{-1}(\hat{H}_N(X_1) - \hat{H}_N(X_2))\big)\Big] - a_N^{-1}\,\mathsf{E}\Big[K\big(a_N^{-1}(\hat{H}_N(X_1) - \hat{H}_N(X_2))\big)\Big] \\
&= 0.
\end{aligned}$$

It is also easy to see that replacing $S_N(\hat{b}_N)$ by $S_N(\hat{b}_N^0)$ and using the scaling factor $N^{\frac{1}{2}}a_N^{-\frac{1}{2}}\sigma_{K,\lambda}^{-1}$ of Theorem 2.2 results in an asymptotically equivalent test, as

$$\begin{aligned}
&\mathsf{E}\Big[S_N(\hat{b}_N) - S_N(\hat{b}_N^0)\Big]^2 \\
&= \mathsf{E}\Big[m^{-1}\sum_{i=1}^{m}\Big[\hat{f}_N - \hat{g}_N\Big] \circ \hat{H}_N(X_i) - m^{-1}\sum_{i=1}^{m}\Big[\hat{f}_N^0 - \hat{g}_N\Big] \circ \hat{H}_N(X_i)\Big]^2 \\
&= \mathsf{E}\Big[m^{-1}\sum_{i=1}^{m}\Big[\hat{f}_N - \hat{f}_N^0\Big] \circ \hat{H}_N(X_i)\Big]^2 \\
&= \mathsf{E}\Big[m^{-1}\sum_{i=1}^{m} -(m-1)^{-1}\Big[\hat{f}_N(\hat{H}_N(X_i)) + a_N^{-1}K(0)\Big]\Big]^2 \\
&= \mathsf{E}\Big[-m^{-1}(m-1)^{-1}\sum_{i=1}^{m}\Big[m^{-1}a_N^{-1}\sum_{j=1}^{m}K\big(a_N^{-1}(\hat{H}_N(X_i) - \hat{H}_N(X_j))\big) + a_N^{-1}K(0)\Big]\Big]^2 \\
&= \mathsf{E}\Big[m^{-2}(m-1)^{-1}a_N^{-1}\sum_{i=1}^{m}\sum_{j=1}^{m}\Big[K\big(a_N^{-1}(\hat{H}_N(X_i) - \hat{H}_N(X_j))\big) + K(0)\Big]\Big]^2 \\
&\leq \mathsf{E}\Big[m^{-2}(m-1)^{-1}a_N^{-1}\sum_{i=1}^{m}\sum_{j=1}^{m}2\,\|K\|\Big]^2 \\
&= \mathsf{E}\big[2\,(m-1)^{-1}a_N^{-1}\|K\|\big]^2 \\
&= 4\,(m-1)^{-2}a_N^{-2}\|K\|^2 \\
&= O(N^{-2}a_N^{-2}),
\end{aligned}$$

and thus

$$\begin{aligned}
&N^{\frac{1}{2}}a_N^{-\frac{1}{2}}\sigma_{K,\lambda}^{-1} \cdot \Big|S_N(\hat{b}_N) - S_N(\hat{b}_N^0)\Big| \\
&= O(N^{\frac{1}{2}}a_N^{-\frac{1}{2}}) \cdot O_P(N^{-1}a_N^{-1}) \\
&= O_P(N^{-\frac{1}{2}}a_N^{-\frac{3}{2}}) \\
&= o_P(1).
\end{aligned}$$

Simulations using the modified $S_N(\hat{b}_N^0)$ with the scaling factor of $N^{\frac{1}{2}}a_N^{-\frac{1}{2}}\sigma_{K,\lambda}^{-1}$ as in Theorem 2.2 are shown in figure 4.

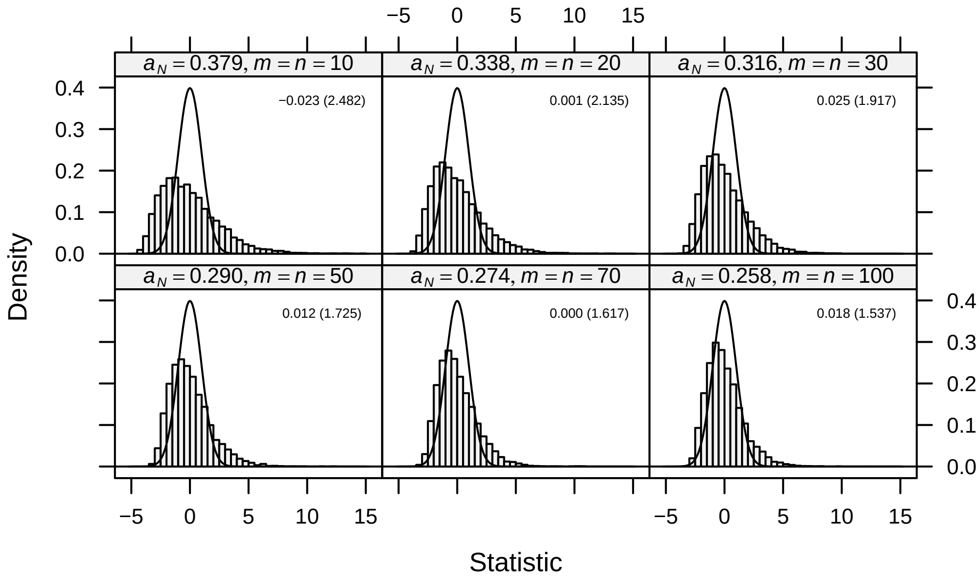

FIGURE 4. Histograms using 10,000 monte-carlo samples each of $S_N(\hat{b}_N^0)$ under $\mathsf{H}_0 : F = G$ after scaling by $N^{\frac{1}{2}} a_N^{-\frac{1}{2}} \sigma_{K,\lambda}^{-1}$ as in Theorem 2.2 with non-restricted kernel density estimators (2.12) and (2.13) using the Parzen-2 kernel for sample sizes $m = n = 10$, 20, 30, 50, 70 and 100 and bandwidth sequence $a_N = 0.625\, N^{-\frac{1}{6}}$. Empirical mean and standard deviation (*mean* (*sd*)) of each set of samples are given in the upper right corner, and the true standard normal density function has been superimposed for comparison.

The simulation results shown in figure 4 show that using $\hat{b}_N^0$ in place of $\hat{b}_N$ solves the centering problem under H_0 as desired, however there are still issues with scaling and distributional convergence as the distribution is still fairly skewed to the right even for large N.

In order to try to isolate the source of the slow convergence and skew, we begin by simulating monte-carlo samples of the non-negligible terms that are responsible for the asymptotic distribution of the statistic S_N under H_0. That is, since we know from the proof of Theorem 2.2 that under H_0

$$S_N(\hat{b}_N) = \int \bar{f}_N \circ H_N(x) \Big[\hat{F}_m(dx) - F(dx)\Big] \tag{2.23}$$

$$- \int \bar{f}_N \circ H_N(x) \Big[\hat{G}_n(dx) - G(dx)\Big] \tag{2.24}$$

$$+ O_P(N^{-1} a_N^{-2})$$

we generate monte-carlo samples of terms (2.23) and (2.24) under H_0 alone without the asymptotically negligible rest parts of the statistic to see whether the structure of these terms is the source of the scaling and skew problems.

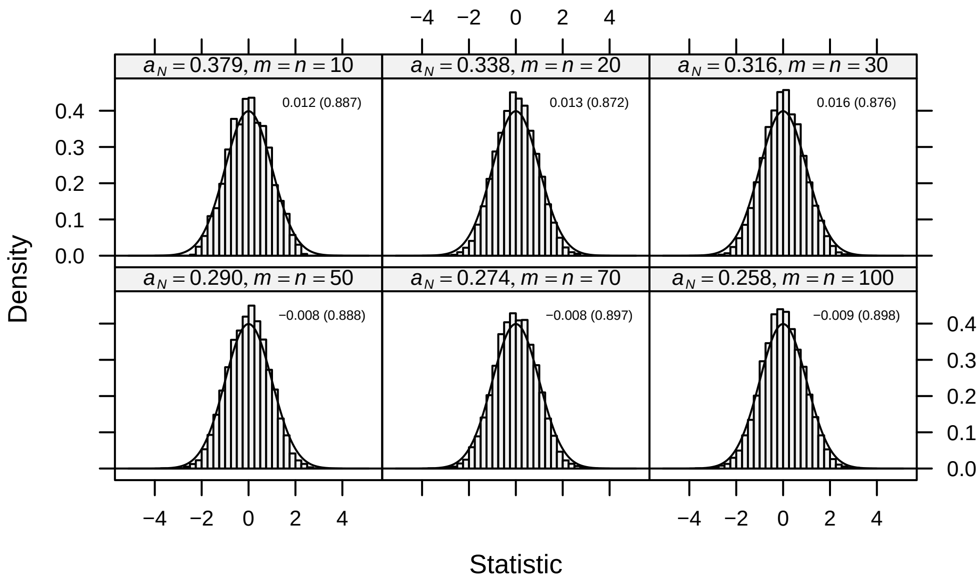

FIGURE 5. Histograms using 10,000 monte-carlo samples each of terms (2.23) and (2.24) of S_N under $\mathsf{H}_0 : F = G$ after scaling by $N^{\frac{1}{2}} a_N^{-\frac{1}{2}} \sigma_{K,\lambda}^{-1}$ as in Theorem 2.2 with non-restricted kernel density estimators (2.12) and (2.13) using the Parzen-2 kernel for sample sizes $m = n =$ 10, 20, 30, 50, 70 and 100 and bandwidth sequence $a_N = 0.625\, N^{-\frac{1}{6}}$. Empirical mean and standard deviation (*mean* (*sd*)) of each set of samples are given in the upper right corner, and the true standard normal density function has been superimposed for comparison.

From figure 5 we see that the distribution of (2.23) and (2.24) is centered and symmetric, but that scaling using the asymptotic variance in Theorem 2.2 does seem to be overestimating variance for small N and convergence to 1 appears quite slow.

Looking at the proof of Theorem 2.2 we see that the variance of (2.23) and (2.24) under H_0 for finite N is actually

$$\sigma_N^2 = N^{-1} a_N \left[\lambda_N^{-1} + (1-\lambda_N)^{-1}\right] \cdot \left[2\int_{-1}^{0}\left[\int_{-1}^{x} K(v)\,dv\right]^2 dx - 4\,a_N\left[\int_0^1 vK(v)\,dv\right]^2\right].$$

The term $4\,a_N[\int_0^1 vK(v)\,dv]^2$ is vanishing, since $a_N \to 0$, and thus doesn't play a role in the asymptotic variance shown in Theorem 2.2. However, the bandwidth sequence a_N is required to converge to zero quite slowly ($Na_N^5 \to \infty$), so that this term does still play an important role in the variance of S_N even for large finite N, and failing to include it in the expression above leads to the overestimation of variance seen in figure 5.

In order to confirm this, we can simulate the distribution of (2.23) and (2.24), as in figure 5, this time scaling by

$$\sigma_N^{-1} = N^{\frac{1}{2}} a_N^{-\frac{1}{2}} \left[\left[\lambda_N^{-1} + (1-\lambda_N)^{-1}\right] \cdot \left[2\int_{-1}^{0}\left[\int_{-1}^{x} K(v)\,dv\right]^2 dx - 4\,a_N\left[\int_0^1 vK(v)\,dv\right]^2\right]\right]^{-\frac{1}{2}}. \tag{3.4}$$

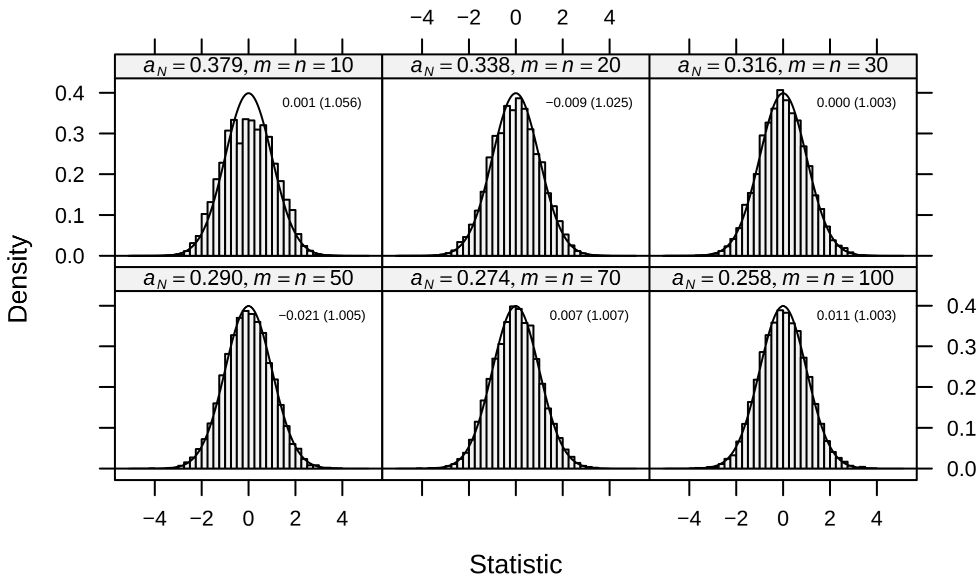

FIGURE 6. Histograms using 10,000 monte-carlo samples each of terms (2.23) and (2.24) of S_N under $\mathsf{H}_0 : F = G$ after scaling by σ_N^{-1} as in (3.4) with non-restricted kernel density estimators (2.12) and (2.13) using the Parzen-2 kernel for sample sizes $m = n = 10$, 20, 30, 50, 70 and 100 and bandwidth sequence $a_N = 0.625\, N^{-\frac{1}{6}}$. Empirical mean and standard deviation (*mean* (*sd*)) of each set of samples are given in the upper right corner, and the true standard normal density function has been superimposed for comparison.

After using the modified variance estimate σ_N^2 including the term $4\, a_N[\int_0^1 vK(v)\, dv]^2$ we see from the results in figure 6 that σ_N^2 gives a correct variance for the terms (2.23) and (2.24) responsible for the distribution of $S_N(\hat{b}_N^0)$ under H_0 even for small N.

This leads us to look at simulations of the full sum $S_N(\hat{b}_N^0)$ scaled by the modified σ_N^{-1}, which are included in figure 7.

The results in figure 7 demonstrate that using the corrected variance estimate σ_N^2 when scaling the complete statistic S_N doesn't bring the same dramatic improvement as far as scaling as it does when used with the asymptotically relevant terms (2.23) and (2.24), and skew is, of course, unaffected by altering the scaling factor so that this problem remains as well.

From the results in figures 6 and 7, we must conclude that the source of the scaling and skew problems is found in the asymptotically negligible terms of $\sigma_N^{-1} S_N(\hat{b}_N^0)$. These were shown in the proof of Theorem 2.2 to be $O_P(N^{-\frac{1}{2}} a_N^{-\frac{5}{2}})$, which is asymptotically negligible, since we require that $Na_N^5 \to \infty$, but convergence can, in real applications, be quite slow, so that these terms still play an important role in the distribution of $\sigma_N^{-1} S_N$ even for large finite N.

By simulations analogous to figure 8 for each of the negligible terms (2.35), (2.36), (2.45) making up the asymptotically negligible $O_{\mathbb{P}}(N^{-1} a_N^{-2})$ part of S_N that does not vanish under H_0, we can isolate (2.35) as the term that is vanishing so slowly (see figure 8) that its influence on the distribution of the test statistic is still important for finite N.

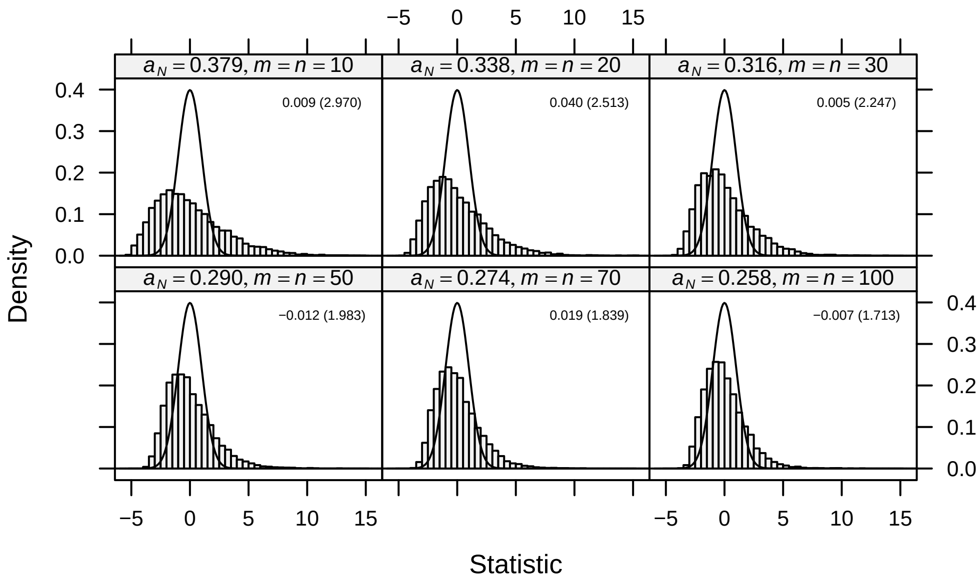

FIGURE 7. Histograms using 10,000 monte-carlo samples each of $S_N(\hat{b}_N^0)$ under $\mathsf{H}_0 : F = G$ after scaling by σ_N^{-1} with non-restricted kernel density estimators (2.12) and (2.13) using the Parzen-2 kernel for sample sizes $m = n = 10$, 20, 30, 50, 70 and 100 and bandwidth sequence $a_N = 0.625\,N^{-\frac{1}{6}}$. Empirical mean and standard deviation (*mean* (*sd*)) of each set of samples are given in the upper right corner, and the true standard normal density function has been superimposed for comparison.

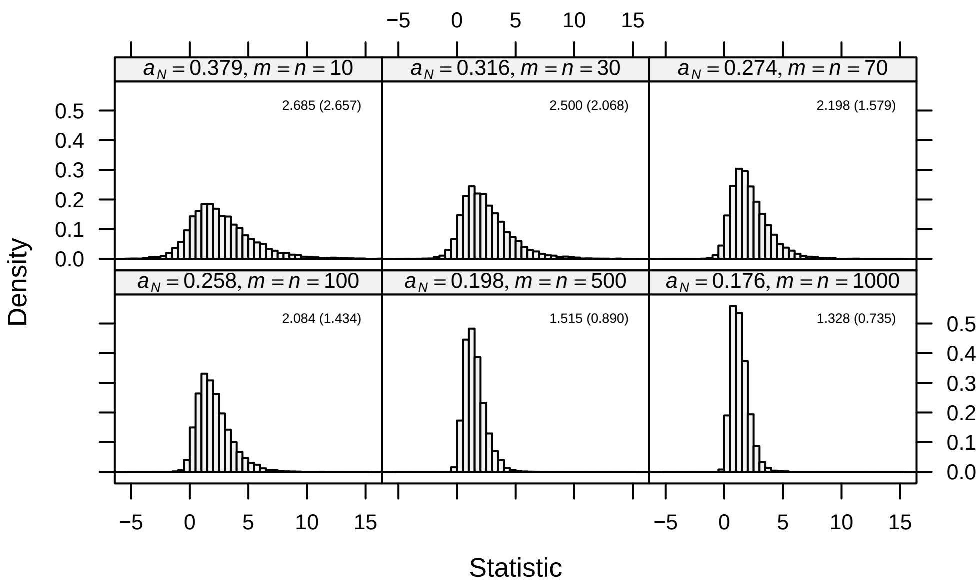

FIGURE 8. Histograms using 10,000 monte-carlo samples each of (2.35) under $\mathsf{H}_0 : F = G$ after scaling by $N^{\frac{1}{2}} a_N^{-\frac{1}{2}} \sigma_{K,\lambda}^{-1}$ as in Theorem 2.2 with non-restricted kernel density estimators (2.12) and (2.13) using the Parzen-2 kernel for sample sizes $m = n = 10$, 30, 100, 500 and 1000 and bandwidth sequence $a_N = 0.625\,N^{-\frac{1}{6}}$. Empirical mean and standard deviation (*mean* (*sd*)) of each set of samples are given in the upper right corner.

The results in figure 8 show how (2.35) contributes to the variance and skew of the distribution of S_N, and that although it is vanishing as $N \to \infty$, actual convergence is very slow with sizable variance even for sample sizes as large as $m = n = 1000$.

Looking more closely at the form of (2.35) under H_0 we find

$$
\begin{aligned}
&\int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big] \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big] \\
&= \int \Big[\hat{f}_N - \hat{g}_N\Big] \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big] \\
&= \int \Big[m^{-1} a_N^{-1} \sum_{j=1}^{m} K\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_j))\big) \\
&\qquad - n^{-1} a_N^{-1} \sum_{k=1}^{n} K\big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))\big)\Big]\Big[\hat{F}_m(dx) - F(dx)\Big] \\
&= \int m^{-1} a_N^{-1} \sum_{j=1}^{m} K\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_j))\big)\Big[\hat{F}_m(dx) - F(dx)\Big] \\
&\quad - \int n^{-1} a_N^{-1} \sum_{k=1}^{n} K\big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))\big)\Big[\hat{F}_m(dx) - F(dx)\Big] \\
&= m^{-1} \sum_{i=1}^{m} \Big[m^{-1} a_N^{-1} \sum_{j=1}^{m} K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(X_j))\big) \\
&\qquad - \int m^{-1} a_N^{-1} \sum_{j=1}^{m} K\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_j))\big)\, F(dx)\Big] \\
&\quad - m^{-1} \sum_{i=1}^{m} \Big[n^{-1} a_N^{-1} \sum_{k=1}^{n} K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(Y_k))\big) \\
&\qquad - \int n^{-1} a_N^{-1} \sum_{k=1}^{n} K\big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))\big)\, F(dx)\Big] \\
&= m^{-2} a_N^{-1} \sum_{i=1}^{m}\sum_{j=1}^{m} \Big[K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(X_j))\big) - \int K\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_j))\big)\, F(dx)\Big] \\
&\quad - m^{-1} n^{-1} a_N^{-1} \sum_{i=1}^{m}\sum_{k=1}^{n} \Big[K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(Y_k))\big) - \int K\big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))\big)\, F(dx)\Big] \\
&= m^{-2} a_N^{-1} \sum_{i=1}^{m}\sum_{j=1}^{m} \Big[K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(X_j))\big) - \int K_0^1\big(a_N^{-1}(v - \hat{H}_N(X_j))\big)\, dv\Big] \\
&\quad - m^{-1} n^{-1} a_N^{-1} \sum_{i=1}^{m}\sum_{k=1}^{n} \Big[K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(Y_k))\big) - \int_0^1 K\big(a_N^{-1}(v - \hat{H}_N(Y_k))\big)\, dv\Big].
\end{aligned}
$$

From this, we see that the first sum above making up (2.35) comprises summands of the form

$$
K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(X_j))\big) - \int_0^1 K\big(a_N^{-1}(v - \hat{H}_N(X_j))\big)\, dv
$$

which is simply the difference between a kernel with bandwidth a_N centered at $\hat{H}_N(X_j) = N^{-1}R_{1j}$ evaluated at $H_N(X_i)$ and the area under the same kernel contained within the interval $[0, 1]$. The form of these summands turns out to be the source of the right skew and slow convergence to 0 of (2.35).

In samples where, for example, the X_i occupy most of the smaller positions in the total sample (i.e. where almost all R_{1i} are smaller than the R_{2k}) large portions of many of the kernels $K(a_N^{-1}(t - N^{-1}R_{1j}))$ will not be contained on $[0,1]$ making $\int_0^1 K(a_N^{-1}(v - N^{-1}R_{1j}))\,dv$ small while at the same time many of the $H_N(X_i)$ will be close to the centers of the bell-shaped kernels at $N^{-1}R_{1j}$ where they reach their maximum making the $K(a_N^{-1}(H_N(X_i) - N^{-1}R_{1j}))$ large. The effect when X_i occupy most of the larger positions in the total sample is the same by analogy. This allows (2.35) to become quite large and disappear slowly, since such samples occur with some probability even under H_0. Reducing the bandwidth a_N in order to allow more of the kernels $K(a_N^{-1}(t - N^{-1}R_{1j}))$ to be contained on $[0,1]$ unfortunately doesn't improve the situation, since a_N^{-1} is a factor in the sum as well leading immediately to kernels with higher peaks, which can exacerbate the problem detailed above.

Since the convergence problem is, in essence, caused by the relative difference between the maximum height of the bell-shaped K at its peak and areas like $\int_0^1 K(a_N^{-1}(v - N^{-1}R_{1j}))\,dv$, we can attempt to reduce these differences and improve convergence by switching from a fairly steep bell-shaped kernel like the Parzen-2 kernel to a much flatter K that still fulfills (2.8) through (2.11). For this purpose, we introduce a parametric family of kernels of the following form which reach a value of γ at $x = 0$, while still fulfilling $\int_0^1 K(x)\,dx = 1$:

$$K_{\beta,\gamma}(x) = \begin{cases} (1-\gamma(1-\beta)-\gamma)\cdot\left[120\,\beta^{-5}\,k_\beta(|x|\cdot\beta(1-\beta)^{-1})+1\right]+\gamma & \text{if } |x|\leq 1-\beta \\ -120\,\beta^{-5}\,(1-\gamma(1-\beta))\cdot k_\beta(|x|-(1-\beta)) & \text{if } 1-\beta<|x|\leq 1 \\ 0 & \text{if } |x|>1 \end{cases} \tag{3.5}$$

with

$$k_\beta(x) = \frac{x^5}{20} - \frac{\beta\,x^4}{8} + \frac{\beta^2\,x^3}{12} - \frac{\beta^5}{120}.$$

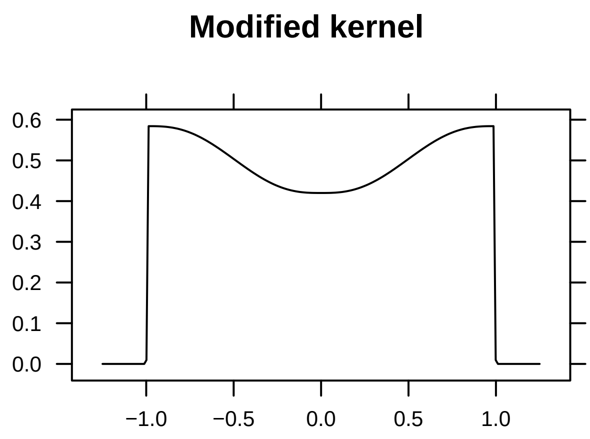

FIGURE 9. Flat kernel defined in (3.5).

The results of using the flatter kernel $K_{\beta,\gamma}$ in place of the bell-shaped Parzen-2 kernel in the centered $S_N(\hat{b}_N^0)$ scaled by $N^{\frac{1}{2}}a_N^{-\frac{1}{2}}\sigma_{K,\lambda}$ are shown in figure 10.

As predicted, figure 10 shows improvement even for small sample sizes with barely noticeable skew, and scaling by $N^{\frac{1}{2}}a_N^{-\frac{1}{2}}\sigma_{K,\lambda}^{-1}$ seem to actually be over-estimating the variance. Switching to the improved variance estimate of σ_N^2 seems to significantly under-estimate the variance in this case (see figure 11).

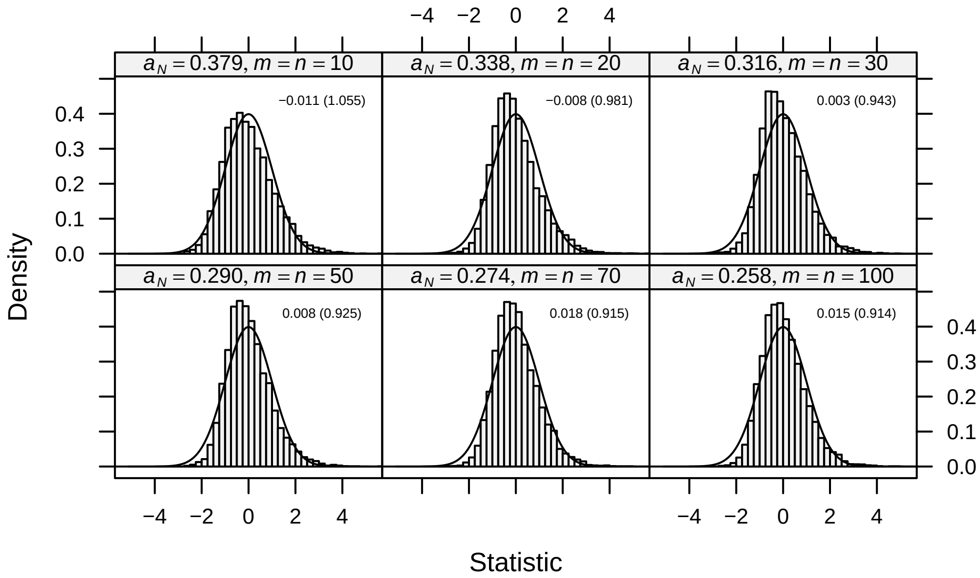

FIGURE 10. Histograms using 10,000 monte-carlo samples each of $S_N(\hat{b}_N^0)$ under H_0 : $F = G$ after scaling by $N^{\frac{1}{2}} a_N^{-\frac{1}{2}} \sigma_{K,\lambda}^{-1}$ as in Theorem 2.2 with non-restricted kernel density estimators (2.12) and (2.13) using the modified flattened kernel $K_{\beta,\gamma}$ with $\beta = 0.01$ and $\gamma = 0.42$ for sample sizes $m = n = 10$, 20, 30, 50, 70 and 100 and bandwidth sequence $a_N = 0.625\, N^{-\frac{1}{6}}$. Empirical mean and standard deviation (*mean* (*sd*)) of each set of samples are given in the upper right corner, and the true standard normal density function has been superimposed for comparison.

Since we know that the distribution of $S_N(\hat{b}_N^0)$ is determined under H_0 for finite N by the terms (2.23), (2.24) and (2.35) we can try to find a more accurate variance estimate for $S_N(\hat{b}_N^0)$ by attempting to incorporate the variance of (2.35) for finite N even though this term is asymptotically negligible. In order to do this, define σ_{2N}^2 as the combined variance of (2.23) and (2.24) and the theoretical analog of (2.35). That is, let

$$\sigma_{2N}^2 = \operatorname{Var}\Big[\int \bar{f}_N \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big] - \int \bar{f}_N \circ H_N(x)\Big[\hat{G}_n(dx) - G(dx)\Big] \tag{3.6}$$

$$+ m^{-2} a_N^{-1} \sum_{i=1}^{m}\sum_{j=1}^{m}\Big[K\big(a_N^{-1}(H_N(X_i) - H_N(X_j))\big) - \int K\big(a_N^{-1}(H_N(x) - H_N(X_j))\big)\, F(dx)\Big] \tag{3.7}$$

$$- m^{-1} n^{-1} a_N^{-1} \sum_{i=1}^{m}\sum_{k=1}^{n}\Big[K\big(a_N^{-1}(H_N(X_i) - H_N(Y_k))\big) - \int K\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx)\Big]\Big] \tag{3.8}$$

under H_0.

We already know from lemmas 5.33 and 5.34 that the variance of (3.6) is equal to σ_N^2. Lemma 5.36 shows that the covariance between (3.6) and (3.7) (3.8) vanishes under H_0 for all N and lemma 5.35 gives the variance of (3.7) and (3.8) under H_0, so that combining these results we have

$$\sigma_{2N}^2 = \sigma_N^2 + m^{-1}(m-1)^{-1}\Big[\Big[a_N^{-1}\int_{-1}^{1} K^2(v)\, dv - 2\int_0^1 vK^2(v)\, dv\Big]$$
$$+ (2n + m - 1)n^{-1}\Big[1 - 4\, a_N \int_0^1 vK(v)\, dv + 4\, a_N^2\Big[\int_0^1 vK(v)\, dv\Big]^2\Big]$$

$$- (1 + 2n^{-1}) \left[1 + 2\, a_N \int_{-1}^{0} \left[\int_{-1}^{x} K(v)\, dv \right]^2 dx - 4\, a_N \int_{0}^{1} vK(v)\, dv \right] \right] \qquad (3.9)$$

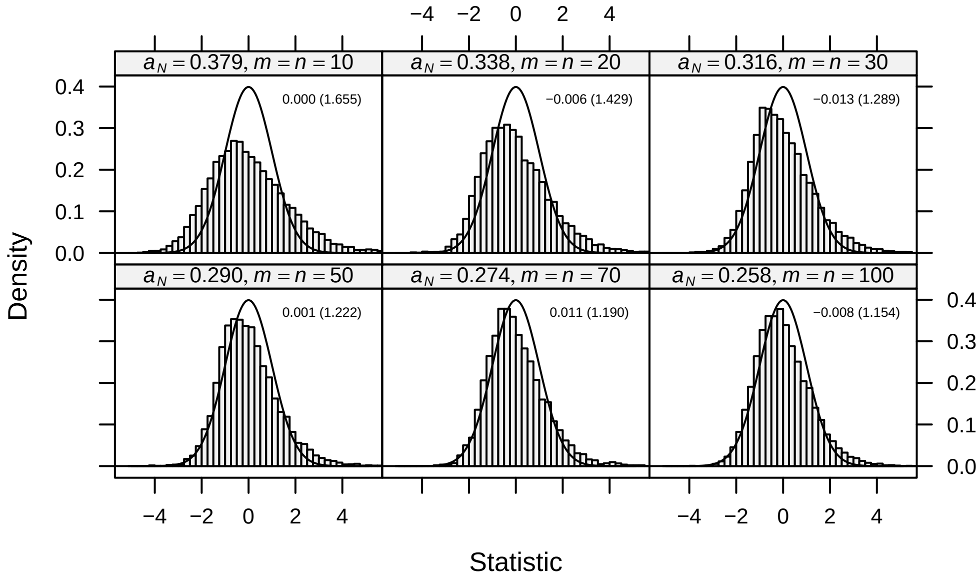

FIGURE 11. Histograms using 10,000 monte-carlo samples each of $S_N(\hat{b}_N^0)$ under H_0 : $F = G$ after scaling by σ_N^{-1} with non-restricted kernel density estimators (2.12) and (2.13) using the modified flattened kernel $K_{\beta,\gamma}$ with $\beta = 0.01$ and $\gamma = 0.42$ for sample sizes $m = n = 10$, 20, 30, 50, 70 and 100 and bandwidth sequence $a_N = 0.625\, N^{-\frac{1}{6}}$. Empirical mean and standard deviation (*mean* (*sd*)) of each set of samples are given in the upper right corner, and the true standard normal density function has been superimposed for comparison.

Using σ_{2N}^{-1} in place of σ_N^{-1} to scale $S_N(\hat{b}_N^0)$ together with flatter kernels of the form $K_{\beta,\gamma}$ leads to a statistic that is much improved as far as skew and scaling across a broad range of sample sizes (see figure 12), so that $N(0, 1)$ could plausibly be considered for calculating critical values and p-values as desired, however, as the simulations in chapter 4 will show, the flatter kernels can lead to a substantial loss of power.

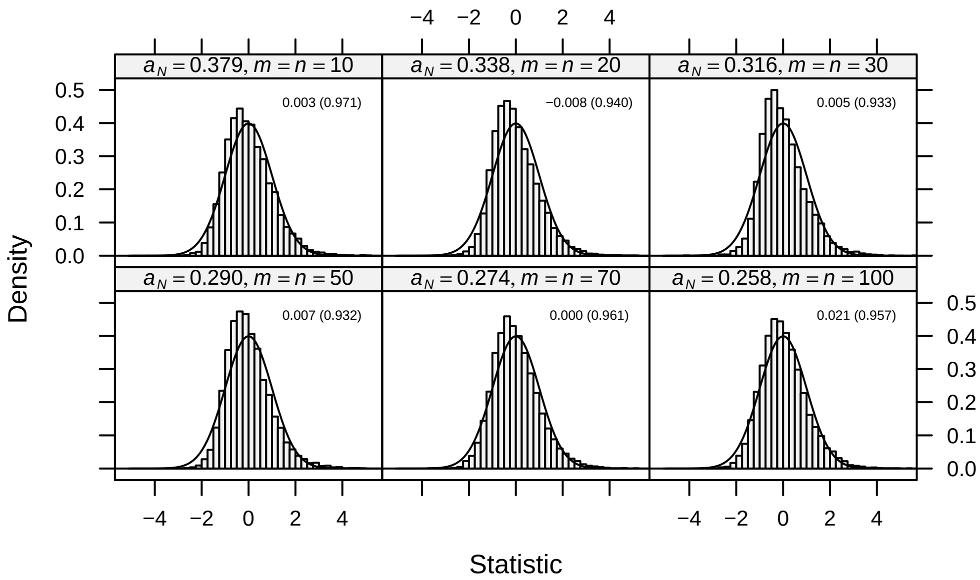

FIGURE 12. Histograms using 10,000 monte-carlo samples each of $S_N(\hat{b}_N^0)$ under H_0 : $F = G$ after scaling by σ_{2N}^{-1} with non-restricted kernel density estimators (2.12) and (2.13) using the modified flattened kernel $K_{\beta,\gamma}$ with $\beta = 0.01$ and $\gamma = 0.42$ for sample sizes $m = n = 10$, 20, 30, 50, 70 and 100 and bandwidth sequence $a_N = 0.625\,N^{-\frac{1}{6}}$. Empirical mean and standard deviation (*mean* (*sd*)) of each set of samples are given in the upper right corner, and the true standard normal density function has been superimposed for comparison.

CHAPTER 4

A simulation study

In the following we give the results of a series of simulations using different implementations of the rank statistic S_N with varying choices regarding the adaptive score function, scaling, kernel function K and bandwidth sequence a_N (see table 3). Of main interest will be comparisons between the statistic $S_N(\hat{\hat{b}}_N)$ using restricted kernel estimators (see (3.1) and (3.2)) as proposed by Behnen et al. (1983) and the modified statistic $S_N(\hat{b}_N^0)$ as proposed in chapter 3 with scaling using the improved variance estimate σ_{2N}^2 given in (3.9). We also include simulations using the fixed bandwidth sequence $a_N = 0.4$ as recommended in Behnen and Neuhaus (1989).

Since the simulations under H_0 in chapter 3 clearly showed in almost all cases that the standard normal distribution cannot be used to set valid critical values or calculate p-values, except where otherwise noted critical values were determined either by calculating the exact distribution of the test statistic for small sample sizes ($m = n = 10$) or by first using a set of 100,000 monte-carlo replications of the test statistic under H_0 to determine monte-carlo critical values for larger sample sizes ($m = n = 20$ or 30).

Table 4 shows the rejection rates of the various tests under H_0. To explore the power of the proposed tests under different kinds of non-trivial alternatives, we follow along the lines of Behnen and Neuhaus (1989) and consider monte-carlo simulations under a collection of generalized shift alternatives that include the classical exact shift model as well as alternatives that concentrate the shift between F and G in the lower, central or upper part of the distribution (see figure 1).

lower shift	$G(x) = F(x - (1 - F(x)))$
central shift	$G(x) = F(x - 4F(x)(1 - F(x)))$
upper shift	$G(x) = F(x - F(x))$
exact shift	$G(x) = F(x - 1)$

Table 1. Distribution functions of the lower, central, upper and exact shift alternatives for an underlying distribution function F.

While the alternative G resulting from an exact shift is always a valid distribution function, this is not immediately obvious for the other three generalized shifts. In the case of the lower, central and upper shifts we see that as continous functions of the distribution function F, each of the alternative G are right continuous with left limits, and that

$$\lim_{x\to-\infty} G(x) = 0 \qquad \text{and} \qquad \lim_{x\to\infty} G(x) = 1,$$

since

$$\lim_{x\to\infty} x - (1 - F(x)) = \infty$$

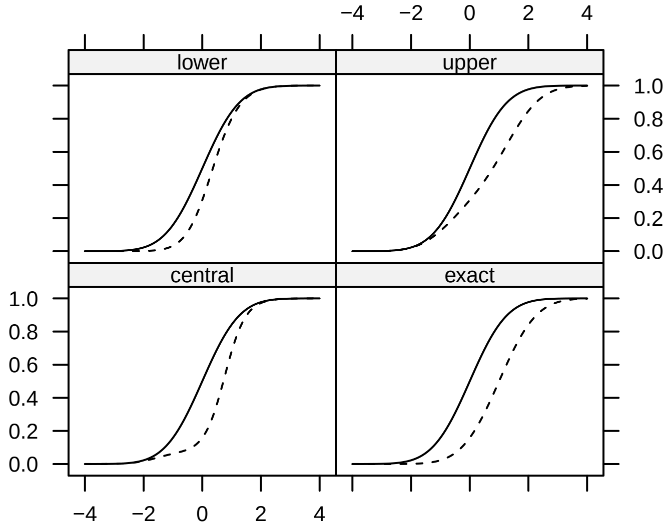

FIGURE 1. Cumulative distribution functions illustrating the exact, lower, central and upper shifts (dashed line) for underlying standard normal F (solid line).

$$\lim_{x\to\infty} x - 4F(x)(1-F(x)) = \infty$$

$$\lim_{x\to\infty} x - F(x) = \infty,$$

and

$$\lim_{x\to-\infty} x - (1-F(x)) = -\infty$$

$$\lim_{x\to-\infty} x - 4F(x)(1-F(x)) = -\infty$$

$$\lim_{x\to-\infty} x - F(x) = -\infty.$$

Thus, the generalized shift functions $1-F(x)$, $4F(x)(1-F(x))$ and $F(x)$ will yield valid distribution functions in the alternatives as long as we can ensure nondecreasing monotonicity of the resulting shifted G. For the lower shift using $1-F(x)$ this is always the case, since $x-(1-F(x))$ is monotonically nondecreasing for any distribution function F.

In the case of the central and upper shifts, we can make certain the shift functions are not increasing too quickly by requiring that F be continuous with Lebesgue-density F' such that

$$\sup_x F'(x) \le 1, \tag{4.1}$$

$$\sup_x F'(x)(1-2F(x)) \le \frac{1}{4}. \tag{4.2}$$

Then for the upper shift we have for $x_1 \le x_2$

$$F(x_2) - F(x_1) = \int_{x_1}^{x_2} F'(u)\,du \le \int_{x_1}^{x_2} 1\,du = x_2 - x_1,$$

so that

$$x_1 - F(x_1) \le x_2 - F(x_2)$$

which ensures nondecreasing monotonicity of $F(x-F(x))$.

And in the case of the central shift with $G(x) = F(x - 4F(x)(1 - F(x)))$ we have for $x_1 \le x_2$

$$\begin{aligned}
4\big[F(x_2)(1 - F(x_2)) - F(x_1)(1 - F(x_1))\big] &= 4\int_{x_1}^{x_2} \big[F'(u)(1 - F(u)) + F(u)(-F'(u))\big]\,du \\
&= 4\int_{x_1}^{x_2} F'(u)(1 - 2F(u))\,du \\
&\le 4\int_{x_1}^{x_2} \frac{1}{4}\,du \\
&= x_2 - x_1,
\end{aligned}$$

so that

$$x_1 - 4F(x_1)(1 - F(x_1)) \le x_2 - 4F(x_2)(1 - F(x_2))$$

which ensures monotonicity of $F(x - 4F(x)(1 - F(x)))$.

For the underlying distribution function F we use the standard normal $N(0,1)$, Logistic$(0,1)$ and Cauchy$(0,1)$ distributions (see table 2). (4.1) is easily verified for these F, since their densities are symmetric about 0, attaining a maximum $F'(0)$ which is less than 1.

When verifying (4.2), we once again use the fact that each of the underlying F' are bounded by their maximum at $F'(0)$.

For Logistic$(0,1)$ we actually have

$$F'(0) = \frac{1}{4},$$

so that (4.2) is fulfilled immediately, as $1 - 2F(x) \le 1$ everywhere.

For $N(0,1)$ and Cauchy$(0,1)$ we only need to be concerned with x such that $x < F^{-1}(\frac{1}{2} - \frac{1}{8}F'(0)^{-1})$, since for $x \ge F^{-1}(\frac{1}{2} - \frac{1}{8}F'(0)^{-1})$ we have

$$\begin{aligned}
F'(x)(1 - 2F(x)) &\le F'(x)\left[1 - 2F\left(F^{-1}\left(\frac{1}{2} - \frac{1}{8}F'(0)^{-1}\right)\right)\right] \\
&\le F'(0)\left[1 - 2\left(\frac{1}{2} - \frac{1}{8}F'(0)^{-1}\right)\right] \\
&= F'(0)\,\frac{1}{4}F'(0)^{-1} \\
&= \frac{1}{4}.
\end{aligned}$$

For any x such that $F'(x) \le \frac{1}{4}$ we see that (4.2) is fulfilled as well, since $1 - 2F(x) \le 1$ for all x. This means that in the case of distributions such as $N(0,1)$ and Cauchy$(0,1)$ whose densities are monotonically increasing on the interval $(-\infty, F^{-1}(0))$, (4.2) is fulfilled, when we can verify that the bound in (4.2) holds for any x such that

$$\inf\left\{x : F'(x) \ge \frac{1}{4}\right\} < x < F^{-1}\left(\frac{1}{2} - \frac{1}{8}F'(0)^{-1}\right). \tag{4.3}$$

As $1 - 2F(x)$ is monotonically nonincreasing everywhere and the three underlying densities used here are monotonically increasing on the interval $(-\infty, F^{-1}(0))$, we know that on the interval (4.3)

$$F'(x)(1 - 2F(x)) \le F'\left(F^{-1}\left(\frac{1}{2} - \frac{1}{8}F'(0)^{-1}\right)\right)\cdot\left[1 - 2F\left(\inf\left\{x : F'(x) \ge \frac{1}{4}\right\}\right)\right],$$

which gives us an easy way to check (4.2).

For $N(0,1)$ we have

$$\inf\left\{x : F'(x) \geq \frac{1}{4}\right\} \approx -0.96664 \qquad \text{and} \qquad F^{-1}\left(\frac{1}{2} - \frac{1}{8}F'(0)^{-1}\right) \approx -0.890229,$$

so that

$$\begin{aligned} &F'\left(F^{-1}\left(\frac{1}{2} - \frac{1}{8}F'(0)^{-1}\right)\right) \cdot \left[1 - 2F\left(\inf\left\{x : F'(x) \geq \frac{1}{4}\right\}\right)\right] \\ &\approx F'(-0.8902299) \cdot (1 - 2F(-0.96664)) \\ &\approx 0.26842 \cdot 0.66628 \\ &\leq \frac{1}{4}. \end{aligned}$$

And for Cauchy$(0,1)$ we have

$$\inf\left\{x : F'(x) \geq \frac{1}{4}\right\} \approx -0.5225 \qquad \text{and} \qquad F^{-1}\left(\frac{1}{2} - \frac{1}{8}F'(0)^{-1}\right) \approx -2.85329,$$

so that the interval in (4.3) is empty and (4.2) holds, since for all x either $x \geq F^{-1}(\frac{1}{2} - \frac{1}{8}F'(0)^{-1})$ or $F'(x) \leq \frac{1}{4}$.

$N(0,1)$	$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{1}{2}\,y^2\right)\,dy$
Logistic$(0,1)$	$F(x) = \frac{\exp(x)}{1+\exp(x)}$
Cauchy$(0,1)$	$F(x) = \frac{1}{2} + \frac{1}{\pi}\arctan(x)$

TABLE 2. Underlying distribution functions F used with each of the lower, central, upper and exact shift alternatives.

Tables 5 through 7 and figures 2 through 4 give empirical rejection rates under these alternatives for the test statistics as defined in table 3 using nominal type I error probabilities $\alpha = 0.01, 0.02, \ldots, 0.10$ on the basis of 10,000 replications each. The non-adaptive Wilcoxon rank-sum test has been included as well for comparison.

Legend	Score function	Scaling factor	K	a_N	Method
S_1	$\hat{\hat{b}}_N$	$ma_N^{\frac{1}{2}}[2\int K^2(x)\,dx]^{-\frac{1}{2}}$	Parzen-2	$0.625\,x^{-\frac{1}{7}}$	exact ($m = n = 10$) or monte-carlo ($m = n = 20, 30$)
S_2	$\hat{\hat{b}}_N$	$ma_N^{\frac{1}{2}}[2\int K^2(x)\,dx]^{-\frac{1}{2}}$	Parzen-2	0.40	exact ($m = n = 10$) or monte-carlo ($m = n = 20, 30$)
S_3	$\hat{b}_N^0$	$N^{\frac{1}{2}}a_N^{-\frac{1}{2}}\sigma_{K,\lambda}^{-1}$	Parzen-2	$0.625\,x^{-\frac{1}{6}}$	exact ($m = n = 10$) or monte-carlo ($m = n = 20, 30$)
S_4	$\hat{b}_N^0$	σ_{2N}^{-1}	$K_{\gamma,\beta}$	$0.625\,x^{-\frac{1}{6}}$	exact ($m = n = 10$) or monte-carlo ($m = n = 20, 30$)
S_5	$\hat{b}_N^0$	σ_{2N}^{-1}	$K_{\gamma,\beta}$	$0.625\,x^{-\frac{1}{6}}$	asymptotic
S_6	Rank-sum test				exact ($m = n = 10$) or asymptotic ($m = n = 20, 30$)

TABLE 3. Score functions, scaling, kernel functions K, bandwidth sequences a_N and method of generating critical values for the test statistics $S_1, S_2, \dots, S_6$ included in the simulation study. For the kernel $K_{\gamma,\beta}$ we used $\gamma = 0.42,\ \beta = 0.01$.

		Test					
m, n	α	S_1	S_2	S_3	S_4	S_5	S_6
10	0.01	0.009	0.009	0.009	0.009	0.012	0.007
	0.05	0.047	0.047	0.048	0.048	0.043	0.044
	0.10	0.098	0.098	0.098	0.100	0.079	0.091
20	0.01	0.010	0.010	0.010	0.002	0.012	0.009
	0.05	0.048	0.048	0.049	0.017	0.039	0.049
	0.10	0.098	0.098	0.099	0.054	0.071	0.101
30	0.01	0.012	0.012	0.012	0.002	0.013	0.010
	0.05	0.050	0.051	0.051	0.024	0.038	0.055
	0.10	0.100	0.102	0.100	0.062	0.069	0.108

TABLE 4. Rates of rejection for the test statistics $S_1, S_2, \dots S_6$ using 10,000 monte-carlo simulations under H_0 for nominal $\alpha = 0.01, 0.05$ and 0.10.

From table 4 we see that for the statistics S_1 through S_4, which used either the exact distribution or a large number (100,000) of monte-carlo simulations under H_0 to derive critical boundaries, the observed rejection rates correspond to the nominal α-levels as expected. The Wilcoxon rank-sum test is slightly conservative for small sample sizes ($m = n = 10$) due to the discreteness of the exact distribution, where the test isn't able to completely exhaust the nominal α. Interesting is that for S_5, the adaptive rank statistic using a flat kernel with asymptotic critical boundaries, the critical boundaries derived from the asymptotic distribution lead to a test that is too conservative for $\alpha = 0.05$ and 0.10. This will also be noticeable later in the simulations under the alternatives defined above, as S_5 will lag behind in power in many situations.

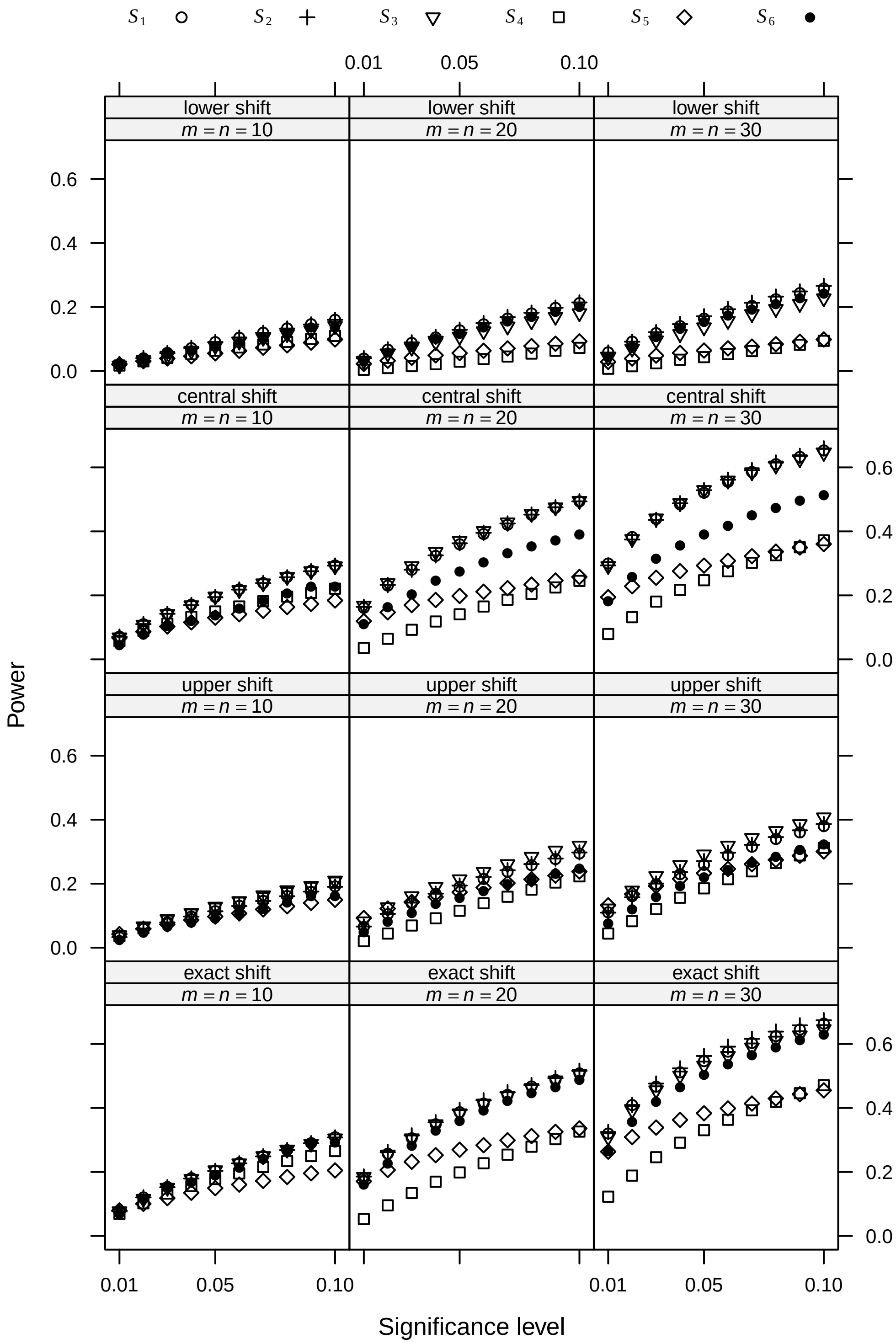

FIGURE 2. Empirical power of the test statistics $S_1, S_2, \ldots, S_6$ using 10,000 monte-carlo simulations under exact, lower, central and upper shift alternatives as defined in table 1 for nominal $\alpha = 0.01, 0.02, \ldots 0.10$ with underlying Cauchy(0, 1) F.

Shift	m, n	α	Test S_1	S_2	S_3	S_4	S_5	S_6
lower	10	0.01	0.02	0.02	0.02	0.02	0.02	0.02
		0.05	0.09	0.09	0.08	0.06	0.06	0.08
		0.10	0.16	0.16	0.14	0.11	0.10	0.14
	20	0.01	0.04	0.04	0.03	0.00	0.02	0.03
		0.05	0.13	0.13	0.11	0.03	0.06	0.12
		0.10	0.21	0.21	0.18	0.07	0.09	0.20
	30	0.01	0.06	0.06	0.04	0.01	0.03	0.04
		0.05	0.16	0.17	0.14	0.04	0.06	0.15
		0.10	0.26	0.27	0.23	0.09	0.10	0.24
central	10	0.01	0.07	0.07	0.07	0.06	0.07	0.04
		0.05	0.20	0.19	0.20	0.15	0.13	0.14
		0.10	0.29	0.29	0.29	0.22	0.18	0.23
	20	0.01	0.16	0.16	0.17	0.04	0.12	0.11
		0.05	0.36	0.36	0.37	0.14	0.20	0.27
		0.10	0.49	0.49	0.50	0.25	0.26	0.39
	30	0.01	0.30	0.29	0.29	0.08	0.19	0.18
		0.05	0.52	0.53	0.53	0.25	0.29	0.39
		0.10	0.65	0.66	0.65	0.37	0.36	0.51
upper	10	0.01	0.03	0.03	0.04	0.04	0.04	0.02
		0.05	0.11	0.11	0.13	0.12	0.10	0.09
		0.10	0.19	0.19	0.21	0.20	0.15	0.16
	20	0.01	0.06	0.07	0.08	0.02	0.09	0.05
		0.05	0.19	0.19	0.21	0.12	0.17	0.16
		0.10	0.29	0.30	0.32	0.22	0.24	0.25
	30	0.01	0.11	0.11	0.12	0.04	0.13	0.08
		0.05	0.26	0.27	0.29	0.19	0.23	0.22
		0.10	0.38	0.39	0.41	0.31	0.30	0.32
exact	10	0.01	0.08	0.08	0.08	0.07	0.08	0.07
		0.05	0.20	0.20	0.21	0.18	0.15	0.19
		0.10	0.31	0.31	0.30	0.27	0.20	0.29
	20	0.01	0.18	0.19	0.18	0.05	0.17	0.16
		0.05	0.39	0.39	0.38	0.20	0.27	0.36
		0.10	0.51	0.52	0.51	0.33	0.34	0.49
	30	0.01	0.32	0.32	0.31	0.12	0.26	0.26
		0.05	0.55	0.56	0.53	0.33	0.38	0.50
		0.10	0.66	0.67	0.65	0.47	0.46	0.63

TABLE 5. Empirical power of the test statistics $S_1, S_2, \ldots, S_6$ using 10,000 monte-carlo simulations under exact, lower, central and upper shift alternatives as defined in table 1 for nominal $\alpha = 0.01, 0.02, \ldots 0.10$ with underlying Cauchy(0, 1) F.

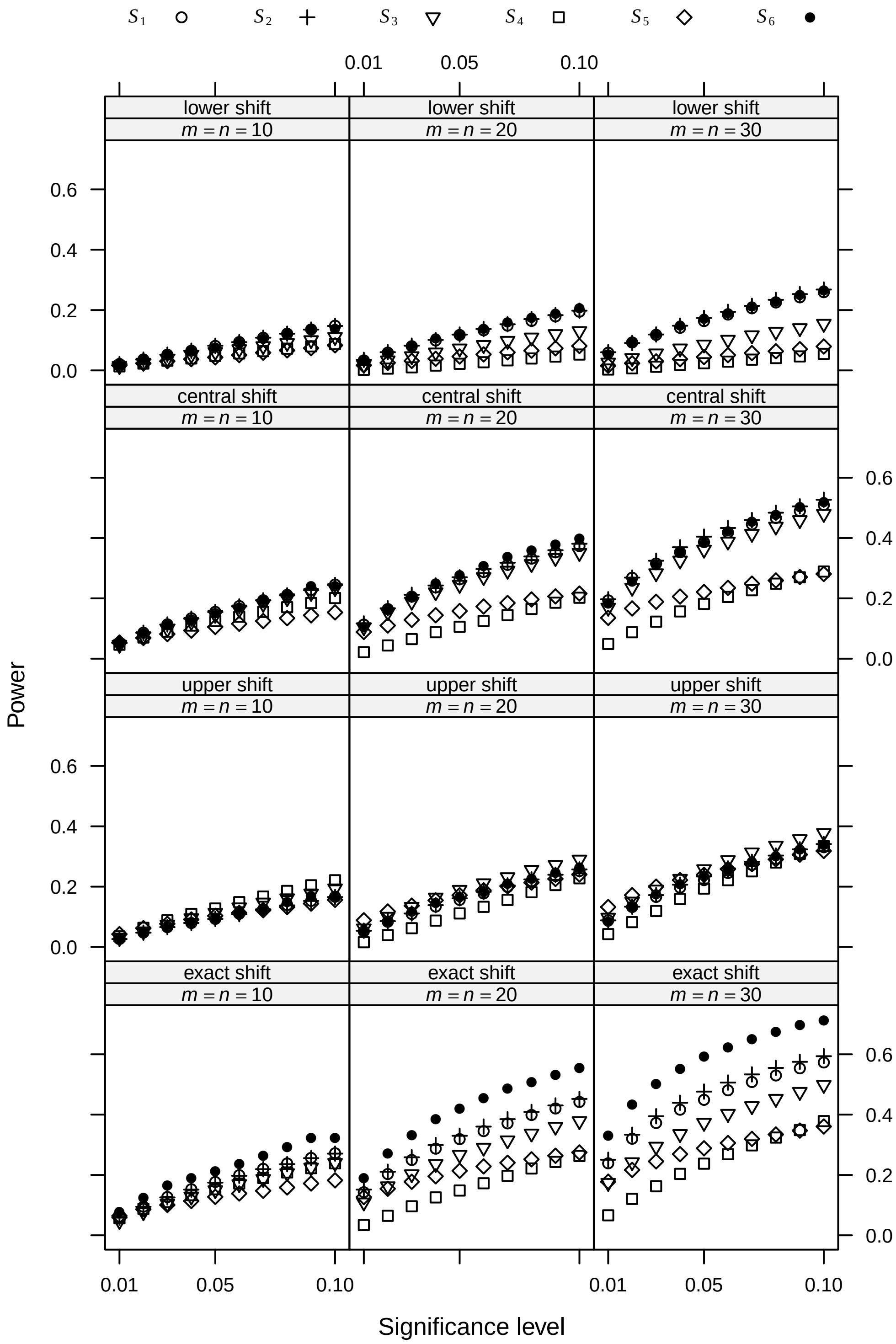

FIGURE 3. Empirical power of the test statistics $S_1, S_2, \ldots, S_6$ using 10,000 monte-carlo simulations under exact, lower, central and upper shift alternatives as defined in table 1 for nominal $\alpha = 0.01, 0.02, \ldots 0.10$ with underlying Logistic(0, 1) F.

Shift	m, n	α	Test S_1	S_2	S_3	S_4	S_5	S_6
lower	10	0.01	0.02	0.02	0.01	0.01	0.02	0.02
		0.05	0.08	0.08	0.06	0.05	0.04	0.08
		0.10	0.15	0.15	0.11	0.09	0.08	0.14
	20	0.01	0.03	0.03	0.02	0.00	0.02	0.03
		0.05	0.12	0.12	0.07	0.02	0.05	0.12
		0.10	0.19	0.20	0.13	0.05	0.08	0.21
	30	0.01	0.06	0.06	0.03	0.00	0.02	0.05
		0.05	0.16	0.17	0.09	0.02	0.04	0.17
		0.10	0.26	0.27	0.15	0.05	0.08	0.27
central	10	0.01	0.05	0.05	0.05	0.05	0.05	0.05
		0.05	0.16	0.15	0.15	0.12	0.11	0.15
		0.10	0.25	0.24	0.23	0.20	0.15	0.24
	20	0.01	0.11	0.12	0.10	0.02	0.09	0.10
		0.05	0.26	0.27	0.24	0.11	0.16	0.28
		0.10	0.37	0.38	0.35	0.20	0.22	0.40
	30	0.01	0.19	0.20	0.17	0.05	0.14	0.18
		0.05	0.39	0.40	0.36	0.18	0.22	0.39
		0.10	0.51	0.53	0.48	0.29	0.28	0.52
upper	10	0.01	0.03	0.03	0.03	0.04	0.04	0.03
		0.05	0.09	0.09	0.11	0.13	0.10	0.10
		0.10	0.17	0.17	0.19	0.22	0.16	0.17
	20	0.01	0.05	0.05	0.06	0.02	0.09	0.05
		0.05	0.16	0.16	0.19	0.11	0.17	0.16
		0.10	0.25	0.26	0.29	0.23	0.24	0.26
	30	0.01	0.09	0.09	0.10	0.04	0.13	0.08
		0.05	0.22	0.23	0.26	0.19	0.24	0.24
		0.10	0.33	0.34	0.38	0.33	0.32	0.34
exact	10	0.01	0.06	0.06	0.05	0.06	0.06	0.08
		0.05	0.18	0.17	0.15	0.15	0.13	0.21
		0.10	0.27	0.27	0.24	0.24	0.18	0.32
	20	0.01	0.14	0.15	0.11	0.03	0.13	0.19
		0.05	0.32	0.33	0.27	0.15	0.21	0.42
		0.10	0.44	0.45	0.38	0.26	0.27	0.55
	30	0.01	0.24	0.25	0.17	0.07	0.18	0.33
		0.05	0.45	0.48	0.37	0.24	0.29	0.59
		0.10	0.57	0.59	0.50	0.38	0.36	0.71

TABLE 6. Empirical power of the test statistics $S_1, S_2, \ldots, S_6$ using 10,000 monte-carlo simulations under exact, lower, central and upper shift alternatives as defined in table 1 for nominal $\alpha = 0.01, 0.02, \ldots 0.10$ with underlying Logistic(0, 1) F.

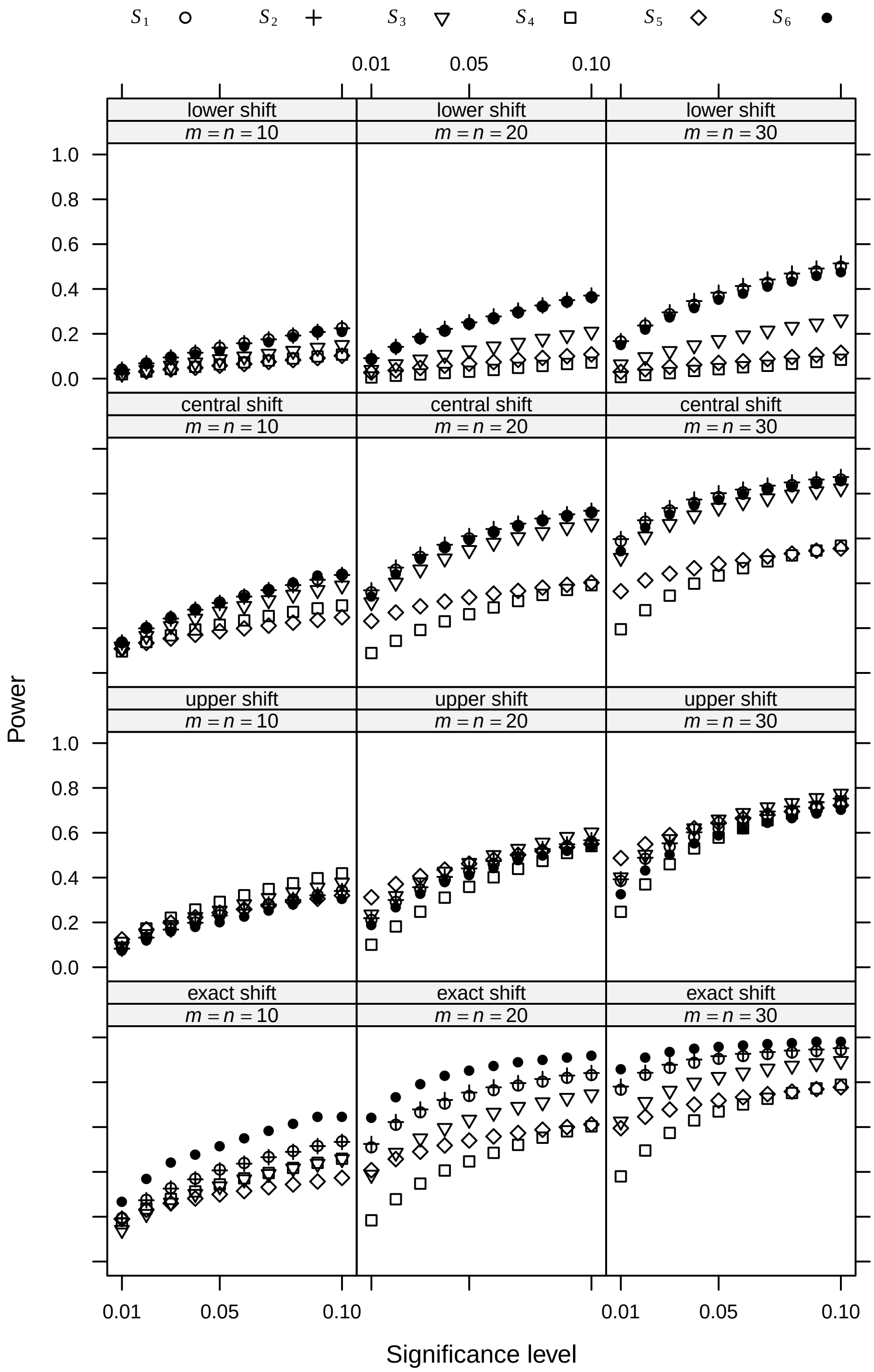

FIGURE 4. Empirical power of the test statistics $S_1, S_2, \ldots, S_6$ using 10,000 monte-carlo simulations under exact, lower, central and upper shift alternatives as defined in table 1 for nominal $\alpha = 0.01, 0.02, \ldots 0.10$ with underlying standard normal $N(0,1)$ F.

Shift	m, n	α	Test S_1	S_2	S_3	S_4	S_5	S_6
lower	10	0.01	0.04	0.04	0.02	0.02	0.02	0.04
		0.05	0.14	0.14	0.09	0.06	0.06	0.12
		0.10	0.23	0.22	0.15	0.11	0.10	0.21
	20	0.01	0.09	0.09	0.04	0.01	0.03	0.09
		0.05	0.24	0.25	0.13	0.03	0.07	0.25
		0.10	0.36	0.37	0.21	0.07	0.11	0.36
	30	0.01	0.17	0.17	0.06	0.01	0.03	0.15
		0.05	0.37	0.38	0.17	0.04	0.07	0.35
		0.10	0.50	0.51	0.26	0.08	0.12	0.48
central	10	0.01	0.14	0.14	0.11	0.10	0.11	0.13
		0.05	0.31	0.31	0.27	0.21	0.19	0.31
		0.10	0.44	0.44	0.39	0.30	0.25	0.44
	20	0.01	0.36	0.37	0.32	0.09	0.23	0.34
		0.05	0.60	0.61	0.55	0.26	0.34	0.59
		0.10	0.72	0.72	0.67	0.39	0.40	0.72
	30	0.01	0.59	0.60	0.51	0.19	0.36	0.54
		0.05	0.78	0.80	0.74	0.43	0.49	0.77
		0.10	0.86	0.87	0.82	0.57	0.55	0.85
upper	10	0.01	0.08	0.08	0.09	0.11	0.12	0.07
		0.05	0.23	0.23	0.25	0.29	0.24	0.20
		0.10	0.34	0.34	0.38	0.42	0.32	0.31
	20	0.01	0.21	0.22	0.24	0.10	0.31	0.19
		0.05	0.43	0.44	0.47	0.36	0.46	0.41
		0.10	0.56	0.57	0.60	0.54	0.55	0.54
	30	0.01	0.39	0.39	0.40	0.25	0.49	0.32
		0.05	0.62	0.64	0.66	0.58	0.64	0.59
		0.10	0.73	0.75	0.77	0.74	0.72	0.70
exact	10	0.01	0.20	0.19	0.14	0.17	0.19	0.27
		0.05	0.41	0.41	0.34	0.34	0.30	0.51
		0.10	0.54	0.53	0.46	0.46	0.37	0.65
	20	0.01	0.51	0.52	0.39	0.18	0.41	0.64
		0.05	0.74	0.75	0.63	0.45	0.54	0.85
		0.10	0.83	0.84	0.75	0.60	0.61	0.92
	30	0.01	0.77	0.78	0.62	0.38	0.59	0.86
		0.05	0.90	0.92	0.82	0.67	0.72	0.96
		0.10	0.94	0.95	0.89	0.79	0.78	0.98

TABLE 7. Empirical power of the test statistics $S_1, S_2, \ldots, S_6$ using 10,000 monte-carlo simulations under exact, lower, central and upper shift alternatives as defined in table 1 for nominal $\alpha = 0.01, 0.02, \ldots 0.10$ with underlying standard normal $N(0,1)$ F.

Of first interest in the simulation results is a comparison between the performance of the adaptive rank statistics and the popular non-adaptive Wilcoxon rank-sum test. The Wilcoxon rank-sum test performs essentially as well or better than all of the adaptive tests across all kinds of shifts with an underlying logistic distribution, which is not surprising, since the test can be derived as the optimal linear rank test for alternatives involving exact location shifts of logistic distributions, and the power differences are most pronounced in exactly this case (see figure 3 exact shift). Of note is also that the adaptive statistics S_3 and S_4 seem to have a very slight power advantage for larger α in the case of a shift in the upper range of the distribution.

The case for the underlying normal distribution is essentially the same, most likely due to this distribution's similarity to the logistic distribution. For alternatives using an underlying Cauchy distribution where less mass is concentrated in the tails of the distribution the situation is reversed, however, with many of the adaptive statistics consistently outperforming the rank-sum test, especially in the case of the central shift.

Secondly, we would like to look at the differences between the various adaptive rank statistics S_1 through S_5 proposed here. In general, the statistics S_4 and S_5 using the flattened kernels $K_{\gamma,\beta}$ pay a heavy price for the improved asymptotic behavior under H_0 and suffer a significant loss of power compared to the ranks-sum statistic and adaptive statistics using the bell-shaped Parzen-2 kernel. In most scenarios, the differences become more severe as the nominal significance level α increases.

It is also interesting to note that in many cases it does not seem to matter much whether the restricted kernel estimators $\hat{\hat{f}}_N$ and $\hat{\hat{g}}_N$ as proposed by Behnen et al. (1983) or the non-restricted kernel estimators $\hat{f}_N$ and $\hat{g}_N$ proposed here are used when forming the test statistic S_N, as long as we are using exact or monte-carlo critical boundaries and are not concerned with asymptotics.

For an underlying Cauchy distribution, there were no real differences in the performance of S_1 and S_2 based on a score function $\hat{\hat{b}}_N$ and S_3 based on $\hat{b}_N$, and S_3 even seemed to have a slight advantage over the restricted estimators in the case of an upper shift. For underlying normal and logistic F the results were much the same except for a distinct loss in power in S_3 relative to S_1 and S_2 on the order of around 0.10 across all significance levels examined in the case of a lower or exact shift.

CHAPTER 5

Proofs

5.1. Leading terms of S_N

In our proof of the representation of $S_N(\hat{b}_N)$ shown in Theorem 2.1 we showed in a first step that $S_N(\hat{b}_N)$ can be separated into a combination of leading terms that play a role in power and in the asymptotic distribution of the test statistic and a collection of asymptotically negigible rest terms:

$$S_N = \int \left[\bar{f}_N - \bar{g}_N\right] \circ H_N(x) \left[\hat{F}_m(dx) - F(dx)\right] \tag{2.38}$$

$$+ \int \left[\bar{f}_N - \bar{g}_N\right] \circ H_N(x)\, F(dx) \tag{2.41}$$

$$+ \int \left[\bar{f}_N - \bar{g}_N\right]' \circ H_N(x) \cdot \left[\hat{H}_N(x) - H_N(x)\right] F(dx) \tag{2.42}$$

$$+ \int \left[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\right] \circ H_N(x)\, F(dx) \tag{2.44}$$

$$+ O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1}).$$

Of these leading terms (2.38) and (2.42) comprise i.i.d. sums, while (2.41) is a deterministic component responsible for power under H_1. Thus, it remains to work further with the remaining leading term (2.44) to complete our linearization of S_N.

In the following, we will show that (2.44) can also be written as a sum of i.i.d. variables plus negligible rest. We can also note from the representation above that (2.38), (2.41) and (2.42) all vanish under H_0, so that (2.44) alone determines the asymptotic distribution of $S_N(\hat{b}_N)$ under the null hypothesis.

Now, we can separate (2.44) into two simpler terms

$$\int \left[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\right] \circ H_N(x)\, F(dx)$$

$$= \int \left[\hat{f}_N - \bar{f}_N\right] \circ H_N(x)\, F(dx) \tag{5.1}$$

$$- \int \left[\hat{g}_N - \bar{g}_N\right] \circ H_N(x)\, F(dx). \tag{5.2}$$

In the following we will derive i.i.d sums from (5.1) and (5.2) and combine these to get an i.i.d. sum for (2.44) plus negligible rest. All results in this section are proven using the same assumptions on K and a_N and definitions as in Theorem 2.1. We begin with (5.1). The work with (5.2) will be completely analogous.

First, recall the definitions of the kernel estimators $\hat{f}_N$ and $\hat{g}_N$ of the densities f_N and g_N and the functions $\bar{f}_N$ and $\bar{g}_N$:

$$\hat{f}_N(t) = (a_N \cdot m)^{-1} \cdot \sum_{i=1}^{m} K\left(\frac{t - \hat{H}_N(X_i)}{a_N}\right),$$

$$\hat{g}_N(t) = (a_N \cdot n)^{-1} \cdot \sum_{k=1}^{n} K\left(\frac{t - \hat{H}_N(Y_k)}{a_N}\right),$$

$$\bar{f}_N(t) = a_N^{-1} \int K\left(\frac{t - H_N(y)}{a_N}\right) F(dy), \quad 0 \le t \le 1,$$

$$\bar{g}_N(t) = a_N^{-1} \int K\left(\frac{t - H_N(y)}{a_N}\right) G(dy), \quad 0 \le t \le 1.$$

Then

$$\begin{aligned}
&\int \left[\hat{f}_N - \bar{f}_N\right] \circ H_N(x)\, F(dx) \\
&= \int \Bigg[(m \cdot a_N)^{-1} \cdot \sum_{i=1}^{m} K\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_i))\big) \\
&\qquad - a_N^{-1} \int K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\Bigg] F(dx) \\
&= \sum_{i=1}^{m} (m \cdot a_N)^{-1} \Bigg[\int K\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_i))\big)\, F(dx) \\
&\qquad - \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\, F(dx)\Bigg].
\end{aligned}$$

Using the Taylor expansion of the kernel function K about each of the $a_N^{-1}(H_N(x) - H_N(X_i))$ then yields

$$\begin{aligned}
&\sum_{i=1}^{m} (m \cdot a_N)^{-1} \Bigg[\int K\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \\
&\qquad - \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\, F(dx)\Bigg] && (5.3) \\
&\quad + a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \cdot \big(H_N(X_i) - \hat{H}_N(X_i)\big) && (5.4) \\
&\quad + a_N^{-1} \cdot m^{-1} \cdot \sum_{i=1}^{m} \iint_{a_N^{-1}(H_N(x)-H_N(X_i))}^{a_N^{-1}(H_N(x)-\hat{H}_N(X_i))} \big(a_N^{-1}(H_N(x) - \hat{H}_N(X_i)) - t\big) \cdot K''(t)\, dt\, F(dx). && (5.5)
\end{aligned}$$

It is immediately apparent that (5.3) is already a sum of centered i.i.d. variables. Thus, it remains to work on deriving an i.i.d. sum from (5.4). Recalling the definition of the pooled empirical d.f. $\hat{H}_N$ we see that (5.4) is equal to

$$\begin{aligned}
&a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \cdot \left[H_N(X_i) - N^{-1}\left[\sum_{j=1}^{m} 1_{\{X_j \le X_i\}} + \sum_{k=1}^{n} 1_{\{Y_k \le X_i\}}\right]\right] \\
&= a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \cdot H_N(X_i)
\end{aligned}$$

$$
\begin{aligned}
&- a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \cdot N^{-1} \cdot \sum_{j=1}^{m} 1_{\{X_j \le X_i\}} \\
&- a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \cdot N^{-1} \cdot \sum_{k=1}^{n} 1_{\{Y_k \le X_i\}} \\
= {}& a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \cdot H_N(X_i) \qquad (5.6) \\
&- \lambda_N \cdot a_N^{-2} \cdot m^{-2} \cdot \sum_{i=1}^{m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \qquad (5.7) \\
&- \lambda_N \cdot a_N^{-2} \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \cdot 1_{\{X_j \le X_i\}} \qquad (5.8) \\
&- (1 - \lambda_N) \cdot a_N^{-2} \cdot m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \cdot 1_{\{Y_k \le X_i\}}. \qquad (5.9)
\end{aligned}
$$

We see that (5.6) is already an i.i.d. sum and (5.7) is negligible, since

$$
\begin{aligned}
&\left| \lambda_N \cdot a_N^{-2} \cdot m^{-2} \cdot \sum_{i=1}^{m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \right| \\
&\le \lambda_N \cdot a_N^{-2} \cdot m^{-2} \cdot \sum_{i=1}^{m} \int \big| K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big) \big|\, F(dx) \\
&\le \lambda_N \cdot a_N^{-2} \cdot m^{-2} \cdot \sum_{i=1}^{m} \int \big\| K' \big\| \\
&= \lambda_N \big\| K' \big\| \cdot a_N^{-2} \cdot m^{-1} \\
&= O(a_N^{-2} \cdot N^{-1}). \qquad (5.10)
\end{aligned}
$$

Further, (5.8) and (5.9) are a U-statistic and generalized U-statistic scaled by $\lambda_N(m-1)m^{-1}$ and $(1-\lambda_N)$ respectively. We will proceed by finding projections of (5.8) and (5.9) onto the space of i.i.d. sums which we can continue to work with.

Lemma 5.1.

$$
\begin{aligned}
& a_N^{-2} \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \cdot 1_{\{X_j \le X_i\}} \qquad (5.11) \\
&= a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \Big[\iint K'\left(a_N^{-1}(H_N(x) - H_N(X_i))\right)\, F(dx) \cdot 1_{\{y \le X_i\}}\, F(dy) \\
&\quad + \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx) \cdot 1_{\{X_i \le y\}}\, F(dy) \\
&\quad - \iiint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx) \cdot 1_{\{z \le y\}}\, F(dy)\, F(dz) \Big] + O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1}).
\end{aligned}
$$

Proof. Define

$$
u_N(r,s) = a_N^{-2} \int K'\left(a_N^{-1}(H_N(x) - H_N(r))\right) F(dx) \cdot 1_{\{s \le r\}},
$$

and define the U-statistic U_m as

$$U_m = m^{-1}(m-1)^{-1} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j).$$

Then

$$a_N^{-2} \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) \cdot 1_{\{X_j \le X_i\}} = (m-1)m^{-1} \cdot U_m.$$

Further, let $\hat{U}_m$ be the Hájek projection of U_m as defined in Lemma A.2:

$$\begin{aligned}
\hat{U}_m &= m^{-1} \cdot \sum_{i=1}^{m} \Big[\int u_N(X_i, y)\, F(dy) + \int u_N(z, X_i)\, F(dz) - \iint u_N(y,z)\, F(dy)\, F(dz) \Big] \\
&= a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \Big[\iint K'\left(a_N^{-1}(H_N(x) - H_N(X_i))\right) F(dx) \cdot 1_{\{y \le X_i\}}\, F(dy) \\
&\qquad + \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx) \cdot 1_{\{X_i \le y\}}\, F(dy) \\
&\qquad - \iiint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx) \cdot 1_{\{z \le y\}}\, F(dy)\, F(dz) \Big].
\end{aligned}$$

In order to complete the proof it only remains to show that

$$\mathbb{E}\Big[(m-1)m^{-1} \cdot U_m - \hat{U}_m\Big]^2 = O(a_N^{-4} \cdot N^{-2}).$$

Applying the inequality from Lemma A.2 we have

$$\begin{aligned}
\mathbb{E}\Big[(m-1)m^{-1} \cdot U_m - \hat{U}_m\Big]^2 &= \mathbb{E}\Big[(m-1)m^{-1} \cdot U_m - U_m + U_m - \hat{U}_m\Big]^2 \\
&\le 2 \cdot \mathbb{E}\Big[(m-1)m^{-1} \cdot U_m - U_m\Big]^2 + 2 \cdot \mathbb{E}\Big[U_m - \hat{U}_m\Big]^2 \\
&\le 2m^{-2} \cdot \mathbb{E}\big[U_m\big]^2 + 4(m-1)m^{-3} \cdot \mathbb{E}\big[u_N^*(X_1, X_2)\big]^2
\end{aligned}$$

for u_N^* defined as

$$u_N^*(r,s) = u_N(r,s) - \int u_N(r,y)\, F(dy) - \int u_N(z,s)\, F(dz) + \iint u_N(z,y)\, F(dy)\, F(dz).$$

Now, the kernel function u_N is uniformly bounded:

$$\|u_N\| \le \|K'\| a_N^{-2}$$

which means for the first expectation we can write

$$\mathbb{E}\big[U_m\big]^2 = \mathbb{E}\Big[m^{-1}(m-1)^{-1} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j)\Big]^2 \le \|u_N\|^2 \le \|K'\|^2 a_N^{-4}.$$

Thus, it remains only to bound the second expectation $\mathbb{E}\big[u_N^*(X_1, X_2)\big]^2$:

$$\begin{aligned}
\mathbb{E}\big[u_N^*(X_1, X_2)\big]^2 \le 4 \cdot \mathbb{E}\Big[\big[u_N(X_1, X_2)\big]^2 &+ \Big[\int u_N(X_1, y)\, F(dy)\Big]^2 \\
&+ \Big[\int u_N(z, X_2)\, F(dz)\Big]^2 + \Big[\iint u_N(z,y)\, F(dy)\, F(dz)\Big]^2\Big]
\end{aligned}$$

$$
\begin{aligned}
&\le 4 \cdot \mathbb{E}\Big[\big[u_N(X_1,X_2)\big]^2 + \int \big[u_N(X_1,y)\big]^2\,F(dy) \\
&\qquad + \int \big[u_N(z,X_2)\big]^2\,F(dz) + \iint \big[u_N(z,y)\big]^2\,F(dy)\,F(dz)\Big] \\
&\le 4 \cdot \mathbb{E}\Big[4 \cdot \Big[\|K'\| a_N^{-2}\Big]^2\Big] \\
&= 4^2 \big\|K'\big\|^2 a_N^{-4}.
\end{aligned}
$$

Altogether this gives us

$$
\begin{aligned}
\mathbb{E}\Big[(m-1)m^{-1} \cdot U_m - \hat{U}_m\Big]^2 &\le 2m^{-2} \cdot \mathbb{E}\big[U_m\big]^2 + 4(m-1)m^{-3} \cdot \mathbb{E}\big[u_N^*(X_1,X_2)\big]^2 \\
&\le 2m^{-2} \cdot \big\|K'\big\|^2 a_N^{-4} + 4(m-1)m^{-3} \cdot 4^2 \big\|K'\big\|^2 a_N^{-4} \\
&= 2\big\|K'\big\|^2 a_N^{-4} \cdot m^{-2} + 4^3 \big\|K'\big\|^2 a_N^{-4} \cdot (m-1)m^{-3} \\
&= O(a_N^{-4} \cdot N^{-2})
\end{aligned}
$$

which completes the proof. □

Using the following lemma, we can replace (5.9) by a projection as well.

Lemma 5.2.

$$
\begin{aligned}
&a_N^{-2} \cdot m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\,F(dx) \cdot 1_{\{Y_k \le X_i\}} \qquad (5.12) \\
&= a_N^{-2} \cdot \Big[m^{-1} \cdot \sum_{i=1}^{m} \iint K'\left(a_N^{-1}(H_N(x) - H_N(X_i))\right)\,F(dx) \cdot 1_{\{y \le X_i\}}\,G(dy) \\
&\quad + n^{-1} \cdot \sum_{k=1}^{n} \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\,F(dx) \cdot 1_{\{Y_k \le y\}}\,F(dy) \\
&\quad - \iiint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\,F(dx) \cdot 1_{\{z \le y\}}\,F(dy)\,G(dz)\Big] + O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1}).
\end{aligned}
$$

Proof. Define

$$
u_N(r,s) = a_N^{-2} \int K'\left(a_N^{-1}(H_N(x) - H_N(r))\right) F(dx) \cdot 1_{\{s \le r\}},
$$

and define the generalized U-statistic $U_{m,n}$ as

$$
U_{m,n} = m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} u_N(X_i, Y_k).
$$

Then

$$
a_N^{-2} \cdot m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\,F(dx) \cdot 1_{\{Y_k \le X_i\}} = U_{m,n}.
$$

Further, let $\hat{U}_{m,n}$ be the Hájek projection of $U_{m,n}$ as defined in Lemma A.3:

$$\begin{aligned}\hat{U}_{m,n} &= m^{-1} \cdot \sum_{i=1}^{m} \int u_N(X_i, y)\, G(dy) + n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x, Y_k)\, F(dx) - \iint u_N(x,y)\, F(dx)G(dy) \\ &= a_N^{-2} \cdot \Bigg[m^{-1} \cdot \sum_{i=1}^{m} \iint K'\left(a_N^{-1}(H_N(x) - H_N(X_i))\right)\, F(dx) \cdot 1_{\{y \le X_i\}}\, G(dy) \\ &\quad + n^{-1} \cdot \sum_{k=1}^{n} \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx) \cdot 1_{\{Y_k \le y\}}\, F(dy) \\ &\quad - \iiint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx) \cdot 1_{\{z \le y\}}\, F(dy)\, G(dz) \Bigg].\end{aligned}$$

In order to complete the proof it only remains to show that

$$\mathbb{E}\Big[U_{m,n} - \hat{U}_{m,n}\Big]^2 = O(a_N^{-4} \cdot N^{-2}).$$

Applying the equality from Lemma A.3 we have

$$\mathbb{E}\Big[U_{m,n} - \hat{U}_{m,n}\Big]^2 = m^{-1} n^{-1} \cdot \mathbb{E}\big[u_N^*(X_1, Y_1)\big]^2$$

for u_N^* defined as

$$u_N^*(r,s) = u(r,s) - \int u(r,y)\, G(dy) - \int u(x,s)\, F(dx) + \iint u(x,y)\, F(dx)\, G(dy).$$

Since the kernel function u_N is uniformly bounded:

$$\|u_N\| \le \|K'\| a_N^{-2}$$

we can write

$$\begin{aligned}\mathbb{E}\big[u_N^*(X_1,Y_1)\big]^2 &\le 4 \cdot \mathbb{E}\Bigg[\big[u_N(X_1,Y_1)\big]^2 + \Big[\int u_N(X_1, y)\, G(dy)\Big]^2 \\ &\quad + \Big[\int u_N(z, Y_1)\, F(dz)\Big]^2 + \Big[\iint u_N(z,y)\, G(dy)\, F(dz)\Big]^2 \Bigg] \\ &\le 4 \cdot \mathbb{E}\Bigg[\big[u_N(X_1,Y_1)\big]^2 + \int \big[u_N(X_1,y)\big]^2\, G(dy) \\ &\quad + \int \big[u_N(z,Y_1)\big]^2\, F(dz) + \iint \big[u_N(z,y)\big]^2\, G(dy)\, F(dz) \Bigg] \\ &\le 4 \cdot \mathbb{E}\Big[4 \cdot \Big[\|K'\| a_N^{-2}\Big]^2\Big] \\ &= 4^2 \|K'\|^2 a_N^{-4}.\end{aligned}$$

Altogether this gives us

$$\begin{aligned}\mathbb{E}\Big[U_{m,n} - \hat{U}_{m,n}\Big]^2 &= m^{-1} n^{-1} \cdot \mathbb{E}\big[u_N^*(X_1,Y_1)\big]^2 \\ &\le m^{-1} n^{-1} \cdot 4^2 \|K'\|^2 a_N^{-4} \\ &= 4^2 \|K'\|^2 a_N^{-4} \cdot m^{-1} n^{-1} \\ &= O(a_N^{-4} \cdot N^{-2})\end{aligned}$$

which completes the proof. □

LEMMA 5.3.

$$a_N^{-1} \cdot m^{-1} \cdot \sum_{i=1}^{m} \iint_{a_N^{-1}(H_N(x)-H_N(X_i))}^{a_N^{-1}(H_N(x)-\hat{H}_N(X_i))} \left(a_N^{-1}(H_N(x) - \hat{H}_N(X_i)) - t\right) \cdot K''(t)\, dt\, F(dx) = O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1}). \tag{5.13}$$

PROOF.

$$\begin{aligned}
&\left| a_N^{-1} \cdot m^{-1} \cdot \sum_{i=1}^{m} \iint_{a_N^{-1}(H_N(x)-H_N(X_i))}^{a_N^{-1}(H_N(x)-\hat{H}_N(X_i))} \left(a_N^{-1}(H_N(x) - \hat{H}_N(X_i)) - t\right) \cdot K''(t)\, dt\, F(dx) \right| \\
&= \left| a_N^{-1} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int_0^1 \int_{a_N^{-1}(v-H_N(X_i))}^{a_N^{-1}(v-\hat{H}_N(X_i))} \left(a_N^{-1}(v - \hat{H}_N(X_i)) - t\right) \cdot K''(t)\, dt\, f_N(v)\, dv \right| \\
&= a_N^{-1} \cdot m^{-1} \cdot \Bigg| \sum_{\substack{1 \le i \le m \\ H_N(X_i) > \hat{H}_N(X_i)}} \int_0^1 \int_{-1}^1 1_{\{a_N^{-1}(v-H_N(X_i)) \,<\, t \,<\, a_N^{-1}(v-\hat{H}_N(X_i))\}} \\
&\qquad \times \left(a_N^{-1}(v - \hat{H}_N(X_i)) - t\right) \cdot K''(t)\, dt\, f_N(v)\, dv \\
&\quad + \sum_{\substack{1 \le i \le m \\ H_N(X_i) \le \hat{H}_N(X_i)}} \int_0^1 \int_{-1}^1 1_{\{a_N^{-1}(v-\hat{H}_N(X_i)) \,<\, t \,<\, a_N^{-1}(v-H_N(X_i))\}} \\
&\qquad \times \left(a_N^{-1}(v - \hat{H}_N(X_i)) - t\right) \cdot K''(t)\, dt\, f_N(v)\, dv \Bigg| \\
&= a_N^{-1} \cdot m^{-1} \cdot \Bigg| \sum_{\substack{1 \le i \le m \\ H_N(X_i) > \hat{H}_N(X_i)}} \int_0^1 \int_{-1}^1 1_{\{\hat{H}_N(X_i))+a_N \cdot t \,<\, v \,<\, H_N(X_i))+a_N \cdot t\}} \\
&\qquad \times \left(a_N^{-1}(v - \hat{H}_N(X_i)) - t\right) \cdot K''(t)\, dt\, f_N(v)\, dv \\
&\quad + \sum_{\substack{1 \le i \le m \\ H_N(X_i) \le \hat{H}_N(X_i)}} \int_0^1 \int_{-1}^1 1_{\{H_N(X_i))+a_N \cdot t \,<\, v \,<\, \hat{H}_N(X_i))+a_N \cdot t\}} \\
&\qquad \times \left(a_N^{-1}(v - \hat{H}_N(X_i)) - t\right) \cdot K''(t)\, dt\, f_N(v)\, dv \Bigg| \\
&\le a_N^{-1} \cdot m^{-1} \cdot \Bigg[\sum_{\substack{1 \le i \le m \\ H_N(X_i) > \hat{H}_N(X_i)}} \int_{-1}^1 \int_0^1 1_{\{\hat{H}_N(X_i))+a_N \cdot t \,<\, v \,<\, H_N(X_i))+a_N \cdot t\}} \\
&\qquad \times \left| a_N^{-1}(v - \hat{H}_N(X_i)) - t \right| \cdot f_N(v)\, dv\, \left| K''(t) \right| dt \\
&\quad + \sum_{\substack{1 \le i \le m \\ H_N(X_i) \le \hat{H}_N(X_i)}} \int_{-1}^1 \int_0^1 1_{\{H_N(X_i))+a_N \cdot t \,<\, v \,<\, \hat{H}_N(X_i))+a_N \cdot t\}} \\
&\qquad \times \left| a_N^{-1}(v - \hat{H}_N(X_i)) - t \right| \cdot f_N(v)\, dv\, \left| K''(t) \right| dt \Bigg]
\end{aligned}$$

$$
\begin{aligned}
&\leq a_N^{-1} \cdot m^{-1} \cdot \Bigg[\sum_{\substack{1\leq i\leq m \\ H_N(X_i)>\hat{H}_N(X_i)}} \int_{-1}^{1} \|f_N\| \int_0^1 1_{\{\hat{H}_N(X_i))+a_N\cdot t\,<\,v\,<\,H_N(X_i))+a_N\cdot t\}} \\
&\qquad \times \left|a_N^{-1}(v-\hat{H}_N(X_i))-t\right| dv \left|K''(t)\right| dt \\
&\qquad + \sum_{\substack{1\leq i\leq m \\ H_N(X_i)\leq\hat{H}_N(X_i)}} \int_{-1}^{1} \|f_N\| \int_0^1 1_{\{H_N(X_i))+a_N\cdot t\,<\,v\,<\,\hat{H}_N(X_i))+a_N\cdot t\}} \\
&\qquad \times \left|a_N^{-1}(v-\hat{H}_N(X_i))-t\right| dv \left|K''(t)\right| dt\Bigg] \\
&\leq a_N^{-1} \cdot m^{-1} \cdot \|f_N\| \cdot \Bigg[\sum_{\substack{1\leq i\leq m \\ H_N(X_i)>\hat{H}_N(X_i)}} \int_{-1}^{1} \left|a_N^{-1}(H_N(X_i)-\hat{H}_N(X_i))\right| \\
&\qquad \times \int_0^1 1_{\{\hat{H}_N(X_i))+a_N\cdot t\,<\,v\,<\,H_N(X_i))+a_N\cdot t\}}\, dv \left|K''(t)\right| dt \\
&\qquad + \sum_{\substack{1\leq i\leq m \\ H_N(X_i)\leq\hat{H}_N(X_i)}} \int_{-1}^{1} \left|a_N^{-1}(H_N(X_i)-\hat{H}_N(X_i))\right| \int_0^1 1_{\{H_N(X_i))+a_N\cdot t\,<\,v\,<\,\hat{H}_N(X_i))+a_N\cdot t\}}\, dv \left|K''(t)\right| dt\Bigg] \\
&\leq a_N^{-2} \cdot m^{-1} \cdot \|f_N\| \cdot \Bigg[\sum_{\substack{1\leq i\leq m \\ H_N(X_i)>\hat{H}_N(X_i)}} \left|H_N(X_i)-\hat{H}_N(X_i)\right| \\
&\qquad \times \int_{-1}^{1}\int_0^1 1_{\{\hat{H}_N(X_i))+a_N\cdot t\,<\,v\,<\,H_N(X_i))+a_N\cdot t\}}\, dv \left|K''(t)\right| dt \\
&\qquad + \sum_{\substack{1\leq i\leq m \\ H_N(X_i)\leq\hat{H}_N(X_i)}} \left|H_N(X_i)-\hat{H}_N(X_i)\right| \cdot \int_{-1}^{1}\int_0^1 1_{\{H_N(X_i))+a_N\cdot t\,<\,v\,<\,\hat{H}_N(X_i))+a_N\cdot t\}}\, dv \left|K''(t)\right| dt\Bigg] \\
&\leq a_N^{-2} \cdot m^{-1} \cdot \|f_N\| \cdot \sum_{i=1}^{m} \left|H_N(X_i)-\hat{H}_N(X_i)\right|^2 \cdot 2\|K''\| \\
&\leq 2\|K''\| a_N^{-2} \cdot \|f_N\| \cdot \|\hat{H}_N - H_N\|^2 \\
&= O_{\mathbb{P}}(a_N^{-2}\cdot N^{-1}),
\end{aligned}
$$

due to the D-K-W bound on $\|\hat{H}_N - H_N\|$ and the fact that $\|f_N\| = O(1)$ (see Lemma A.1). □

We can apply (5.3), (5.6) and (5.10) together with Lemmas 5.1, 5.2 and 5.3 to express (5.1) as an i.i.d sum plus negligible rest terms, which we record in the following lemma.

LEMMA 5.4.

$$
\begin{aligned}
&\int \left[\hat{f}_N - \bar{f}_N\right] \circ H_N(x)\, F(dx) = \\
&\quad = a_N^{-1} \cdot m^{-1} \cdot \sum_{i=1}^{m} \Bigg[\int K\big(a_N^{-1}(H_N(x)-H_N(X_i))\big)\, F(dx) \\
&\qquad - \iint K\big(a_N^{-1}(H_N(x)-H_N(y))\big)\, F(dy)\, F(dx)\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
&- \lambda_N \cdot a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \Big[\iint K' \left(a_N^{-1}(H_N(x) - H_N(y))\right) \, F(dx) \cdot 1_{\{X_i \le y\}} \, F(dy) \\
&\quad - \iint K' \left(a_N^{-1}(H_N(x) - H_N(y))\right) \, F(dx) \cdot F(y) \, F(dy) \Big] \\
&- (1-\lambda_N) \cdot a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \Big[\iint K' \left(a_N^{-1}(H_N(x) - H_N(y))\right) \, F(dx) \cdot 1_{\{Y_k \le y\}} \, F(dy) \\
&\quad - \iint K' \left(a_N^{-1}(H_N(x) - H_N(y))\right) \, F(dx) \cdot G(y) \, F(dy) \Big] \\
&+ O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1}).
\end{aligned}
$$

PROOF. Combine (5.3), (5.6) and (5.10) together with Lemmas 5.1 and 5.2 to get

$$
\begin{aligned}
&\int \left[\hat{f}_N - \bar{f}_N\right] \circ H_N(x) \, F(dx) \\
&= a_N^{-1} \cdot m^{-1} \cdot \sum_{i=1}^{m} \Big[\int K\big(a_N^{-1}(H_N(x) - H_N(X_i))\big) \, F(dx) \\
&\quad - \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big) \, F(dy) \, F(dx) \Big] \\
&+ a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big) \, F(dx) \cdot H_N(X_i) \\
&- \lambda_N \cdot a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \Big[\iint K' \left(a_N^{-1}(H_N(x) - H_N(X_i))\right) \, F(dx) \cdot 1_{\{y \le X_i\}} \, F(dy) \\
&\quad + \iint K' \left(a_N^{-1}(H_N(x) - H_N(y))\right) \, F(dx) \cdot 1_{\{X_i \le y\}} \, F(dy) \\
&\quad - \iiint K' \left(a_N^{-1}(H_N(x) - H_N(y))\right) \, F(dx) \cdot 1_{\{z \le y\}} \, F(dy) \, F(dz) \Big] \\
&- (1-\lambda_N) \cdot a_N^{-2} \cdot \Big[m^{-1} \cdot \sum_{i=1}^{m} \iint K' \left(a_N^{-1}(H_N(x) - H_N(X_i))\right) \, F(dx) \cdot 1_{\{y \le X_i\}} \, G(dy) \\
&\quad + n^{-1} \cdot \sum_{k=1}^{n} \iint K' \left(a_N^{-1}(H_N(x) - H_N(y))\right) \, F(dx) \cdot 1_{\{Y_k \le y\}} \, F(dy) \\
&\quad - \iiint K' \left(a_N^{-1}(H_N(x) - H_N(y))\right) \, F(dx) \cdot 1_{\{z \le y\}} \, F(dy) \, G(dz) \Big] \\
&+ O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1}) \\
&= a_N^{-1} \cdot m^{-1} \cdot \sum_{i=1}^{m} \Big[\int K\big(a_N^{-1}(H_N(x) - H_N(X_i))\big) \, F(dx) \\
&\quad - \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big) \, F(dy) \, F(dx) \Big] \\
&+ a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big) \, F(dx) \cdot H_N(X_i) \\
&- a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int K' \left(a_N^{-1}(H_N(x) - H_N(X_i))\right) \, F(dx) \cdot \lambda_N F(X_i)
\end{aligned}
$$

$$- \lambda_N \cdot a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx) \cdot 1_{\{X_i \le y\}} F(dy)$$

$$+ \lambda_N \cdot a_N^{-2} \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx) \cdot F(y) F(dy)$$

$$- a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int K'\left(a_N^{-1}(H_N(x) - H_N(X_i))\right) F(dx) \cdot (1-\lambda_N) G(X_i)$$

$$- (1-\lambda_N) \cdot a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx) \cdot 1_{\{Y_k \le y\}} F(dy)$$

$$+ (1-\lambda_N) \cdot a_N^{-2} \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx) \cdot G(y) F(dy)$$

$$+ O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1}).$$

Now, since

$$\begin{aligned}
& a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big) F(dx) \cdot H_N(X_i) \\
&\quad = a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int K'\left(a_N^{-1}(H_N(x) - H_N(X_i))\right) F(dx) \cdot \lambda_N F(X_i) \\
&\qquad + a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \int K'\left(a_N^{-1}(H_N(x) - H_N(X_i))\right) F(dx) \cdot (1-\lambda_N) G(X_i)
\end{aligned}$$

this simplifies to

$$\begin{aligned}
& \int \left[\hat{f}_N - \bar{f}_N\right] \circ H_N(x) F(dx) \\
&= a_N^{-1} \cdot m^{-1} \cdot \sum_{i=1}^{m} \Big[\int K\big(a_N^{-1}(H_N(x) - H_N(X_i))\big) F(dx) \\
&\quad - \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big) F(dy) F(dx) \Big] \\
&\quad - \lambda_N \cdot a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx) \cdot 1_{\{X_i \le y\}} F(dy) \\
&\quad + \lambda_N \cdot a_N^{-2} \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx) \cdot F(y) F(dy) \\
&\quad - (1-\lambda_N) \cdot a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx) \cdot 1_{\{Y_k \le y\}} F(dy) \\
&\quad + (1-\lambda_N) \cdot a_N^{-2} \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx) \cdot G(y) F(dy) \\
&\quad + O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1}) \\
&= a_N^{-1} \cdot m^{-1} \cdot \sum_{i=1}^{m} \Big[\int K\big(a_N^{-1}(H_N(x) - H_N(X_i))\big) F(dx) \\
&\quad - \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big) F(dy) F(dx) \Big]
\end{aligned}$$

$$
\begin{aligned}
&- \lambda_N \cdot a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \Big[\iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) \, F(dx) \cdot 1_{\{X_i \le y\}} \, F(dy) \\
&\quad - \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) \, F(dx) \cdot F(y) \, F(dy) \Big] \\
&- (1-\lambda_N) \cdot a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \Big[\iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) \, F(dx) \cdot 1_{\{Y_k \le y\}} \, F(dy) \\
&\quad - \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) \, F(dx) \cdot G(y) \, F(dy) \Big] \\
&+ O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1})
\end{aligned}
$$

which completes the proof. □

To derive an i.i.d sum from (5.2) we will use very similar arguments to those which we used to work with (5.1). We begin by deriving a sum representation of (5.2).

$$
\begin{aligned}
&\int \Big[\hat{g}_N - \bar{g}_N\Big] \circ H_N(x) \, F(dx) \\
&= \int \Big[(n \cdot a_N)^{-1} \cdot \sum_{k=1}^{n} K\big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))\big) \\
&\quad - a_N^{-1} \int K\big(a_N^{-1}(H_N(x) - H_N(y))\big) \, G(dy)\Big] \, F(dx) \\
&= \sum_{k=1}^{n} (n \cdot a_N)^{-1} \Big[\int K\big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))\big) \, F(dx) \\
&\quad - \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big) \, G(dy) \, F(dx) \Big].
\end{aligned}
$$

Using the Taylor expansion of the kernel function K about each of the $a_N^{-1}(H_N(x) - H_N(Y_k))$ then yields

$$
\begin{aligned}
&\sum_{k=1}^{n} (n \cdot a_N)^{-1} \Big[\int K\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big) \, F(dx) \\
&\quad - \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big) \, G(dy) \, F(dx) \Big] && (5.14) \\
&+ a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big) \, F(dx) \cdot \big(H_N(Y_k) - \hat{H}_N(Y_k)\big) && (5.15) \\
&+ a_N^{-1} \cdot n^{-1} \cdot \sum_{k=1}^{n} \iint_{a_N^{-1}(H_N(x) - H_N(Y_k))}^{a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))} \big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k)) - t\big) \cdot K''(t) \, dt \, F(dx). && (5.16)
\end{aligned}
$$

It is immediately apparent that (5.14) is already a sum of centered i.i.d. variables. Thus, it remains to work on deriving an i.i.d. sum from (5.15). Recalling the definition of the pooled empirical d.f. $\hat{H}_N$ we see that (5.15) is equal to

$$
\begin{aligned}
&a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big) \, F(dx) \cdot \Big[H_N(Y_k) - N^{-1}\Big[\sum_{j=1}^{m} 1_{\{X_j \le Y_k\}} + \sum_{l=1}^{n} 1_{\{Y_l \le Y_k\}}\Big]\Big] \\
&= a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big) \, F(dx) \cdot H_N(Y_k)
\end{aligned}
$$

$$- a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) \cdot N^{-1} \cdot \sum_{j=1}^{m} 1_{\{X_j \le Y_k\}}$$

$$- a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) \cdot N^{-1} \cdot \sum_{l=1}^{n} 1_{\{Y_l \le Y_k\}}$$

$$= a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) \cdot H_N(Y_k) \tag{5.17}$$

$$- \lambda_N \cdot a_N^{-2} \cdot m^{-1} n^{-1} \cdot \sum_{k=1}^{n} \sum_{j=1}^{m} \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) \cdot 1_{\{X_j \le Y_k\}} \tag{5.18}$$

$$- (1 - \lambda_N) \cdot a_N^{-2} \cdot n^{-2} \cdot \sum_{k=1}^{n} \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) \tag{5.19}$$

$$- (1 - \lambda_N) \cdot a_N^{-2} \cdot n^{-2} \cdot \sum_{1 \le k \ne l \le m} \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) \cdot 1_{\{Y_l \le Y_k\}}. \tag{5.20}$$

We see that (5.17) is already an i.i.d. sum and

$$(1 - \lambda_N) \cdot a_N^{-2} \cdot n^{-2} \cdot \sum_{k=1}^{n} \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) = O(a_N^{-2} \cdot N^{-1}) \tag{5.21}$$

by (5.10).

Further, (5.18) and (5.20) are a generalized U-statistic and U-statistic scaled by λ_N and $(1-\lambda_N)\cdot(n-1)n^{-1}$ respectively. We will proceed by finding projections of (5.18) and (5.20) onto the space of i.i.d. sums which we can continue to work with.

Lemma 5.5.

$$a_N^{-2} \cdot m^{-1} n^{-1} \cdot \sum_{k=1}^{n} \sum_{j=1}^{m} \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) \cdot 1_{\{X_j \le Y_k\}} \tag{5.22}$$

$$\begin{aligned} = a_N^{-2} \cdot \Big[& m^{-1} \cdot \sum_{i=1}^{m} \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx) \cdot 1_{\{X_i \le y\}}\, G(dy) \\ & + n^{-1} \cdot \sum_{k=1}^{n} \iint K'\left(a_N^{-1}(H_N(x) - H_N(Y_k))\right)\, F(dx) \cdot 1_{\{y \le Y_k\}}\, F(dy) \\ & - \iiint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx) \cdot 1_{\{z \le y\}}\, G(dy)\, F(dz) \Big] + O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1}). \end{aligned}$$

Proof. Define

$$u_N(r, s) = a_N^{-2} \int K'\left(a_N^{-1}(H_N(x) - H_N(r))\right) F(dx) \cdot 1_{\{s \le r\}},$$

and define the generalized U-statistic $U_{m,n}$ as

$$U_{m,n} = m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} u_N(Y_k, X_i).$$

The rest of the proof is completely analogous to the proof of Lemma 5.2 with the kernel function $u_N(X_i, Y_k)$ replaced by $u_N(Y_k, X_i)$. $\square$

LEMMA 5.6.

$$a_N^{-2} \cdot n^{-2} \cdot \sum_{1\le k\neq l\le m} \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) \cdot 1_{\{Y_l \le Y_k\}} \tag{5.23}$$
$$= a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \Big[\iint K'\left(a_N^{-1}(H_N(x) - H_N(Y_k))\right)\, F(dx) \cdot 1_{\{y \le Y_k\}}\, G(dy)$$
$$+ \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx) \cdot 1_{\{Y_k \le y\}}\, G(dy)$$
$$- \iiint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx) \cdot 1_{\{z \le y\}}\, G(dy)\, G(dz)\Big] + O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1}).$$

PROOF. Define

$$u_N(r, s) = a_N^{-2} \int K'\left(a_N^{-1}(H_N(x) - H_N(r))\right) F(dx) \cdot 1_{\{s \le r\}},$$

and define the U-statistic U_m as

$$U_n = n^{-1}(n-1)^{-1} \cdot \sum_{1\le k\neq l\le m} u_N(Y_k, Y_l).$$

Then

$$a_N^{-2} \cdot n^{-2} \cdot \sum_{1\le k\neq l\le m} \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) \cdot 1_{\{Y_l \le Y_k\}} = (n-1)n^{-1} \cdot U_n.$$

The rest of the proof is identical to the proof of Lemma 5.1 with m, i, j, X_i, X_j, $F(dy)$ and $F(dz)$ replaced by n, k, l, Y_k, Y_l, $G(dy)$ and $G(dz)$ respectively. □

LEMMA 5.7.

$$a_N^{-1} \cdot n^{-1} \cdot \sum_{k=1}^{n} \iint_{a_N^{-1}(H_N(x) - H_N(Y_k))}^{a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))} \left(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k)) - t\right) \cdot K''(t)\, dt\, F(dx) = O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1}). \tag{5.24}$$

PROOF. The proof is identical to the proof of Lemma 5.3 with k, n and Y_k in place of i, m, and X_i. □

We can apply (5.14), (5.17) and (5.21) together with Lemmas 5.5, 5.6 and 5.7 to express (5.2) as an i.i.d sum plus negligible rest terms:

LEMMA 5.8.

$$\int [\hat{g}_N - \bar{g}_N] \circ H_N(x)\, F(dx) =$$
$$= a_N^{-1} \cdot n^{-1} \cdot \sum_{k=1}^{n} \Big[\int K\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx)$$

$$
\begin{aligned}
&- \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\, F(dx)\Big] \\
&- (1-\lambda_N)\cdot a_N^{-2}\cdot n^{-1}\cdot \sum_{k=1}^{n} \Big[\iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx)\cdot 1_{\{Y_k \le y\}}\, G(dy) \\
&- \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx)\cdot G(y)\, G(dy)\Big] \\
&- \lambda_N \cdot a_N^{-2}\cdot m^{-1}\cdot \sum_{i=1}^{m} \Big[\iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx)\cdot 1_{\{X_i \le y\}}\, G(dy) \\
&- \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx)\cdot F(y)\, G(dy)\Big] \\
&+ O_{\mathbb{P}}(a_N^{-2}\cdot N^{-1}).
\end{aligned}
$$

Proof. Combine (5.14), (5.17) and (5.21) together with Lemmas 5.5 and 5.6. The proof is identical to the proof of Lemma 5.4 with n, m, k, i, Y_k, X_i, λ_N, $(1-\lambda_N)$, $G(dy)$, $G(dz)$, and $F(dy)$ in place of m, n, i, k, X_i, Y_k, λ_N and $(1-\lambda_N)$, $F(dy)$, $F(dz)$, and $G(dy)$ respectively. □

Finally, we combine Lemmas 5.4 and 5.8 to get the desired i.i.d. sum representation for (2.44), which yields

Lemma 5.9.

$$
\begin{aligned}
&\int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big]\circ H_N(x)\, F(dx) \\
&= a_N^{-1}\cdot m^{-1}\cdot \sum_{i=1}^{m} \Big[\int K\big(a_N^{-1}(H_N(x) - H_N(X_i))\big)\, F(dx) && (5.25)\\
&- \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\, F(dx)\Big] && (5.26)\\
&- \lambda_N \cdot a_N^{-2}\cdot m^{-1}\cdot \sum_{i=1}^{m} \Big[\iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx)\cdot 1_{\{X_i \le y\}}\, F(dy) && (5.27)\\
&- \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx)\cdot F(y)\, F(dy)\Big] && (5.28)\\
&- (1-\lambda_N)\cdot a_N^{-2}\cdot n^{-1}\cdot \sum_{k=1}^{n} \Big[\iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx)\cdot 1_{\{Y_k \le y\}}\, F(dy) && (5.29)\\
&- \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx)\cdot G(y)\, F(dy)\Big] && (5.30)\\
&- a_N^{-1}\cdot n^{-1}\cdot \sum_{k=1}^{n} \Big[\int K\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) && (5.31)\\
&- \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\, F(dx)\Big] && (5.32)\\
&+ (1-\lambda_N)\cdot a_N^{-2}\cdot n^{-1}\cdot \sum_{k=1}^{n} \Big[\iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx)\cdot 1_{\{Y_k \le y\}}\, G(dy) && (5.33)\\
&- \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right)\, F(dx)\cdot G(y)\, G(dy)\Big] && (5.34)
\end{aligned}
$$

$$+ \lambda_N \cdot a_N^{-2} \cdot m^{-1} \cdot \sum_{i=1}^{m} \Big[\iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx) \cdot 1_{\{X_i \le y\}} G(dy) \tag{5.35}$$

$$- \iint K'\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx) \cdot F(y)\, G(dy)\Big] \tag{5.36}$$

$$+ O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1}),$$

which can also be expressed more simply using integral notation as

$$\begin{aligned}
&\int \bar{f}_N \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big] \\
&- \int \bar{f}_N \circ H_N(x)\Big[\hat{G}_m(dx) - G(dx)\Big] \\
&+ \int \bar{f}'_N \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\Big[F(dx) - G(dx)\Big] \\
&+ O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1}).
\end{aligned}$$

5.2. Negligible terms

The proofs in this section will deal with terms which are *asymptotically negligible* under H_1 as well as under H_0, meaning that they are stochastically bounded so that they converge in probability to 0 even after $S_N(\hat{b}_N)$ has been properly scaled to ensure convergence in distribution.

We begin by bounding the Taylor rest terms generated in our expansion of $S_N(\hat{b}_N)$. The following sections will deal with further terms from the expansion that turn out to be negligible as well.

5.2.1. Taylor rest terms from the expansion of S_N. The following lemma shows that the four Taylor rest terms (2.37), (2.40), (2.43) and (2.46) that appear in the expansion of $S_N(\hat{b}_N)$ shown in the proof of Theorem 2.1 can be combined into a simpler single integral representation which is shown to be asymptotically negligible.

LEMMA 5.10. *Let F and G be continuous distribution functions and H_N and $\hat{H}_N$ be mixed theoretical and empirical distribution functions for sample sizes m and n defined as in* (2.4) *and* (2.16). *Further, let K be a kernel on* $(-1, 1)$ *and a_N a bandwidth sequence satisfying* (2.5) *through* (2.11) *and let $\hat{f}_N$ and $\hat{g}_N$ be kernel estimators* (2.12) *and* (2.13) *and $\bar{f}_N$ and $\bar{g}_N$ be functions on the interval* $(0, 1)$ *defined as in* (2.17) *and* (2.18) *respectively. Then*

$$\begin{aligned}
&\int \int_{H_N(x)}^{\hat{H}_N(x)} \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big]''(t) \cdot \big(\hat{H}_N(x) - t\big)\, dt \Big[\hat{F}_m(dx) - F(dx)\Big] \\
&+ \int \int_{H_N(x)}^{\hat{H}_N(x)} \Big[\bar{f}_N - \bar{g}_N\Big]''(t) \cdot \big(\hat{H}_N(x) - t\big)\, dt \Big[\hat{F}_m(dx) - F(dx)\Big] \\
&+ \int \int_{H_N(x)}^{\hat{H}_N(x)} \Big[\bar{f}_N - \bar{g}_N\Big]''(t) \cdot \big(\hat{H}_N(x) - t\big)\, dt\, F(dx)
\end{aligned}$$

$$+\int\int_{H_N(x)}^{\hat{H}_N(x)}\Big[\hat{f}_N-\hat{g}_N-(\bar{f}_N-\bar{g}_N)\Big]''(t)\cdot\big(\hat{H}_N(x)-t\big)\,dt\,F(dx)$$
$$=O_{\mathbb{P}}(a_N^{-2}\cdot N^{-1}).$$

PROOF.

$$\int\int_{H_N(x)}^{\hat{H}_N(x)}\Big[\hat{f}_N-\hat{g}_N-(\bar{f}_N-\bar{g}_N)\Big]''(t)\cdot\big(\hat{H}_N(x)-t\big)\,dt\,\Big[\hat{F}_m(dx)-F(dx)\Big]$$
$$+\int\int_{H_N(x)}^{\hat{H}_N(x)}\Big[\bar{f}_N-\bar{g}_N\Big]''(t)\cdot\big(\hat{H}_N(x)-t\big)\,dt\,\Big[\hat{F}_m(dx)-F(dx)\Big]$$
$$+\int\int_{H_N(x)}^{\hat{H}_N(x)}\Big[\bar{f}_N-\bar{g}_N\Big]''(t)\cdot\big(\hat{H}_N(x)-t\big)\,dt\,F(dx)$$
$$+\int\int_{H_N(x)}^{\hat{H}_N(x)}\Big[\hat{f}_N-\hat{g}_N-(\bar{f}_N-\bar{g}_N)\Big]''(t)\cdot\big(\hat{H}_N(x)-t\big)\,dt\,F(dx)$$
$$=\int\int_{H_N(x)}^{\hat{H}_N(x)}\Big[\hat{f}_N-\hat{g}_N\Big]''(t)\cdot\big(\hat{H}_N(x)-t\big)\,dt\,\hat{F}_m(dx)\tag{5.37}$$
$$=m^{-1}\cdot\sum_{i=1}^{m}\int_{H_N(X_i)}^{\hat{H}_N(X_i)}\Big[\hat{f}_N-\hat{g}_N\Big]''(t)\cdot\big(\hat{H}_N(X_i)-t\big)\,dt$$
$$=m^{-1}\cdot\sum_{i=1}^{m}\int_{H_N(X_i)}^{\hat{H}_N(X_i)}\Big[a_N^{-3}\cdot m^{-1}\cdot\sum_{j=1}^{m}K''\big(a_N^{-1}(t-\hat{H}_N(X_j))\big)$$
$$-a_N^{-3}\cdot n^{-1}\cdot\sum_{k=1}^{n}K''\big(a_N^{-1}(t-\hat{H}_N(Y_k))\big)\Big]\cdot\big(\hat{H}_N(X_i)-t\big)\,dt$$
$$=a_N^{-3}\cdot m^{-2}\cdot\sum_{i=1}^{m}\sum_{j=1}^{m}\int_{H_N(X_i)}^{\hat{H}_N(X_i)}K''\big(a_N^{-1}(t-\hat{H}_N(X_j))\big)\cdot\big(\hat{H}_N(X_i)-t\big)\,dt\tag{5.38}$$
$$-a_N^{-3}\cdot m^{-1}n^{-1}\cdot\sum_{i=1}^{m}\sum_{k=1}^{n}\int_{H_N(X_i)}^{\hat{H}_N(X_i)}K''\big(a_N^{-1}(t-\hat{H}_N(Y_k))\big)\cdot\big(\hat{H}_N(X_i)-t\big)\,dt.\tag{5.39}$$

We consider (5.38) and (5.39) in turn, showing that each are bounded by $O_{\mathbb{P}}(a_N^{-2}\cdot N^{-1})$ sequences. For (5.38) we have:

$$a_N^{-3}\cdot m^{-2}\cdot\sum_{i=1}^{m}\sum_{j=1}^{m}\int_{H_N(X_i)}^{\hat{H}_N(X_i)}K''\big(a_N^{-1}(t-\hat{H}_N(X_j))\big)\cdot\big(\hat{H}_N(X_i)-t\big)\,dt$$
$$=a_N^{-3}\cdot m^{-2}\cdot\sum_{i=1}^{m}\int_{H_N(X_i)}^{\hat{H}_N(X_i)}K''\big(a_N^{-1}(t-\hat{H}_N(X_i))\big)\cdot\big(\hat{H}_N(X_i)-t\big)\,dt\tag{5.40}$$
$$+a_N^{-3}\cdot m^{-2}\cdot\sum_{1\le i\ne j\le m}\int_{H_N(X_i)}^{\hat{H}_N(X_i)}K''\big(a_N^{-1}(t-\hat{H}_N(X_j))\big)\cdot\big(\hat{H}_N(X_i)-t\big)\,dt.\tag{5.41}$$

Now bounding (5.40) we have

$$\Big|a_N^{-3}\cdot m^{-2}\cdot\sum_{i=1}^{m}\int_{H_N(X_i)}^{\hat{H}_N(X_i)}K''\big(a_N^{-1}(t-\hat{H}_N(X_i))\big)\cdot\big(\hat{H}_N(X_i)-t\big)\,dt\Big|$$

$$
\begin{aligned}
&= \Bigg| a_N^{-3} \cdot m^{-2} \Bigg[\sum_{\substack{1 \le i \le m \\ H_N(X_i) \le \hat{H}_N(X_i)}} \int_{H_N(X_i)}^{\hat{H}_N(X_i)} K''\big(a_N^{-1}(t - \hat{H}_N(X_i))\big) \cdot \big(\hat{H}_N(X_i) - t\big)\, dt \\
&\quad - \sum_{\substack{1 \le i \le m \\ H_N(X_i) > \hat{H}_N(X_i)}} \int_{\hat{H}_N(X_i)}^{H_N(X_i)} K''\big(a_N^{-1}(t - \hat{H}_N(X_i))\big) \cdot \big(\hat{H}_N(X_i) - t\big)\, dt \Bigg] \Bigg| \\
&\le a_N^{-3} \cdot m^{-2} \Bigg[\sum_{\substack{1 \le i \le m \\ H_N(X_i) \le \hat{H}_N(X_i)}} \int_{H_N(X_i)}^{\hat{H}_N(X_i)} \big|K''\big(a_N^{-1}(t - \hat{H}_N(X_i))\big)\big| \cdot \big|\hat{H}_N(X_i) - t\big|\, dt \\
&\quad + \sum_{\substack{1 \le i \le m \\ H_N(X_i) > \hat{H}_N(X_i)}} \int_{\hat{H}_N(X_i)}^{H_N(X_i)} \big|K''\big(a_N^{-1}(t - \hat{H}_N(X_i))\big)\big| \cdot \big|\hat{H}_N(X_i) - t\big|\, dt \Bigg] \\
&\le a_N^{-3} \cdot m^{-2} \Bigg[\sum_{\substack{1 \le i \le m \\ H_N(X_i) \le \hat{H}_N(X_i)}} \big\|K''\big\| \cdot \int_{H_N(X_i)}^{\hat{H}_N(X_i)} \big|\hat{H}_N(X_i) - t\big|\, dt \\
&\quad + \sum_{\substack{1 \le i \le m \\ H_N(X_i) > \hat{H}_N(X_i)}} \big\|K''\big\| \cdot \int_{\hat{H}_N(X_i)}^{H_N(X_i)} \big|\hat{H}_N(X_i) - t\big|\, dt \Bigg] \\
&\le a_N^{-3} \cdot m^{-2} \cdot \sum_{i=1}^{m} \big\|K''\big\| \cdot \big|\hat{H}_N(X_i) - H_N(X_i)\big|^2 \\
&\le a_N^{-3} \cdot m^{-1} \cdot \big\|K''\big\| \cdot \big\|\hat{H}_N - H_N\big\|^2 \\
&= O_{\mathbb{P}}(a_N^{-3} \cdot N^{-2}).
\end{aligned}
$$

We use the L^1 norm to show that (5.41) is $O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1})$. First, define $\hat{H}_N^*$ to be equal to $\hat{H}_N$ with X_1, X_2, Y_1 and Y_2 removed from the sample. That is,

$$
\hat{H}_N^*(x) = N^{-1} \cdot \Bigg[\sum_{i=3}^{m} 1_{\{X_i \le x\}} + \sum_{k=3}^{n} 1_{\{Y_k \le x\}} \Bigg].
$$

Now recalling that K is zero outside of $(-1, 1)$ we have for (5.41):

$$
\begin{aligned}
&\mathbb{E}\Bigg[\Bigg| a_N^{-3} \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \int_{H_N(X_i)}^{\hat{H}_N(X_i)} K''\big(a_N^{-1}(t - \hat{H}_N(X_j))\big) \cdot \big(\hat{H}_N(X_i) - t\big)\, dt \Bigg| \Bigg] \\
&\quad \le a_N^{-3} \cdot m^{-2} \cdot \mathbb{E}\Bigg[\sum_{\substack{1 \le i \ne j \le m \\ H_N(X_i) \le \hat{H}_N(X_i)}} \Bigg| \int_{H_N(X_i)}^{\hat{H}_N(X_i)} K''\big(a_N^{-1}(t - \hat{H}_N(X_j))\big) \cdot \big(\hat{H}_N(X_i) - t\big)\, dt \Bigg| \\
&\qquad + \sum_{\substack{1 \le i \ne j \le m \\ H_N(X_i) > \hat{H}_N(X_i)}} \Bigg| \int_{\hat{H}_N(X_i)}^{H_N(X_i)} K''\big(a_N^{-1}(t - \hat{H}_N(X_j))\big) \cdot \big(\hat{H}_N(X_i) - t\big)\, dt \Bigg| \Bigg] \\
&\quad = a_N^{-3} \cdot m^{-2} \cdot \mathbb{E}\Bigg[\sum_{\substack{1 \le i \ne j \le m \\ H_N(X_i) \le \hat{H}_N(X_i)}} \Bigg| \int_{H_N(X_i)}^{\hat{H}_N(X_i)} 1_{\{\hat{H}_N(X_j) - a_N \,<\, t \,<\, \hat{H}_N(X_j) + a_N\}}
\end{aligned}
$$

$$\times K''\big(a_N^{-1}(t-\hat{H}_N(X_j))\big)\cdot\big(\hat{H}_N(X_i)-t\big)\,dt\Big|$$

$$+\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)>\hat{H}_N(X_i)}}\Big|\int_{\hat{H}_N(X_i)}^{H_N(X_i)}1_{\{\hat{H}_N(X_j)-a_N\,<\,t\,<\,\hat{H}_N(X_j)+a_N\}}\cdot K''\big(a_N^{-1}(t-\hat{H}_N(X_j))\big)\cdot\big(\hat{H}_N(X_i)-t\big)\,dt\Big|\Big]$$

$$\le a_N^{-3}\cdot m^{-2}\cdot\mathbb{E}\Big[\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)\le\hat{H}_N(X_i)}}\|K''\|\cdot\int_{H_N(X_i)}^{\hat{H}_N(X_i)}1_{\{\hat{H}_N(X_j)-a_N\,<\,t\,<\,\hat{H}_N(X_j)+a_N\}}\cdot\big|\hat{H}_N(X_i)-t\big|\,dt$$

$$+\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)>\hat{H}_N(X_i)}}\|K''\|\cdot\int_{\hat{H}_N(X_i)}^{H_N(X_i)}1_{\{\hat{H}_N(X_j)-a_N\,<\,t\,<\,\hat{H}_N(X_j)+a_N\}}\cdot\big|\hat{H}_N(X_i)-t\big|\,dt\Big]$$

$$=\|K''\|\cdot a_N^{-3}\cdot m^{-2}\cdot\Big[\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)\le\hat{H}_N(X_i)}}\mathbb{E}\Big[\int_{H_N(X_1)}^{\hat{H}_N(X_1)}1_{\{\hat{H}_N(X_2)-a_N\,<\,t\,<\,\hat{H}_N(X_2)+a_N\}}\cdot\big|\hat{H}_N(X_1)-t\big|\,dt\Big]$$

$$+\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)>\hat{H}_N(X_i)}}\mathbb{E}\Big[\int_{\hat{H}_N(X_1)}^{H_N(X_1)}1_{\{\hat{H}_N(X_2)-a_N\,<\,t\,<\,\hat{H}_N(X_2)+a_N\}}\cdot\big|\hat{H}_N(X_1)-t\big|\,dt\Big]\Big]$$

$$\le\|K''\|\cdot a_N^{-3}\cdot m^{-2}\cdot\Big[\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)\le\hat{H}_N(X_i)}}\mathbb{E}\Big[\big|\hat{H}_N(X_1)-H_N(X_1)\big|$$

$$\times\int 1_{\{H_N(X_1)\,<\,t\,<\,\hat{H}_N(X_1)\}}\cdot 1_{\{\hat{H}_N(X_2)-a_N\,<\,t\,<\,\hat{H}_N(X_2)+a_N\}}\,dt\Big]$$

$$+\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)>\hat{H}_N(X_i)}}\mathbb{E}\Big[\big|\hat{H}_N(X_1)-H_N(X_1)\big|\cdot\int 1_{\{\hat{H}_N(X_1)\,<\,t\,<\,H_N(X_1)\}}\cdot 1_{\{\hat{H}_N(X_2)-a_N\,<\,t\,<\,\hat{H}_N(X_2)+a_N\}}\,dt\Big]\Big]$$

$$\le\|K''\|\cdot a_N^{-3}\cdot m^{-2}\cdot\Big[\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)\le\hat{H}_N(X_i)}}\mathbb{E}\Big[\Big[\big|\hat{H}_N(X_1)-\hat{H}_N^*(X_1)\big|+\big|\hat{H}_N^*(X_1)-H_N(X_1)\big|\Big]$$

$$\times\int 1_{\{H_N(X_1)\,<\,t\,<\,\hat{H}_N^*(X_1)+|\hat{H}_N(X_1)-\hat{H}_N^*(X_1)|\}}\cdot 1_{\{t-a_N\,<\,\hat{H}_N(X_2)\,<\,t+a_N\}}\,dt\Big]$$

$$+\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)>\hat{H}_N(X_i)}}\mathbb{E}\Big[\Big[\big|\hat{H}_N(X_1)-\hat{H}_N^*(X_1)\big|+\big|\hat{H}_N^*(X_1)-H_N(X_1)\big|\Big]$$

$$\times\int 1_{\{\hat{H}_N^*(X_1)-|\hat{H}_N(X_1)-\hat{H}_N^*(X_1)|\,<\,t\,<\,H_N(X_1)\}}\cdot 1_{\{t-a_N\,<\,\hat{H}_N(X_2)\,<\,t+a_N\}}\,dt\Big]\Big]$$

$$\le\|K''\|\cdot a_N^{-3}\cdot m^{-2}\cdot\Big[\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)\le\hat{H}_N(X_i)}}\mathbb{E}\Big[\Big[\|\hat{H}_N^*-H_N\|+4N^{-1}\Big]\cdot\int 1_{\{H_N(X_1)\,<\,t\,<\,\hat{H}_N^*(X_1)+4N^{-1}\}}$$

$$\times 1_{\{t-a_N-|H_N(X_2)-\hat{H}_N^*(X_2)|-|\hat{H}_N^*(X_2)-\hat{H}_N(X_2)|\,<\,H_N(X_2)\}}$$

$$\times 1_{\{H_N(X_2)\,<\,t+a_N+|H_N(X_2)-\hat{H}_N^*(X_2)|+|\hat{H}_N^*(X_2)-\hat{H}_N(X_2)|\}}\,dt\Big]$$

$$
\begin{aligned}
&+\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)>\hat H_N(X_i)}} \mathbb{E}\Bigg[\Big[\big\|\hat H_N^* - H_N\big\| + 4N^{-1}\Big]\cdot \int 1_{\{\hat H_N^*(X_1)-4N^{-1} < t < H_N(X_1)\}}\\
&\qquad \times 1_{\{t-a_N-|H_N(X_2)-\hat H_N^*(X_2)|-|\hat H_N^*(X_2)-\hat H_N(X_2)| < H_N(X_2)\}}\\
&\qquad \times 1_{\{H_N(X_2) < t+a_N+|H_N(X_2)-\hat H_N^*(X_2)|+|\hat H_N^*(X_2)-\hat H_N(X_2)|\}}\,dt\Bigg]\Bigg]\\
&\le \big\|K''\big\|\cdot a_N^{-3}\cdot m^{-2}\cdot \Bigg[\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)\le\hat H_N(X_i)}} \mathbb{E}\Bigg[\Big[\big\|\hat H_N^* - H_N\big\| + 4N^{-1}\Big]\cdot \int 1_{\{H_N(X_1) < t < \hat H_N^*(X_1)+4N^{-1}\}}\\
&\qquad \times 1_{\{t-a_N-\|H_N-\hat H_N^*\|-4N^{-1} < H_N(X_2) < t+a_N+\|H_N-\hat H_N^*\|+4N^{-1}|\}}\,dt\Bigg]\\
&\quad +\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)>\hat H_N(X_i)}} \mathbb{E}\Bigg[\Big[\big\|\hat H_N^* - H_N\big\| + 4N^{-1}\Big]\cdot \int 1_{\{\hat H_N^*(X_1)-4N^{-1} < t < H_N(X_1)\}}\\
&\qquad \times 1_{\{t-a_N-\|H_N-\hat H_N^*\|-4N^{-1} < H_N(X_2) < t+a_N+\|H_N-\hat H_N^*\|+4N^{-1}|\}}\,dt\Bigg]\Bigg]\\
&= \big\|K''\big\|\cdot a_N^{-3}\cdot m^{-2}\cdot \Bigg[\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)\le\hat H_N(X_i)}} \mathbb{E}\Bigg[\mathbb{E}\Bigg[\Big[\big\|\hat H_N^* - H_N\big\| + 4N^{-1}\Big]\cdot \int 1_{\{H_N(X_1) < t < \hat H_N^*(X_1)+4N^{-1}\}}\\
&\qquad \times 1_{\{t-a_N-\|H_N-\hat H_N^*\|-4N^{-1} < H_N(X_2) < t+a_N+\|H_N-\hat H_N^*\|+4N^{-1}|\}}\,dt\,\Bigg|\,X_1,X_3,\dots,X_m,Y_3,\dots,Y_n\Bigg]\Bigg]\\
&\quad +\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)>\hat H_N(X_i)}} \mathbb{E}\Bigg[\mathbb{E}\Bigg[\Big[\big\|\hat H_N^* - H_N\big\| + 4N^{-1}\Big]\cdot \int 1_{\{\hat H_N^*(X_1)-4N^{-1} < t < H_N(X_1)\}}\\
&\qquad \times 1_{\{t-a_N-\|H_N-\hat H_N^*\|-4N^{-1} < H_N(X_2) < t+a_N+\|H_N-\hat H_N^*\|+4N^{-1}|\}}\,dt\,\Bigg|\,X_1,X_3,\dots,X_m,Y_3,\dots,Y_n\Bigg]\Bigg]\Bigg]\\
&= \big\|K''\big\|\cdot a_N^{-3}\cdot m^{-2}\cdot \Bigg[\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)\le\hat H_N(X_i)}} \mathbb{E}\Bigg[\Big[\big\|\hat H_N^* - H_N\big\| + 4N^{-1}\Big]\cdot \int_0^1\int 1_{\{H_N(X_1) < t < \hat H_N^*(X_1)+4N^{-1}\}}\\
&\qquad \times 1_{\{t-a_N-\|H_N-\hat H_N^*\|-4N^{-1} < v < t+a_N+\|H_N-\hat H_N^*\|+4N^{-1}|\}}\,dt\,f_N(v)\,dv\Bigg]\\
&\quad +\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)>\hat H_N(X_i)}} \mathbb{E}\Bigg[\Big[\big\|\hat H_N^* - H_N\big\| + 4N^{-1}\Big]\cdot \int_0^1\int 1_{\{\hat H_N^*(X_1)-4N^{-1} < t < H_N(X_1)\}}\\
&\qquad \times 1_{\{t-a_N-\|H_N-\hat H_N^*\|-4N^{-1} < v < t+a_N+\|H_N-\hat H_N^*\|+4N^{-1}|\}}\,dt\,f_N(v)\,dv\Bigg]\Bigg]\\
&= \big\|K''\big\|\cdot a_N^{-3}\cdot m^{-2}\cdot \Bigg[\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)\le\hat H_N(X_i)}} \mathbb{E}\Bigg[\Big[\big\|\hat H_N^* - H_N\big\| + 4N^{-1}\Big]\cdot \int 1_{\{H_N(X_1) < t < \hat H_N^*(X_1)+4N^{-1}\}}\\
&\qquad \times \int_0^1 1_{\{t-a_N-\|H_N-\hat H_N^*\|-4N^{-1} < v < t+a_N+\|H_N-\hat H_N^*\|+4N^{-1}|\}}\cdot f_N(v)\,dv\,dt\Bigg]\\
&\quad +\sum_{\substack{1\le i\neq j\le m\\ H_N(X_i)>\hat H_N(X_i)}} \mathbb{E}\Bigg[\Big[\big\|\hat H_N^* - H_N\big\| + 4N^{-1}\Big]\cdot \int 1_{\{\hat H_N^*(X_1)-4N^{-1} < t < H_N(X_1)\}}
\end{aligned}
$$

$$
\begin{aligned}
&\quad \times \int_0^1 1_{\{t-a_N-\|H_N-\hat{H}_N^*\|-4N^{-1} < v < t+a_N+\|H_N-\hat{H}_N^*\|+4N^{-1}|\}} \cdot f_N(v)\, dv\, dt\Big]\Big] \\
&\le \big\|K''\big\| \cdot a_N^{-3} \cdot m^{-2} \cdot \Bigg[\sum_{\substack{1\le i\neq j\le m \\ H_N(X_i)\le \hat{H}_N(X_i)}} \mathbb{E}\Big[\Big[\big\|\hat{H}_N^* - H_N\big\| + 4N^{-1}\Big] \\
&\qquad \times \Big[\hat{H}_N^*(X_1) - H_N(X_1) + 4N^{-1}\Big] \cdot 2\big\|f_N\big\| \cdot \Big[a_N + \big\|H_N - \hat{H}_N^*\big\| + 4N^{-1}\Big]\Big] \\
&\quad + \sum_{\substack{1\le i\neq j\le m \\ H_N(X_i)> \hat{H}_N(X_i)}} \mathbb{E}\Big[\Big[\big\|\hat{H}_N^* - H_N\big\| + 4N^{-1}\Big] \\
&\qquad \times \Big[H_N(X_1) - \hat{H}_N^*(X_1) + 4N^{-1}\Big] \cdot 2\big\|f_N\big\| \cdot \Big[a_N + \big\|H_N - \hat{H}_N^*\big\| + 4N^{-1}\Big]\Big]\Bigg] \\
&\le \big\|K''\big\| \cdot a_N^{-3} \cdot m(m-1)m^{-2} \cdot \mathbb{E}\Big[\Big[\big\|\hat{H}_N^* - H_N\big\| + 4N^{-1}\Big]^2 \cdot 2\big\|f_N\big\| \cdot \Big[a_N + \big\|H_N - \hat{H}_N^*\big\| + 4N^{-1}\Big]\Big] \\
&\le \big\|K''\big\| \cdot a_N^{-3} \cdot m(m-1)m^{-2} \cdot 2\big\|f_N\big\| \\
&\qquad \times \Big[\mathbb{E}\Big[\big\|\hat{H}_N^* - H_N\big\| + 4N^{-1}\Big]^4\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\Big[a_N + \big\|H_N - \hat{H}_N^*\big\| + 4N^{-1}\Big]^2\Big]^{\frac{1}{2}} \\
&\le \big\|K''\big\| \cdot a_N^{-3} \cdot m(m-1)m^{-2} \cdot 2\big\|f_N\big\| \\
&\qquad \times \Big[8\,\mathbb{E}\Big[\big\|\hat{H}_N^* - H_N\big\|^4 + 4N^{-4}\Big]\Big]^{\frac{1}{2}} \cdot \Big[4\,\mathbb{E}\Big[a_N^2 + \big\|H_N - \hat{H}_N^*\big\|^2 + 4N^{-2}\Big]\Big]^{\frac{1}{2}} \\
&= \big\|K''\big\| \cdot a_N^{-3} \cdot m(m-1)m^{-2} \cdot 2\big\|f_N\big\| \\
&\qquad \times 8^{\frac{1}{2}} \Big[\mathbb{E}\Big[\big\|\hat{H}_N^* - H_N\big\|\Big]^4 + 4N^{-4}\Big]^{\frac{1}{2}} \cdot 2\,\Big[\mathbb{E}\Big[a_N\Big]^2 + \mathbb{E}\Big[\big\|H_N - \hat{H}_N^*\big\|\Big]^2 + 4N^{-2}\Big]^{\frac{1}{2}} \\
&= \big\|K''\big\| \cdot a_N^{-3} \cdot O(1) \cdot \Big[O(N^{-2}) + O(N^{-4})\Big]^{\frac{1}{2}} \cdot \Big[O(a_N^2) + O(N^{-1}) + O(N^{-2})\Big]^{\frac{1}{2}} \\
&= a_N^{-3} \cdot O(1) \cdot O(N^{-1}) \cdot O(a_N) \\
&= O(a_N^{-2} \cdot N^{-1}).
\end{aligned}
$$

Thus, we have shown that (5.40) and (5.41) are $O_{\mathbb{P}}(a_N^{-3} \cdot N^{-2})$ and $O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1})$ respectively so that the first of the two sums that make up the total rest term (5.37) is $O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1})$.

To see that the second sum (5.39) is $O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1})$ as well simply replace the scaled summation $m^{-2} \cdot \sum_{1\le i\neq j\le m}$ with $m^{-1}n^{-1} \cdot \sum_{1\le i\le m,\, 1\le k\le n}$, and X_j, X_2 and f_N by Y_k, Y_1 and g_N respectively in the above proof showing that (5.41) is $O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1})$. Then, altogether, we have (5.37) is $O_{\mathbb{P}}(a_N^{-2} \cdot N^{-1})$ as claimed. □

The terms (2.35), (2.36), (2.39) and (2.45) are leading terms from the expansion of $S_N(\hat{b}_N)$ in the proof of Theorem 2.1 that will turn out to be negligible as well. In the following sections we will consider each of these terms in turn and use similar techniques in each case to properly bound them.

All results in this section are proven using the same assumptions on K and a_N and definitions as in Lemma 5.10.

5.2.2. First bounded term. Beginning with (2.35) we can write

$$\int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big] \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big]$$

$$= \int \Big[\hat{f}_N - \bar{f}_N\Big] \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big] \tag{5.42}$$

$$- \int \Big[\hat{g}_N - \bar{g}_N\Big] \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big] \tag{5.43}$$

We will first work at bounding (5.42). The proof for (5.43) follows along similar lines. The approach will be to show that our favorable choice of centering functions $\bar{f}_N$ and $\bar{g}_N$ will mean that (5.42) can be written as a degenerate U-statistic plus a negligible rest.

Recall our definitions of $\bar{f}_N$ and $\bar{g}_N$:

$$\bar{f}_N(t) = a_N^{-1} \int K\left(\frac{t - H_N(y)}{a_N}\right) F(dy), \quad 0 \le t \le 1, \tag{5.44}$$

$$\bar{g}_N(t) = a_N^{-1} \int K\left(\frac{t - H_N(y)}{a_N}\right) G(dy), \quad 0 \le t \le 1. \tag{5.45}$$

Then

$$\begin{aligned}
&\int \Big[\hat{f}_N - \bar{f}_N\Big] \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big] \\
&= m^{-1} \cdot \sum_{i=1}^{m} \Big[\Big[\hat{f}_N - \bar{f}_N\Big] \circ H_N(X_i) - \int \Big[\hat{f}_N - \bar{f}_N\Big] \circ H_N(x)\, F(dx)\Big] \\
&= m^{-1} \cdot \sum_{i=1}^{m} \Big[(m \cdot a_N)^{-1} \cdot \sum_{j=1}^{m} K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(X_j))\big) - \bar{f}_N \circ H_N(X_i) \\
&\qquad - \int (m \cdot a_N)^{-1} \cdot \sum_{j=1}^{m} K\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_j))\big)\, F(dx) + \int \bar{f}_N \circ H_N(x)\, F(dx)\Big] \\
&= m^{-2} \cdot \sum_{i=1}^{m}\sum_{j=1}^{m} \Big[a_N^{-1} K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(X_j))\big) - \bar{f}_N \circ H_N(X_i) \\
&\qquad - a_N^{-1} \int K\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_j))\big)\, F(dx) + \int \bar{f}_N \circ H_N(x)\, F(dx)\Big] \\
&= m^{-2} \cdot \sum_{i=1}^{m}\sum_{j=1}^{m} \Big[a_N^{-1} K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(X_j))\big) - a_N^{-1} \int K\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, F(dy) \\
&\qquad - a_N^{-1} \int K\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_j))\big)\, F(dx) + a_N^{-1} \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\, F(dx)\Big]
\end{aligned}$$

At this point we separate the summands with $i = j$ and use the Taylor expansion of the kernel function K about each of the $a_N^{-1}(H_N(X_i) - H_N(X_j))$ and $a_N^{-1}(H_N(x) - H_N(X_j))$ for the remaining summands with $i \neq j$, which yields

$$\int \Big[\hat{f}_N - \bar{f}_N\Big] \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big]$$

$$
\begin{aligned}
&= a_N^{-1} \cdot m^{-2} \cdot \sum_{i=1}^{m} \Big[K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(X_i))\big) - \int K\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, F(dy) \\
&\qquad - \int K\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_i))\big)\, F(dx) + \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\, F(dx) \Big] \quad (5.46) \\
&\quad + a_N^{-1} \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \Big[K\big(a_N^{-1}(H_N(X_i) - H_N(X_j))\big) - \int K\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, F(dy) \\
&\qquad - \int K\big(a_N^{-1}(H_N(x) - H_N(X_j))\big)\, F(dx) + \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\, F(dx) \Big] \quad (5.47) \\
&\quad + a_N^{-2} \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \Big[K'\big(a_N^{-1}(H_N(X_i) - H_N(X_j))\big)\big(H_N(X_j) - \hat{H}_N(X_j)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x) - H_N(X_j))\big)\, F(dx)\big(H_N(X_j) - \hat{H}_N(X_j)\big) \Big] \quad (5.48) \\
&\quad + a_N^{-1} \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \Big[\int_{a_N^{-1}(H_N(X_i) - H_N(X_j))}^{a_N^{-1}(H_N(X_i) - \hat{H}_N(X_j))} \big(a_N^{-1}(H_N(X_i) - \hat{H}_N(X_j)) - t\big) \cdot K''(t)\, dt \\
&\qquad - \iint_{a_N^{-1}(H_N(x) - H_N(X_j))}^{a_N^{-1}(H_N(x) - \hat{H}_N(X_j))} \big(a_N^{-1}(H_N(x) - \hat{H}_N(X_j)) - t\big) \cdot K''(t)\, dt\, F(dx) \Big] \quad (5.49)
\end{aligned}
$$

Remark 2. In many of the following lemmas (e.g. Lemmas 5.11, 5.12 5.14 and 5.15) we eschew deriving sharper bounds for the terms in question in favor of shorter, simpler proofs providing rough upper bounds. We will invest more effort in deriving bounds for terms (5.49) and (5.76), since these converge more slowly and thus play the role here of the "limiting" terms which determine the overall rates of convergence for (5.42) and (5.43).

Since we have assumed that the kernel function K is bounded, it is easy to see that (5.46) is $O(a_N^{-1} \cdot N^{-1})$, since

$$
\begin{aligned}
&= \Big| a_N^{-1} \cdot m^{-2} \cdot \sum_{i=1}^{m} \Big[K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(X_i))\big) - \int K\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, F(dy) \\
&\qquad - \int K\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_i))\big)\, F(dx) + \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\, F(dx) \Big] \Big| \\
&\le a_N^{-1} \cdot m^{-1} \cdot 4\|K\|. \quad (5.50)
\end{aligned}
$$

In the following lemmas we will derive bounds for the remaining three terms (5.47), (5.48) and (5.49).

Lemma 5.11.

$$
\begin{aligned}
&a_N^{-1} \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \Big[K\big(a_N^{-1}(H_N(X_i) - H_N(X_j))\big) - \int K\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, F(dy) \\
&\qquad - \int K\big(a_N^{-1}(H_N(x) - H_N(X_j))\big)\, F(dx) + \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\, F(dx) \Big] \\
&\quad = O_P(a_N^{-1} \cdot N^{-1}).
\end{aligned}
$$

PROOF. Define

$$u_N(s,\ t) = a_N^{-1} K\big(a_N^{-1}(H_N(s) - H_N(t))\big),$$

and define the U-statistic U_m as

$$U_m = m^{-1}(m-1)^{-1} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j).$$

Let $\hat{U}_m$ be the Hájek projection of U_m as defined in Lemma A.2. Then (5.47) is equal to

$$\frac{m-1}{m} \cdot \Big[U_m - \hat{U}_m\Big].$$

Applying the inequality in Lemma A.2 we have

$$\mathbb{E}\Big[U_m - \hat{U}_m\Big]^2 \le 2(m-1)m^{-3} \cdot \mathbb{E}\big[u_N^*(X_1, X_2)\big]^2$$

with u_N^* defined as

$$u_N^*(r,s) = u_N(r,s) - \int u_N(r,y)\, F(dy) - \int u_N(x,s)\, F(dx) + \iint u_N(x,y)\, F(dx)F(dy).$$

Thus it remains only to bound the expectation $\mathbb{E}\big[u_N^*(X_1,X_2)\big]^2$:

$$\begin{aligned}
&\mathbb{E}\big[u_N^*(X_1, X_2)\big]^2 \\
&\quad = \mathbb{E}\Bigg[a_N^{-1}\bigg[K\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) - \int K\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \\
&\qquad\qquad - \int K\big(a_N^{-1}(H_N(x) - H_N(X_2))\big)\, F(dx) + \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\, F(dx)\bigg]\Bigg]^2 \\
&\quad \le a_N^{-2} \cdot 4 \cdot \mathbb{E}\Bigg[\big[K\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big)\big]^2 + \bigg[\int K\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy)\bigg]^2 \\
&\qquad\qquad + \bigg[\int K\big(a_N^{-1}(H_N(x) - H_N(X_1))\big)\, F(dx)\bigg]^2 + \bigg[\iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\, F(dx)\bigg]^2\Bigg] \\
&\quad \le a_N^{-2} \cdot 4 \cdot \mathbb{E}\big[4 \cdot \big\|K\big\|^2\big] \\
&\quad = 16\big\|K\big\|^2 \cdot a_N^{-2}.
\end{aligned}$$

Altogether, this yields

$$\begin{aligned}
\mathbb{E}\Big[U_m - \hat{U}_m\Big]^2 &\le 2(m-1)m^{-3} \cdot \mathbb{E}\big[u_N^*(X_1, X_2)\big]^2 \\
&\le 2(m-1)m^{-3} \cdot 16\big\|K\big\|^2 \cdot a_N^{-2} \\
&= 32\big\|K\big\|^2 \cdot a_N^{-2} \cdot (m-1)m^{-3} \\
&= O(a_N^{-2} \cdot N^{-2}).
\end{aligned}$$

so that (5.47) is

$$\frac{m-1}{m} \cdot \Big[U_m - \hat{U}_m\Big] = O_P(a_N^{-1} \cdot N^{-1})$$

which completes the proof. □

Lemma 5.12.

$$a_N^{-2} \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \Big[K'\big(a_N^{-1}(H_N(X_i) - H_N(X_j))\big)\big(H_N(X_j) - \hat{H}_N(X_j)\big) - \int K'\big(a_N^{-1}(H_N(x) - H_N(X_j))\big)\, F(dx)\big(H_N(X_j) - \hat{H}_N(X_j)\big)\Big] = O_P(a_N^{-2} \cdot N^{-1}).$$

Proof. Define

$$u_N(s,\, t) = a_N^{-2} \cdot \Big[K'\big(a_N^{-1}(H_N(s) - H_N(t))\big) - \int K'\big(a_N^{-1}(H_N(x) - H_N(t))\big)\, F(dx)\Big].$$

Then

$$\begin{aligned}
& a_N^{-2} \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \Big[K'\big(a_N^{-1}(H_N(X_i) - H_N(X_j))\big)\big(H_N(X_j) - \hat{H}_N(X_j)\big) \\
& \quad - \int K'\big(a_N^{-1}(H_N(x) - H_N(X_j))\big)\, F(dx)\big(H_N(X_j) - \hat{H}_N(X_j)\big)\Big] \\
& = m^{-2} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j) \cdot \big[H_N(X_j) - \hat{H}_N(X_j)\big] \\
& = m^{-2} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j) \cdot \Big[H_N(X_j) - N^{-1} \Big[\sum_{l=1}^{m} 1_{\{X_l \le X_j\}} + \sum_{k=1}^{n} 1_{\{Y_k \le X_j\}} \Big]\Big] \\
& = m^{-2} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j) \cdot H_N(X_j) - m^{-2} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j) \cdot N^{-1} \sum_{l=1}^{m} 1_{\{X_l \le X_j\}} \\
& \quad - m^{-2} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j) \cdot N^{-1} \sum_{k=1}^{n} 1_{\{Y_k \le X_j\}} \\
& = m^{-2} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j) \cdot H_N(X_j) - \lambda_N \cdot m^{-3} \cdot \sum_{1 \le i \ne j \le m} \sum_{l=1}^{m} u_N(X_i, X_j) \cdot 1_{\{X_l \le X_j\}} \\
& \quad - (1 - \lambda_N) \cdot m^{-2} n^{-1} \cdot \sum_{1 \le i \ne j \le m} \sum_{k=1}^{n} u_N(X_i, X_j) \cdot 1_{\{Y_k \le X_j\}} \\
& = m^{-2} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j) \cdot H_N(X_j) - \lambda_N \cdot m^{-3} \cdot \sum_{\substack{1 \le i,\, j,\, l \le m \\ i \ne j,\, j \ne l \text{ and } i \ne l}} u_N(X_i, X_j) \cdot 1_{\{X_l \le X_j\}} \\
& \quad - \lambda_N \cdot m^{-3} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j) \cdot 1_{\{X_i \le X_j\}} - \lambda_N \cdot m^{-3} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j) \\
& \quad - (1 - \lambda_N) \cdot m^{-2} n^{-1} \cdot \sum_{1 \le i \ne j \le m} \sum_{k=1}^{n} u_N(X_i, X_j) \cdot 1_{\{Y_k \le X_j\}}
\end{aligned}$$

Define the U-statistics U_m^1, U_m^2 and the generalized U-statistic $U_{m,n}^3$ as

$$U_m^1 = m^{-1}(m-1)^{-1} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j) \cdot H_N(X_j),$$

$$U_m^2 = m^{-1}(m-1)^{-1}(m-2)^{-1} \cdot \sum_{\substack{1 \le i,\, j,\, l \le m \\ i \neq j,\, j \neq l \text{ and } i \neq l}} u_N(X_i, X_j) \cdot 1_{\{X_l \le X_j\}},$$

$$U_{m,n}^3 = m^{-1}(m-1)^{-1} n^{-1} \cdot \sum_{\substack{1 \le i \neq j \le m \\ 1 \le k \le n}} u_N(X_i, X_j) \cdot 1_{\{Y_k \le X_j\}},$$

and let $\hat{U}_m^1$, $\hat{U}_m^2$ and $\hat{U}_{m,n}^3$ be the Hájek projections of U_m^1, U_m^2 and $U_{m,n}^3$ respectively as defined in Lemmas A.2, A.4 and A.5. Then (5.48) is equal to

$$\begin{aligned} &\frac{m-1}{m} \cdot U_m^1 - \frac{\lambda_N \cdot (m-1)(m-2)}{m^2} \cdot U_m^2 \\ &\quad - \lambda_N \cdot m^{-3} \cdot \sum_{1 \le i \neq j \le m} u_N(X_i, X_j) \cdot 1_{\{X_i \le X_j\}} - \lambda_N \cdot m^{-3} \cdot \sum_{1 \le i \neq j \le m} u_N(X_i, X_j) \\ &\quad - \frac{(1-\lambda_N) \cdot (m-1)}{m} \cdot U_{m,n}^3. \end{aligned} \tag{5.51}$$

Now, the kernel function u_N is bounded:

$$\|u_N\| \le 2\|K'\| a_N^{-2}.$$

Which means for the sums in (5.51) we can write

$$\begin{aligned} &\Big| - \lambda_N \cdot m^{-3} \cdot \sum_{1 \le i \neq j \le m} u_N(X_i, X_j) \cdot 1_{\{X_i \le X_j\}} - \lambda_N \cdot m^{-3} \cdot \sum_{1 \le i \neq j \le m} u_N(X_i, X_j) \Big| \\ &\quad \le \lambda_N \cdot m^{-3} \cdot \sum_{1 \le i \neq j \le m} \big| u_N(X_i, X_j) \cdot 1_{\{X_i \le X_j\}} \big| + \lambda_N \cdot m^{-3} \cdot \sum_{1 \le i \neq j \le m} \big| u_N(X_i, X_j) \big| \\ &\quad \le \lambda_N \cdot m^{-3} \cdot \sum_{1 \le i \neq j \le m} \|u_N\| + \lambda_N \cdot m^{-3} \cdot \sum_{1 \le i \neq j \le m} \|u_N\| \\ &\quad = \frac{2\lambda_N \cdot m(m-1)}{m^3} \cdot \|u_N\| \\ &\quad = O(a_N^{-2} \cdot N^{-1}). \end{aligned} \tag{5.52}$$

Thus, we can partition (5.48) into the sum of three scaled U-statistics and a negligible rest:

$$\frac{m-1}{m} \cdot U_m^1 - \frac{\lambda_N \cdot (m-1)(m-2)}{m^2} \cdot U_m^2 - \frac{(1-\lambda_N) \cdot (m-1)}{m} \cdot U_{m,n}^3 + O(a_N^{-2} \cdot N^{-1}). \tag{5.53}$$

In the following we will show that (5.53) is $O(a_N^{-2} \cdot N^{-1})$ as well, which will complete the proof. Begin by calculating each of the projections $\hat{U}_m^1$, $\hat{U}_m^2$ and $\hat{U}_{m,n}^3$. Firstly,

$$\begin{aligned} \hat{U}_m^1 = m^{-1} \cdot \sum_{i=1}^m \Big[&\int u_N(X_i, y) \cdot H_N(y)\, F(dy) + \int u_N(x, X_i) \cdot H_N(X_i)\, F(dx) \\ &- \iint u_N(x, y) \cdot H_N(y)\, F(dx) F(dy) \Big]. \end{aligned}$$

Nextly, for $\lambda_N \cdot \hat{U}_m^2$ we have

$$\begin{aligned} \lambda_N \cdot \hat{U}_m^2 = \lambda_N \cdot m^{-1} \cdot \sum_{i=1}^m \Big[&\iint u_N(X_i, y) \cdot 1_{\{z \le y\}}\, F(dy) F(dz) + \iint u_N(x, X_i) \cdot 1_{\{z \le X_i\}}\, F(dx) F(dz) \\ &+ \iint u_N(x, y) \cdot 1_{\{X_i \le y\}}\, F(dx) F(dy) - 2 \iiint u_N(x, y) \cdot 1_{\{z \le y\}}\, F(dx) F(dy) F(dz) \Big] \end{aligned}$$

$$= m^{-1} \cdot \sum_{i=1}^{m} \Big[\int u_N(X_i, y) \cdot \lambda_N F(y)\, F(dy) + \int u_N(x, X_i) \cdot \lambda_N F(X_i)\, F(dx)$$

$$+ \lambda_N \cdot \iint u_N(x,y) \cdot 1_{\{X_i \le y\}}\, F(dx)F(dy) - 2 \cdot \iint u_N(x,y) \cdot \lambda_N F(y)\, F(dx)F(dy) \Big]$$

$$= m^{-1} \cdot \sum_{i=1}^{m} \Big[\int u_N(X_i, y) \cdot \lambda_N F(y)\, F(dy) + \int u_N(x, X_i) \cdot \lambda_N F(X_i)\, F(dx)$$

$$- \iint u_N(x,y) \cdot \lambda_N F(y)\, F(dx)F(dy) \Big]$$

$$+ \lambda_N \cdot m^{-1} \cdot \sum_{i=1}^{m} \Big[\iint u_N(x,y) \cdot 1_{\{X_i \le y\}}\, F(dx)F(dy) - \iint u_N(x,y) F(y)\, F(dx)F(dy) \Big].$$

And lastly, for $(1 - \lambda_N) \cdot \hat{U}^3_{m,n}$ we have

$$(1 - \lambda_N) \cdot \hat{U}^3_{m,n}$$

$$= (1 - \lambda_N) \Big[m^{-1} \cdot \sum_{i=1}^{m} \Big[\iint u_N(X_i, y) \cdot 1_{\{z \le y\}}\, F(dy)G(dz) + \iint u_N(x, X_i) \cdot 1_{\{z \le X_i\}}\, F(dx)G(dz) \Big]$$

$$+ n^{-1} \cdot \sum_{k=1}^{n} \iint u_N(x,y) \cdot 1_{\{Y_k \le y\}}\, F(dx)F(dy) - 2 \iiint u_N(x,y) \cdot 1_{\{z \le y\}}\, F(dx)F(dy)G(dz) \Big]$$

$$= m^{-1} \cdot \sum_{i=1}^{m} \Big[\int u_N(X_i, y) \cdot (1 - \lambda_N) G(y)\, F(dy) + \int u_N(x, X_i) \cdot (1 - \lambda_N) G(X_i)\, F(dx)$$

$$- \iint u_N(x,y) \cdot (1 - \lambda_N) G(y)\, F(dx)F(dy) \Big]$$

$$+ (1 - \lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^{n} \Big[\iint u_N(x,y) \cdot 1_{\{Y_k \le y\}}\, F(dx)F(dy) - \iint u_N(x,y) G(y)\, F(dx)F(dy) \Big].$$

Now, since

$$H_N = \lambda_N \cdot F + (1 - \lambda_N) \cdot G,$$

we see that

$$\hat{U}^1_m - \lambda_N \cdot \hat{U}^2_m - (1 - \lambda_N) \cdot \hat{U}^3_{m,n}$$

$$= -\lambda_N \cdot m^{-1} \cdot \sum_{i=1}^{m} \Big[\iint u_N(x,y) \cdot 1_{\{X_i \le y\}}\, F(dx)F(dy) - \iint u_N(x,y) F(y)\, F(dx)F(dy) \Big]$$

$$- (1 - \lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^{n} \Big[\iint u_N(x,y) \cdot 1_{\{Y_k \le y\}}\, F(dx)F(dy) - \iint u_N(x,y) G(y)\, F(dx)F(dy) \Big]$$

$$= 0$$

due to

$$\int u_N(x,y)\, F(dx)$$

$$= \int \Big[a_N^{-2} \cdot \Big[K'\big(a_N^{-1}(H_N(x) - H_N(y))\big) - \int K'\big(a_N^{-1}(H_N(z) - H_N(y))\big)\, F(dz) \Big] \Big] F(dx)$$

$$
\begin{aligned}
&= a_N^{-2} \cdot \left[\int K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dx) - \int K'\big(a_N^{-1}(H_N(z) - H_N(y))\big)\, F(dz) \right] \\
&= 0.
\end{aligned}
$$

Thus, for (5.48) we have

$$
\begin{aligned}
&\frac{m-1}{m} \cdot U_m^1 - \frac{\lambda_N \cdot (m-1)(m-2)}{m^2} \cdot U_m^2 - \frac{(1-\lambda_N)\cdot(m-1)}{m} \cdot U_{m,n}^3 + O(a_N^{-2} \cdot N^{-1}) \\
&= \frac{m-1}{m} \cdot \left[U_m^1 - \lambda_N \cdot U_m^2 - \frac{2\lambda_N}{m} \cdot U_m^2 - (1-\lambda_N) \cdot U_{m,n}^3 \right] + O(a_N^{-2} \cdot N^{-1}) \\
&= \frac{m-1}{m} \cdot \left[U_m^1 - \hat{U}_m^1 - \lambda_N \cdot \left[U_m^2 - \hat{U}_m^2 \right] - \frac{2\lambda_N}{m} \cdot U_m^2 - (1-\lambda_N) \cdot \left[U_{m,n}^3 - \hat{U}_{m,n}^3 \right] \right] + O(a_N^{-2} \cdot N^{-1}),
\end{aligned}
$$

and it remains only to bound $\big[U_m^1 - \hat{U}_m^1\big]$, $\big[U_m^2 - \hat{U}_m^2\big]$ and $\big[U_{m,n}^3 - \hat{U}_{m,n}^3\big]$ and $\frac{2\lambda_N}{m} \cdot U_m^2$. Firstly, using Lemma A.2 we have

$$
\mathbb{E}\Big[U_m^1 - \hat{U}_m^1\Big]^2 \leq 2(m-1)m^{-3} \cdot \mathbb{E}\big[u_{1N}^*(X_1, X_2)\big]^2
$$

for u_{1N}^* defined as

$$
\begin{aligned}
u_{1N}^*(r,s) = u_N(r,s) \cdot H_N(s) &- \int u_N(r,y) \cdot H_N(y)\, F(dy) \\
&- \int u_N(x,s) \cdot H_N(s)\, F(dx) + \iint u_N(x,y) \cdot H_N(y)\, F(dx)F(dy)
\end{aligned}
$$

so that the expectation is easily bounded:

$$
\begin{aligned}
&\mathbb{E}\big[u_{1N}^*(X_1, X_2)\big]^2 \\
&\quad\leq 4 \cdot \mathbb{E}\Bigg[\big[u_N(X_1,X_2) \cdot H_N(X_2)\big]^2 + \left[\int u_N(X_1,y) \cdot H_N(y)\, F(dy) \right]^2 \\
&\qquad + \left[\int u_N(x,X_2) \cdot H_N(X_2)\, F(dx) \right]^2 + \left[\iint u_N(x,y) \cdot H_N(y)\, F(dx)F(dy) \right]^2 \Bigg] \\
&\quad\leq 4 \cdot \mathbb{E}\Bigg[\big[u_N(X_1,X_2)\big]^2 + \int \big[u_N(X_1,y)\big]^2\, F(dy) \\
&\qquad + \left[\int u_N(x,X_2)\, F(dx) \right]^2 + \left[\iint u_N(x,y)\, F(dx) \cdot H_N(y)\, F(dy) \right]^2 \Bigg] \\
&\quad\leq 4 \cdot \mathbb{E}\Bigg[\big[2a_N^{-2} \cdot \|K'\|\,\big]^2 + \int \big[2a_N^{-2} \cdot \|K'\|\,\big]^2\, F(dy) \Bigg] \\
&\quad\leq 32\, \|K'\|^2 \cdot a_N^{-4}.
\end{aligned}
$$

Altogether this yields

$$
\begin{aligned}
\mathbb{E}\Big[U_m^1 - \hat{U}_m^1\Big]^2 &\leq 2(m-1)m^{-3} \cdot \mathbb{E}\big[u_{1N}^*(X_1, X_2)\big]^2 \\
&\leq 2(m-1)m^{-3} \cdot 32\big\|K'\big\|^2 \cdot a_N^{-4} \\
&= 64\big\|K'\big\|^2 \cdot a_N^{-4} \cdot (m-1)m^{-3} \\
&= O(a_N^{-4} \cdot N^{-2}). \qquad (5.54)
\end{aligned}
$$

Similarly, using Lemmas A.4 and A.5 we have

$$\mathbb{E}\Big[U_m^2 - \hat{U}_m^2\Big]^2 = O(m^{-2}) \cdot \big\|u_{2N}^*\big\|^2$$

and

$$\mathbb{E}\Big[U_{m,n}^3 - \hat{U}_{m,n}^3\Big]^2 = \Big[O(m^{-1}n^{-1}) + O(m^{-2})\Big] \cdot \big\|u_{3N}^*\big\|^2$$

for u_{2N}^* and u_{3N}^* defined as

$$\begin{aligned} u_{2N}^*(r,s,t) &= u_N(r,s) \cdot 1_{\{t\le s\}} - \iint u_N(r,y)\cdot 1_{\{z\le y\}}\, F(dy)F(dz) - \iint u_N(x,s)\cdot 1_{\{z\le s\}}\, F(dx)F(dz) \\ &\quad - \iint u_N(x,y)\cdot 1_{\{t\le y\}}\, F(dx)F(dy) + 2\cdot \iiint u_N(x,y)\cdot 1_{\{z\le y\}}\, F(dx)F(dy)F(dz) \end{aligned}$$

and

$$\begin{aligned} u_{3N}^*(r,s,t) &= u_N(r,s) \cdot 1_{\{t\le s\}} - \iint u_N(r,y)\cdot 1_{\{z\le y\}}\, F(dy)G(dz) - \iint u_N(x,s)\cdot 1_{\{z\le s\}}\, F(dx)G(dz) \\ &\quad - \iint u_N(x,y)\cdot 1_{\{t\le y\}}\, F(dx)F(dy) + 2\cdot \iiint u_N(x,y)\cdot 1_{\{z\le y\}}\, F(dx)F(dy)G(dz). \end{aligned}$$

Bounding u_{2N}^* we obtain

$$\begin{aligned} &\big|v_{2N}^*(r,s,t)\big| \\ &= \Big|u_N(r,s) \cdot 1_{\{t\le s\}} - \iint u_N(r,y)\cdot 1_{\{z\le y\}}\, F(dy)F(dz) - \iint u_N(x,s)\cdot 1_{\{z\le s\}}\, F(dx)F(dz) \\ &\quad - \iint u_N(x,y)\cdot 1_{\{t\le y\}}\, F(dx)F(dy) + 2\cdot \iiint u_N(x,y)\cdot 1_{\{z\le y\}}\, F(dx)F(dy)F(dz)\Big| \\ &= \Big|u_N(r,s) \cdot 1_{\{t\le s\}} - \iint u_N(r,y)\cdot 1_{\{z\le y\}}\, F(dy)F(dz) - \iint u_N(x,s)\, F(dx)\cdot 1_{\{z\le s\}}\, F(dz) \\ &\quad - \iint u_N(x,y)\, F(dx)\cdot 1_{\{t\le y\}}\, F(dy) + 2\cdot \iiint u_N(x,y)\, F(dx)\cdot 1_{\{z\le y\}}\, F(dy)F(dz)\Big| \\ &\le \big|u_N(r,s)\big| + \iint \big|u_N(r,y)\big|\, F(dy)F(dz) \\ &\le 2a_N^{-2}\cdot \big\|K'\big\| + 2a_N^{-2}\cdot \big\|K'\big\| \\ &= 4a_N^{-2}\cdot \big\|K'\big\|. \end{aligned}$$

Completely analogous arguments show that

$$\big|v_{3N}^*(r,s,t)\big| \le 4a_N^{-2}\cdot \big\|K'\big\|$$

as well. This gives us

$$\mathbb{E}\Big[U_m^2 - \hat{U}_m^2\Big]^2 = O(m^{-2}) \cdot \big\|u_{2N}^*\big\|^2 = O(m^{-2})\cdot O(a_N^{-4}) = O(a_N^{-4}\cdot N^{-2}) \tag{5.55}$$

and

$$\mathbb{E}\Big[U_{m,n}^3 - \hat{U}_{m,n}^3\Big]^2 = \Big[O(m^{-2}) - O(m^{-1}n^{-1})\Big] \cdot \|u_{3N}^*\|^2 = O(a_N^{-4}\cdot N^{-2}). \tag{5.56}$$

Lastly,

$$\begin{aligned}
\frac{2\lambda_N}{m} \cdot U_m^2 &= \frac{2\lambda_N}{m} \cdot m^{-1}(m-1)^{-1}(m-2)^{-1} \cdot \sum_{\substack{1\le i,\, j,\, l\le m \\ i\ne j,\, j\ne l \text{ and } i\ne l}} u_N(X_i, X_j) \cdot 1_{\{X_l \le X_j\}} \\
&\le \frac{2\lambda_N}{m^2(m-1)(m-2)} \cdot \sum_{\substack{1\le i,\, j,\, l\le m \\ i\ne j,\, j\ne l \text{ and } i\ne l}} \|u_N\| \\
&= \frac{2\lambda_N \cdot \|u_N\|}{m} \\
&\le \frac{2\lambda_N \cdot 2\|K'\| a_N^{-2}}{m} \\
&= O(a_N^{-2} \cdot N^{-1}). \qquad (5.57)
\end{aligned}$$

Combining (5.54), (5.55), (5.56) and (5.57) we see that (5.48) is equal to

$$\begin{aligned}
&\frac{m-1}{m} \cdot \Big[O_P(a_N^{-2} \cdot N^{-1}) - \lambda_N \cdot O_P(a_N^{-2} \cdot N^{-1}) - O(a_N^{-2} \cdot N^{-1}) \\
&\quad - (1-\lambda_N) \cdot O_P(a_N^{-2} \cdot N^{-1}) \Big] + O(a_N^{-2} \cdot N^{-1}) = O_P(a_N^{-2} \cdot N^{-1})
\end{aligned}$$

which completes the proof. □

Lemma 5.13.

$$\begin{aligned}
&a_N^{-1} \cdot m^{-2} \cdot \sum_{1\le i\ne j\le m} \Bigg[\int_{a_N^{-1}(H_N(X_i)-H_N(X_j))}^{a_N^{-1}(H_N(X_i)-\hat H_N(X_j))} \big(a_N^{-1}(H_N(X_i) - \hat H_N(X_j)) - t\big) \cdot K''(t)\, dt \\
&\quad - \iint_{a_N^{-1}(H_N(x)-H_N(X_j))}^{a_N^{-1}(H_N(x)-\hat H_N(X_j))} \big(a_N^{-1}(H_N(x) - \hat H_N(X_j)) - t\big) \cdot K''(t)\, dt\, F(dx) \Bigg] \\
&= O_P(a_N^{-3} \cdot N^{-\frac{5}{4}}).
\end{aligned}$$

Proof. Begin by defining

$$\begin{aligned}
\hat u_N(r,s) = a_N^{-1} \cdot \Bigg[&\int_{a_N^{-1}(H_N(r)-H_N(s))}^{a_N^{-1}(H_N(r)-\hat H_N(s))} \big(a_N^{-1}(H_N(r) - \hat H_N(s)) - t\big) \cdot K''(t)\, dt \\
&- \iint_{a_N^{-1}(H_N(x)-H_N(s))}^{a_N^{-1}(H_N(x)-\hat H_N(s))} \big(a_N^{-1}(H_N(x) - \hat H_N(s)) - t\big) \cdot K''(t)\, dt\, F(dx) \Bigg].
\end{aligned}$$

Then we may write (5.49) as

$$m^{-2} \cdot \sum_{1\le i\ne j\le m} \hat u_N(X_i, X_j). \qquad (5.58)$$

Looking at the second moment of (5.58) in order to derive a bound for (5.49) we find

$$\begin{aligned}
&\mathbb{E}\Big[m^{-2} \cdot \sum_{1\le i\ne j\le m} \hat u_N(X_i, X_j) \Big]^2 \\
&= m^{-4} \cdot \Big[m(m-1) \cdot \mathbb{E}\big[\hat u_N(X_1, X_2)\big]^2 \qquad (5.59)
\end{aligned}$$

$$+ m(m-1) \cdot \mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_2, X_1)\big] \tag{5.60}$$

$$+ 2m(m-1)(m-2) \cdot \mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_3, X_1)\big] \tag{5.61}$$

$$+ m(m-1)(m-2) \cdot \mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_1, X_3)\big] \tag{5.62}$$

$$+ m(m-1)(m-2) \cdot \mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_3, X_2)\big] \tag{5.63}$$

$$+ m(m-1)(m-2)(m-3) \cdot \mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_3, X_4)\big]\Big]. \tag{5.64}$$

In order to derive bounds for some of the expectations in (5.59) through (5.64) we define $\hat{H}_N^*$ to be equal to $\hat{H}_N$ with X_1, X_2, X_3 and X_4 removed from the sample. That is,

$$\hat{H}_N^*(x) = N^{-1} \cdot \Bigg[\sum_{i=5}^{m} 1_{\{X_i \le x\}} + \sum_{k=1}^{n} 1_{\{Y_k \le x\}}\Bigg]. \tag{5.65}$$

Using $\hat{H}_N^*$ we can derive a useful decomposition of $\hat{u}_N$. Define $\hat{u}_{1N}^*$, $\hat{u}_{2N}^*$ and $\hat{u}_{3N}^*$ as

$$\hat{u}_{1N}^*(r,s) = a_N^{-1} \cdot \Bigg[\int_{a_N^{-1}(H_N(r)-H_N(s))}^{a_N^{-1}(H_N(r)-\hat{H}_N^*(s))} \big(a_N^{-1}(H_N(r) - \hat{H}_N^*(s)) - t\big) \cdot K''(t)\, dt$$
$$- \iint_{a_N^{-1}(H_N(x)-H_N(s))}^{a_N^{-1}(H_N(x)-\hat{H}_N^*(s))} \big(a_N^{-1}(H_N(x) - \hat{H}_N^*(s)) - t\big) \cdot K''(t)\, dt\, F(dx)\Bigg], \tag{5.66}$$

$$\hat{u}_{2N}^*(r,s) = a_N^{-1} \cdot \Bigg[\int_{a_N^{-1}(H_N(r)-H_N(s))}^{a_N^{-1}(H_N(r)-\hat{H}_N^*(s))} a_N^{-1}(\hat{H}_N^*(s) - \hat{H}_N(s)) \cdot K''(t)\, dt$$
$$- \iint_{a_N^{-1}(H_N(x)-H_N(s))}^{a_N^{-1}(H_N(x)-\hat{H}_N^*(s))} a_N^{-1}(\hat{H}_N^*(s) - \hat{H}_N(s)) \cdot K''(t)\, dt\, F(dx)\Bigg] \tag{5.67}$$

and

$$\hat{u}_{3N}^*(r,s) = a_N^{-1} \cdot \Bigg[\int_{a_N^{-1}(H_N(r)-\hat{H}_N^*(s))}^{a_N^{-1}(H_N(r)-\hat{H}_N(s))} \big(a_N^{-1}(H_N(r) - \hat{H}_N(s)) - t\big) \cdot K''(t)\, dt$$
$$- \iint_{a_N^{-1}(H_N(x)-\hat{H}_N^*(s))}^{a_N^{-1}(H_N(x)-\hat{H}_N(s))} a_N^{-1}\big(a_N^{-1}(H_N(x) - \hat{H}_N(s)) - t\big) \cdot K''(t)\, dt\, F(dx)\Bigg]. \tag{5.68}$$

Then we may write $\hat{u}_N$ as

$$\hat{u}_N = \hat{u}_{1N}^* + \hat{u}_{2N}^* + \hat{u}_{3N}^*,$$

and for the expectation in (5.64) we may write

$$\begin{aligned}
&\mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_3, X_4)\big] \\
&\quad = \mathbb{E}\big[\big[\hat{u}_{1N}^*(X_1, X_2) + \hat{u}_{2N}^*(X_1, X_2) + \hat{u}_{3N}^*(X_1, X_2)\big] \cdot \big[\hat{u}_{1N}^*(X_3, X_4) + \hat{u}_{2N}^*(X_3, X_4) + \hat{u}_{3N}^*(X_3, X_4)\big]\big] \\
&\quad = \mathbb{E}\big[\hat{u}_{1N}^*(X_1, X_2) \cdot \hat{u}_{1N}^*(X_3, X_4)\big] + \mathbb{E}\big[\hat{u}_{1N}^*(X_1, X_2) \cdot \hat{u}_{2N}^*(X_3, X_4)\big] + \mathbb{E}\big[\hat{u}_{1N}^*(X_1, X_2) \cdot \hat{u}_{3N}^*(X_3, X_4)\big] \\
&\qquad + \mathbb{E}\big[\hat{u}_{2N}^*(X_1, X_2) \cdot \hat{u}_{1N}^*(X_3, X_4)\big] + \mathbb{E}\big[\hat{u}_{2N}^*(X_1, X_2) \cdot \hat{u}_{2N}^*(X_3, X_4)\big] + \mathbb{E}\big[\hat{u}_{2N}^*(X_1, X_2) \cdot \hat{u}_{3N}^*(X_3, X_4)\big] \\
&\qquad + \mathbb{E}\big[\hat{u}_{3N}^*(X_1, X_2) \cdot \hat{u}_{1N}^*(X_3, X_4)\big] + \mathbb{E}\big[\hat{u}_{3N}^*(X_1, X_2) \cdot \hat{u}_{2N}^*(X_3, X_4)\big] + \mathbb{E}\big[\hat{u}_{3N}^*(X_1, X_2) \cdot \hat{u}_{3N}^*(X_3, X_4)\big] \\
&\quad \le \mathbb{E}\big[\hat{u}_{1N}^*(X_1, X_2) \cdot \hat{u}_{1N}^*(X_3, X_4)\big] \\
&\qquad + \Big[\mathbb{E}\big[\hat{u}_{1N}^*(X_1, X_2)\big]^2\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\big[\hat{u}_{2N}^*(X_3, X_4)\big]^2\Big]^{\frac{1}{2}} + \Big[\mathbb{E}\big[\hat{u}_{1N}^*(X_1, X_2)\big]^2\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\big[\hat{u}_{3N}^*(X_3, X_4)\big]^2\Big]^{\frac{1}{2}}
\end{aligned}$$

$$+ \Big[\mathbb{E}\big[\hat{u}^*_{2N}(X_1,X_2)\big]^2\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\big[\hat{u}^*_{1N}(X_3,X_4)\big]^2\Big]^{\frac{1}{2}} + \Big[\mathbb{E}\big[\hat{u}^*_{2N}(X_1,X_2)\big]^2\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\big[\hat{u}^*_{2N}(X_3,X_4)\big]^2\Big]^{\frac{1}{2}}$$

$$+ \Big[\mathbb{E}\big[\hat{u}^*_{2N}(X_1,X_2)\big]^2\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\big[\hat{u}^*_{3N}(X_3,X_4)\big]^2\Big]^{\frac{1}{2}} + \Big[\mathbb{E}\big[\hat{u}^*_{3N}(X_1,X_2)\big]^2\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\big[\hat{u}^*_{1N}(X_3,X_4)\big]^2\Big]^{\frac{1}{2}}$$

$$+ \Big[\mathbb{E}\big[\hat{u}^*_{3N}(X_1,X_2)\big]^2\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\big[\hat{u}^*_{2N}(X_3,X_4)\big]^2\Big]^{\frac{1}{2}} + \Big[\mathbb{E}\big[\hat{u}^*_{3N}(X_1,X_2)\big]^2\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\big[\hat{u}^*_{3N}(X_3,X_4)\big]^2\Big]^{\frac{1}{2}}$$

$$= \mathbb{E}\big[\hat{u}^*_{1N}(X_1,X_2) \cdot \hat{u}^*_{1N}(X_3,X_4)\big] + 2 \cdot \Big[\mathbb{E}\big[\hat{u}^*_{1N}(X_1,X_2)\big]^2\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\big[\hat{u}^*_{2N}(X_1,X_2)\big]^2\Big]^{\frac{1}{2}}$$

$$+ 2 \cdot \Big[\mathbb{E}\big[\hat{u}^*_{1N}(X_1,X_2)\big]^2\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\big[\hat{u}^*_{3N}(X_1,X_2)\big]^2\Big]^{\frac{1}{2}} + \mathbb{E}\big[\hat{u}^*_{2N}(X_1,X_2)\big]^2$$

$$+ 2 \cdot \Big[\mathbb{E}\big[\hat{u}^*_{2N}(X_1,X_2)\big]^2\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\big[\hat{u}^*_{3N}(X_1,X_2)\big]^2\Big]^{\frac{1}{2}} + \mathbb{E}\big[\hat{u}^*_{3N}(X_1,X_2)\big]^2.$$

This means we only need to bound the four expectations $\mathbb{E}\big[\hat{u}^*_{1N}(X_1,X_2)\cdot\hat{u}^*_{1N}(X_3,X_4)\big]$, $\mathbb{E}\big[\hat{u}^*_{1N}(X_1,X_2)\big]^2$, $\mathbb{E}\big[\hat{u}^*_{2N}(X_1,X_2)\big]^2$ and $\mathbb{E}\big[\hat{u}^*_{3N}(X_1,X_2)\big]^2$ in order to bound the expectation in (5.64). Firstly,

$$\begin{aligned}
&\mathbb{E}\big[\hat{u}^*_{1N}(X_1,X_2) \cdot \hat{u}^*_{1N}(X_3,X_4)\big] \\
&\quad = \mathbb{E}\Big[\mathbb{E}\big[\hat{u}^*_{1N}(X_1,X_2) \mid X_2, X_3, \cdots, X_m,\ Y_1, Y_2, \cdots, Y_n\big] \cdot \hat{u}^*_{1N}(X_3,X_4)\Big] \\
&\quad = 0,
\end{aligned}$$

since for the inner expectation

$$\begin{aligned}
&\mathbb{E}\big[\hat{u}^*_{1N}(X_1,X_2) \mid X_2, X_3, \cdots, X_m,\ Y_1, Y_2, \cdots, Y_n\big] \\
&= a_N^{-1} \cdot \Bigg[\iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}^*_N(X_2))} \big(a_N^{-1}(H_N(x) - \hat{H}^*_N(X_2)) - t\big) \cdot K''(t)\, dt\, F(dx) \\
&\qquad - \iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}^*_N(X_2))} \big(a_N^{-1}(H_N(x) - \hat{H}^*_N(X_2)) - t\big) \cdot K''(t)\, dt\, F(dx)\Bigg] \\
&= 0,
\end{aligned}$$

so that the first expectation vanishes completely.

The other three expectations do not vanish, but can be bound adequately. Beginning with $\mathbb{E}\big[\hat{u}^*_{1N}(X_1,X_2)\big]^2$ we have

$$\begin{aligned}
&\mathbb{E}\big[\hat{u}^*_{1N}(X_1,X_2)\big]^2 \\
&= \mathbb{E}\Bigg[a_N^{-1} \cdot \Bigg[\int_{a_N^{-1}(H_N(X_1)-H_N(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}^*_N(X_2))} \big(a_N^{-1}(H_N(X_1) - \hat{H}^*_N(X_2)) - t\big) \cdot K''(t)\, dt \\
&\qquad - \iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}^*_N(X_2))} \big(a_N^{-1}(H_N(x) - \hat{H}^*_N(X_2)) - t\big) \cdot K''(t)\, dt\, F(dx)\Bigg]\Bigg]^2 \\
&\le 2a_N^{-2} \cdot \mathbb{E}\Bigg[\Bigg[\int_{a_N^{-1}(H_N(X_1)-H_N(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}^*_N(X_2))} \big(a_N^{-1}(H_N(X_1) - \hat{H}^*_N(X_2)) - t\big) \cdot K''(t)\, dt\Bigg]^2 \\
&\qquad + \Bigg[\iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}^*_N(X_2))} \big(a_N^{-1}(H_N(x) - \hat{H}^*_N(X_2)) - t\big) \cdot K''(t)\, dt\, F(dx)\Bigg]^2\Bigg] \\
&\le 2a_N^{-2} \cdot \mathbb{E}\Bigg[\Bigg[\|K''\| \int_{a_N^{-1}(H_N(X_1)-H_N(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}^*_N(X_2))} \big|a_N^{-1}(H_N(X_1) - \hat{H}^*_N(X_2)) - t\big|\, dt\Bigg]^2
\end{aligned}$$

$$+ \left[\|K''\| \iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N^*(X_2))} \left| a_N^{-1}(H_N(x) - \hat{H}_N^*(X_2)) - t \right| dt\, F(dx) \right]^2 \Bigg]$$

$$\leq 2a_N^{-2} \cdot \mathbb{E}\Bigg[\Bigg[\|K''\| \cdot \left| a_N^{-1}(H_N(X_2) - \hat{H}_N^*(X_2)) \right| \cdot \int_{a_N^{-1}(H_N(X_1)-H_N(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}_N^*(X_2))} dt \Bigg]^2$$

$$+ \left[\|K''\| \cdot \left| a_N^{-1}(H_N(X_2) - \hat{H}_N^*(X_2)) \right| \cdot \iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N^*(X_2))} dt\, F(dx) \right]^2 \Bigg]$$

$$\leq 2a_N^{-2} \cdot \mathbb{E}\Bigg[\Big[\|K''\| \cdot \left| a_N^{-1}(H_N(X_2) - \hat{H}_N^*(X_2)) \right|^2 \Big]^2$$

$$+ \Big[\|K''\| \cdot \left| a_N^{-1}(H_N(X_2) - \hat{H}_N^*(X_2)) \right|^2 \Big]^2 \Bigg]$$

$$\leq 2a_N^{-2} \cdot \mathbb{E}\Big[2 \Big[\|K''\| \cdot \left| a_N^{-1}(H_N(X_2) - \hat{H}_N^*(X_2)) \right|^2 \Big]^2 \Big]$$

$$\leq 4a_N^{-6} \cdot \|K''\|^2 \cdot \mathbb{E}\Big[\Big[\left| (H_N(X_2) - \hat{H}_N^*(X_2)) \right|^2 \Big]^2 \Big]$$

$$\leq 4a_N^{-6} \cdot \|K''\|^2 \cdot \mathbb{E}\Big[\|H_N - \hat{H}_N^*\|^4 \Big].$$

The expectation is $O(N^{-2})$ by the well-known D-K-W bound on $\|H_N - \hat{H}_N\|$, so that altogether we have

$$\mathbb{E}\big[\hat{u}_{1N}^*(X_1, X_2)\big]^2 = O(a_N^{-6} \cdot N^{-2}). \tag{5.69}$$

For the next expectation $\mathbb{E}\big[\hat{u}_{2N}^*(X_1, X_2)\big]^2$ we obtain

$$\mathbb{E}\big[\hat{u}_{2N}^*(X_1, X_2)\big]^2$$

$$= \mathbb{E}\Bigg[a_N^{-1} \cdot \Bigg[\int_{a_N^{-1}(H_N(X_1)-H_N(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}_N^*(X_2))} a_N^{-1}(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)) \cdot K''(t)\, dt$$

$$- \iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N^*(X_2))} a_N^{-1}(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)) \cdot K''(t)\, dt\, F(dx) \Bigg] \Bigg]^2$$

$$\leq 2a_N^{-2} \cdot \mathbb{E}\Bigg[\Bigg[\int_{a_N^{-1}(H_N(X_1)-H_N(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}_N^*(X_2))} a_N^{-1}(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)) \cdot K''(t)\, dt \Bigg]^2$$

$$+ \left[\iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N^*(X_2))} a_N^{-1}(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)) \cdot K''(t)\, dt\, F(dx) \right]^2 \Bigg]$$

$$\leq 2a_N^{-2} \cdot \mathbb{E}\Bigg[\Bigg[\|K''\| \int_{a_N^{-1}(H_N(X_1)-H_N(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}_N^*(X_2))} \left| a_N^{-1}(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)) \right| dt \Bigg]^2$$

$$+ \left[\|K''\| \iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N^*(X_2))} \left| a_N^{-1}(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)) \right| dt\, F(dx) \right]^2 \Bigg]$$

$$\leq 2a_N^{-4} \cdot \|K''\|^2 \cdot \mathbb{E}\Bigg[\Bigg[\int_{a_N^{-1}(H_N(X_1)-H_N(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}_N^*(X_2))} \left| \hat{H}_N^*(X_2) - \hat{H}_N(X_2) \right| dt \Bigg]^2$$

$$+ \left[\iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N^*(X_2))} \left| \hat{H}_N^*(X_2) - \hat{H}_N(X_2) \right| dt\, F(dx) \right]^2 \Bigg]$$

$$
\begin{aligned}
&\le 4a_N^{-6} \cdot \|K''\|^2 \cdot \mathbb{E}\Big[\big|\hat{H}_N^*(X_2) - \hat{H}_N(X_2)\big| \cdot \big|\hat{H}_N^*(X_2) - H_N(X_2)\big|\Big]^2 \\
&\le 4a_N^{-6} \cdot \|K''\|^2 \cdot \mathbb{E}\Big[4N^{-1} \cdot \|\hat{H}_N^* - H_N\|\Big]^2 \\
&\le 64a_N^{-6} \cdot N^{-2} \cdot \|K''\|^2 \cdot \mathbb{E}\Big[\|\hat{H}_N^* - H_N\|^2\Big].
\end{aligned}
$$

The expectation is $O(N^{-1})$ by the D-K-W bound, so that altogether we have

$$
\mathbb{E}\big[\hat{u}_{2N}^*(X_1, X_2)\big]^2 = O(a_N^{-6} \cdot N^{-3}). \tag{5.70}
$$

Lastly, for the expectation $\mathbb{E}\big[\hat{u}_{3N}^*(X_1, X_2)\big]^2$ we obtain

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_{3N}^*(X_1, X_2)\big]^2 \\
&= \mathbb{E}\Big[a_N^{-1} \cdot \Big[\int_{a_N^{-1}(H_N(X_1)-\hat{H}_N^*(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}_N(X_2))} \big(a_N^{-1}(H_N(X_1) - \hat{H}_N(X_2)) - t\big) \cdot K''(t)\, dt \\
&\quad - \iint_{a_N^{-1}(H_N(x)-\hat{H}_N^*(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N(X_2))} a_N^{-1}\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_2)) - t\big) \cdot K''(t)\, dt\, F(dx)\Big]\Big]^2 \\
&\le 2a_N^{-2} \cdot \mathbb{E}\Big[\Big[\int_{a_N^{-1}(H_N(X_1)-\hat{H}_N^*(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}_N(X_2))} \big(a_N^{-1}(H_N(X_1) - \hat{H}_N(X_2)) - t\big) \cdot K''(t)\, dt\Big]^2 \\
&\quad + \Big[\iint_{a_N^{-1}(H_N(x)-\hat{H}_N^*(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N(X_2))} a_N^{-1}\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_2)) - t\big) \cdot K''(t)\, dt\, F(dx)\Big]^2\Big] \\
&\le 2a_N^{-2} \cdot \mathbb{E}\Big[\Big[\|K''\| \int_{a_N^{-1}(H_N(X_1)-\hat{H}_N^*(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}_N(X_2))} \big|a_N^{-1}(H_N(X_1) - \hat{H}_N(X_2)) - t\big|\, dt\Big]^2 \\
&\quad + \Big[\|K''\| \iint_{a_N^{-1}(H_N(x)-\hat{H}_N^*(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N(X_2))} \big|a_N^{-1}(H_N(x) - \hat{H}_N(X_2)) - t\big|\, dt\, F(dx)\Big]^2\Big] \\
&\le 2a_N^{-2} \cdot \mathbb{E}\Big[\Big[\|K''\| \cdot \big|a_N^{-1}(\hat{H}_N^*(X_2) - \hat{H}_N(X_2))\big| \int_{a_N^{-1}(H_N(X_1)-\hat{H}_N^*(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}_N(X_2))} dt\Big]^2 \\
&\quad + \Big[\|K''\| \cdot \big|a_N^{-1}(\hat{H}_N^*(X_2) - \hat{H}_N(X_2))\big| \iint_{a_N^{-1}(H_N(x)-\hat{H}_N^*(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N(X_2))} dt\, F(dx)\Big]^2\Big] \\
&\le 2a_N^{-2} \cdot \|K''\|^2 \cdot \mathbb{E}\Big[\Big[\big|a_N^{-1}(\hat{H}_N^*(X_2) - \hat{H}_N(X_2))\big|^2\Big]^2 + \Big[\big|a_N^{-1}(\hat{H}_N^*(X_2) - \hat{H}_N(X_2))\big|^2\Big]^2\Big] \\
&\le 4a_N^{-2} \cdot \|K''\|^2 \cdot \mathbb{E}\Big[\Big[\big|a_N^{-1}(\hat{H}_N^*(X_2) - \hat{H}_N(X_2))\big|^2\Big]^2\Big] \\
&= 4a_N^{-6} \cdot \|K''\|^2 \cdot \mathbb{E}\Big[\big|\hat{H}_N^*(X_2) - \hat{H}_N(X_2)\big|^4\Big] \\
&\le 4a_N^{-6} \cdot \|K''\|^2 \cdot \mathbb{E}\Big[(4N^{-1})^4\Big] \\
&= 4^5 \|K''\|^2 \cdot a_N^{-6} \cdot N^{-4}.
\end{aligned}
$$

Thus, for the expectation we have

$$\mathbb{E}\big[\hat{u}_{3N}^{*}(X_1,X_2)\big]^2 = O(a_N^{-6}\cdot N^{-4}). \tag{5.71}$$

Combining (5.69), (5.70) and (5.71), we have shown for the expectation in (5.64) that

$$\begin{aligned}
&\mathbb{E}\big[\hat{u}_N(X_1,X_2)\cdot \hat{u}_N(X_3,X_4)\big]\\
&\quad = 2\cdot\Big[O(a_N^{-6}\cdot N^{-2}\Big]^{\frac{1}{2}}\cdot\Big[O(a_N^{-6}\cdot N^{-3})\Big]^{\frac{1}{2}}\\
&\qquad + 2\cdot\Big[O(a_N^{-6}\cdot N^{-2}\Big]^{\frac{1}{2}}\cdot\Big[O(a_N^{-6}\cdot N^{-4}\Big]^{\frac{1}{2}} + O(a_N^{-6}\cdot N^{-3})\\
&\qquad + 2\cdot\Big[O(a_N^{-6}\cdot N^{-3}\Big]^{\frac{1}{2}}\cdot\Big[O(a_N^{-6}\cdot N^{-4}\Big]^{\frac{1}{2}} + O(a_N^{-6}\cdot N^{-4})\\
&\quad = O(a_N^{-6}\cdot N^{-\frac{5}{2}}) + O(a_N^{-6}\cdot N^{-3}) + O(a_N^{-6}\cdot N^{-\frac{7}{2}}) + O(a_N^{-6}\cdot N^{-4})\\
&\quad = O(a_N^{-6}\cdot N^{-\frac{5}{2}}).
\end{aligned}$$

Using the Cauchy-Schwarz inequality, all of the other expectations in (5.59) through (5.63) are bound by the expectation $\mathbb{E}\big[\hat{u}_N(X_1,X_2)\big]^2$. Bounding this expression we have

$$\begin{aligned}
&\mathbb{E}\big[\hat{u}_N(X_1,X_2)\big]^2\\
&= \mathbb{E}\bigg[a_N^{-1}\cdot\bigg[\int_{a_N^{-1}(H_N(X_1)-H_N(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}_N(X_2))}\big(a_N^{-1}(H_N(X_1)-\hat{H}_N(X_2))-t\big)\cdot K''(t)\,dt\\
&\quad - \iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N(X_2))}\big(a_N^{-1}(H_N(x)-\hat{H}_N(X_2))-t\big)\cdot K''(t)\,dt\,F(dx)\bigg]\bigg]^2\\
&\le 2a_N^{-2}\cdot\mathbb{E}\bigg[\bigg[\int_{a_N^{-1}(H_N(X_1)-H_N(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}_N(X_2))}\big(a_N^{-1}(H_N(X_1)-\hat{H}_N(X_2))-t\big)\cdot K''(t)\,dt\bigg]^2\\
&\quad + \bigg[\iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N(X_2))}\big(a_N^{-1}(H_N(x)-\hat{H}_N(X_2))-t\big)\cdot K''(t)\,dt\,F(dx)\bigg]^2\bigg]\\
&\le 2a_N^{-2}\cdot\mathbb{E}\bigg[\bigg[\big\|K''\big\|\int_{a_N^{-1}(H_N(X_1)-H_N(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}_N(X_2))}\big|a_N^{-1}(H_N(X_1)-\hat{H}_N(X_2))-t\big|\,dt\bigg]^2\\
&\quad + \bigg[\big\|K''\big\|\iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N(X_2))}\big|a_N^{-1}(H_N(x)-\hat{H}_N(X_2))-t\big|\,dt\,F(dx)\bigg]^2\bigg]\\
&\le 2a_N^{-2}\cdot\mathbb{E}\bigg[\bigg[\big\|K''\big\|\cdot\big|a_N^{-1}(H_N(X_2)-\hat{H}_N(X_2))\big|\int_{a_N^{-1}(H_N(X_1)-H_N(X_2))}^{a_N^{-1}(H_N(X_1)-\hat{H}_N(X_2))}dt\bigg]^2\\
&\quad + \bigg[\big\|K''\big\|\cdot\big|a_N^{-1}(H_N(X_2)-\hat{H}_N(X_2))\big|\iint_{a_N^{-1}(H_N(x)-H_N(X_2))}^{a_N^{-1}(H_N(x)-\hat{H}_N(X_2))}dt\,F(dx)\bigg]^2\bigg]\\
&= 2a_N^{-2}\cdot\big\|K''\big\|^2\cdot\mathbb{E}\Big[2\big|a_N^{-1}(H_N(X_2)-\hat{H}_N(X_2))\big|^4\Big]\\
&\le 4a_N^{-6}\cdot\big\|K''\big\|^2\cdot\mathbb{E}\Big[\big\|H_N-\hat{H}_N\big\|^4\Big].
\end{aligned}$$

The expectation is $O(N^{-2})$ by the D-K-W bound, so that altogether we have

$$\mathbb{E}\big[\hat{u}_N(X_1, X_2)\big]^2 = O(a_N^{-6} \cdot N^{-2}), \tag{5.72}$$

which means that the summands (5.59) through (5.63) are all $O(a_N^{-6} \cdot N^{-3})$. Combining this with the fact that (5.64) is $O(a_N^{-6} \cdot N^{-\frac{5}{2}})$ gives us a rate for (5.58), namely

$$m^{-2} \cdot \sum_{1 \le i \ne j \le m} \hat{u}_N(X_i, X_j) = O_P(a_N^{-3} \cdot N^{-\frac{5}{4}}) \tag{5.73}$$

which completes the proof. □

To bound (5.43) we will use very similar arguments to those which we used to show that (5.42) is $O_P(a_N^{-2} \cdot N^{-1})$. We begin by deriving a sum representation of (5.43).

$$\begin{aligned}
&\int \Big[\hat{g}_N - \bar{g}_N\Big] \circ H_N(x) \Big[\hat{F}_m(dx) - F(dx)\Big] \\
&= m^{-1} \cdot \sum_{i=1}^{m} \bigg[\Big[\hat{g}_N - \bar{g}_N\Big] \circ H_N(X_i) - \int \Big[\hat{g}_N - \bar{g}_N\Big] \circ H_N(x)\, F(dx)\bigg] \\
&= m^{-1} \cdot \sum_{i=1}^{m} \bigg[(n \cdot a_N)^{-1} \cdot \sum_{k=1}^{n} K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(Y_k))\big) - \bar{g}_N \circ H_N(X_i) \\
&\qquad - \int (n \cdot a_N)^{-1} \cdot \sum_{k=1}^{n} K\big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))\big)\, F(dx) + \int \bar{g}_N \circ H_N(x)\, F(dx)\bigg] \\
&= m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} \bigg[a_N^{-1} K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(Y_k))\big) - \bar{g}_N \circ H_N(X_i) \\
&\qquad - a_N^{-1} \int K\big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))\big)\, F(dx) + \int \bar{g}_N \circ H_N(x)\, F(dx)\bigg] \\
&= m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} \bigg[a_N^{-1} K\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(Y_k))\big) - a_N^{-1} \int K\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, G(dy) \\
&\qquad - a_N^{-1} \int K\big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))\big)\, F(dx) + a_N^{-1} \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\, F(dx)\bigg]
\end{aligned}$$

Now using the Taylor expansion of the kernel function K about each of the $a_N^{-1}(H_N(X_i) - H_N(Y_k))$ and $a_N^{-1}(H_N(x) - H_N(Y_k))$ then yields

$$\begin{aligned}
&\int \Big[\hat{g}_N - \bar{g}_N\Big] \circ H_N(x) \Big[\hat{F}_m(dx) - F(dx)\Big] \\
&= a_N^{-1} \cdot m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} \bigg[K\big(a_N^{-1}(H_N(X_i) - H_N(Y_k))\big) - \int K\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, G(dy) \\
&\qquad - \int K\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) + \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\, F(dx)\bigg] \qquad (5.74) \\
&\quad + a_N^{-2} \cdot m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} \Big[K'\big(a_N^{-1}(H_N(X_i) - H_N(Y_k))\big)\big(H_N(Y_k) - \hat{H}_N(Y_k)\big)
\end{aligned}$$

$$- \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx)\big(H_N(X_j) - \hat{H}_N(Y_k)\big)\Big] \tag{5.75}$$

$$+ a_N^{-1} \cdot m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \Big[\int_{a_N^{-1}(H_N(X_i)-H_N(Y_k))}^{a_N^{-1}(H_N(X_i)-\hat{H}_N(Y_k))} \big(a_N^{-1}(H_N(X_i) - \hat{H}_N(Y_k)) - t\big) \cdot K''(t)\, dt$$

$$- \iint_{a_N^{-1}(H_N(x)-H_N(Y_k))}^{a_N^{-1}(H_N(x)-\hat{H}_N(Y_k))} \big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k)) - t\big) \cdot K''(t)\, dt\, F(dx)\Big] \tag{5.76}$$

Lemma 5.14.

$$a_N^{-1} \cdot m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \Big[K\big(a_N^{-1}(H_N(X_i) - H_N(Y_k))\big) - \int K\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, G(dy)$$
$$- \int K\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx) + \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\, F(dx)\Big]$$
$$= O_P(a_N^{-1} \cdot N^{-1}).$$

Proof. Define

$$u_N(s,\, t) = a_N^{-1} K\big(a_N^{-1}(H_N(s) - H_N(t))\big),$$

and define the generalized U-statistic $U_{m,n}$ as

$$U_{m,n} = m^{-1}n^{-1} \cdot \sum_{\substack{1\le i\le m\\ 1\le k\le n}} u_N(X_i, Y_k),$$

and let $\hat{U}_{m,n}$ be the Hájek projection of $U_{U,n}$ as defined in Lemma A.3. Then (5.74) is equal to $U_{m,n} - \hat{U}_{m,n}$.

Applying the inequality in Lemma A.3 we have

$$\mathbb{E}\Big[U_{m,n} - \hat{U}_{m,n}\Big]^2 = m^{-1}n^{-1} \cdot \mathbb{E}\big[u_N^*(X_1, Y_1)\big]^2$$

with u_N^* defined as

$$u_N^*(r,s) = u_N(r,s) - \int u_N(r,y)\, G(dy) - \int u_N(x,s)\, F(dx) + \iint u_N(x,y)\, F(dx)G(dy).$$

Thus it remains to bound the expectation $\mathbb{E}\big[u_N^*(X_1, Y_1)\big]^2$:

$$\mathbb{E}\big[u_N^*(X_1, Y_1)\big]^2$$
$$= \mathbb{E}\Big[a_N^{-1}\Big[K\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) - \int K\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, G(dy)$$
$$- \int K\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)\, F(dx) + \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\, F(dx)\Big]\Big]^2$$
$$= a_N^{-2} \cdot \mathbb{E}\Big[K\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) - \int K\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, G(dy)$$
$$- \int K\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)\, F(dx) + \iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\, F(dx)\Big]^2$$

$$\le a_N^{-2} \cdot 4 \cdot \mathbb{E}\Bigg[\big[K\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)\big]^2 + \Big[\int K\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, G(dy)\Big]^2$$

$$+ \Big[\int K\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)\, F(dx)\Big]^2 + \Big[\iint K\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\, F(dx)\Big]^2\Bigg]$$

$$\le a_N^{-2} \cdot 4 \cdot \mathbb{E}\big[4 \cdot \big\|K\big\|^2\big]$$

$$\le 16\big\|K\big\|^2 \cdot a_N^{-2}.$$

This yields

$$\begin{aligned}\mathbb{E}\Big[U_{m,n} - \hat{U}_{m,n}\Big]^2 &= m^{-1}n^{-1} \cdot \mathbb{E}\big[u_N^*(X_1, Y_1)\big]^2 \\ &\le m^{-1}n^{-1} \cdot 16\big\|K\big\|^2 \cdot a_N^{-2} \\ &= 16\big\|K\big\|^2 \cdot a_N^{-2} \cdot m^{-1}n^{-1} \\ &= O(a_N^{-2} \cdot N^{-2})\end{aligned}$$

which completes the proof. □

LEMMA 5.15.

$$a_N^{-2} \cdot m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n}\Big[K'\big(a_N^{-1}(H_N(X_i) - H_N(Y_k))\big)\big(H_N(Y_k) - \hat{H}_N(Y_k)\big)$$

$$- \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx)\big(H_N(Y_k) - \hat{H}_N(Y_k)\big)\Big]$$

$$= O_P(a_N^{-2} \cdot N^{-1}).$$

PROOF. Define

$$u_N(s,\, t) = a_N^{-2} \cdot \Big[K'\big(a_N^{-1}(H_N(s) - H_N(t))\big) - \int K'\big(a_N^{-1}(H_N(x) - H_N(t))\big)\, F(dx)\Big].$$

Then

$$a_N^{-2} \cdot m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n}\Big[K'\big(a_N^{-1}(H_N(X_i) - H_N(Y_k))\big)\big(H_N(Y_k) - \hat{H}_N(Y_k)\big)$$

$$- \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big)\, F(dx)\big(H_N(Y_k) - \hat{H}_N(Y_k)\big)\Big]$$

$$= m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} u_N(X_i, Y_k) \cdot \big[H_N(Y_k) - \hat{H}_N(Y_k)\big]$$

$$= m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} u_N(X_i, Y_k) \cdot \Big[H_N(Y_k) - N^{-1}\Big[\sum_{j=1}^{m} 1_{\{X_j \le Y_k\}} + \sum_{q=1}^{n} 1_{\{Y_q \le Y_k\}}\Big]\Big]$$

$$= m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} u_N(X_i, Y_k) \cdot H_N(Y_k) - m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} u_N(X_i, Y_k) \cdot N^{-1}\sum_{j=1}^{m} 1_{\{X_j \le Y_k\}}$$

$$- m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} u_N(X_i, Y_k) \cdot N^{-1}\sum_{q=1}^{n} 1_{\{Y_q \le Y_k\}}$$

$$= m^{-1}n^{-1} \cdot \sum_{\substack{1\le i\le m\\1\le k\le n}} u_N(X_i, Y_k) \cdot H_N(Y_k) - \lambda_N \cdot m^{-2}n^{-1} \cdot \sum_{\substack{1\le i\ne j\le m\\1\le k\le n}} u_N(X_i, Y_k) \cdot 1_{\{X_j \le Y_k\}}$$
$$- \lambda_N \cdot m^{-2}n^{-1} \cdot \sum_{\substack{1\le i\le m\\1\le k\le n}} u_N(X_i, Y_k) \cdot 1_{\{X_i \le Y_k\}} - (1-\lambda_N) \cdot m^{-1}n^{-2} \cdot \sum_{\substack{1\le i\le m\\1\le k\le n}} u_N(X_i, Y_k)$$
$$- (1-\lambda_N) \cdot m^{-1}n^{-2} \cdot \sum_{\substack{1\le i\le m\\1\le k\ne q\le n}} u_N(X_i, Y_k) \cdot 1_{\{Y_q \le Y_k\}}$$

Define the generalized U-statistics $U^1_{m,n}$, $U^2_{m,n}$, and $U^3_{m,n}$ as

$$U^1_{m,n} = m^{-1}n^{-1} \cdot \sum_{\substack{1\le i\le m\\1\le k\le n}} u_N(X_i, Y_k) \cdot H_N(Y_k),$$
$$U^2_{m,n} = m^{-1}(m-1)^{-1}n^{-1} \cdot \sum_{\substack{1\le i\ne j\le m\\1\le k\le n}} u_N(X_i, Y_k) \cdot 1_{\{X_j \le Y_k\}},$$
$$U^3_{m,n} = m^{-1}n^{-1}(n-1)^{-1} \cdot \sum_{\substack{1\le i\le m\\1\le k\ne q\le n}} u_N(X_i, Y_k) \cdot 1_{\{Y_q \le Y_k\}}.$$

Then (5.75) is equal to

$$U^1_{m,n} - \frac{\lambda_N \cdot (m-1)}{m} \cdot U^2_{m,n}$$
$$- \lambda_N \cdot m^{-2}n^{-1} \cdot \sum_{\substack{1\le i\le m\\1\le k\le n}} u_N(X_i, Y_k) \cdot 1_{\{X_i \le Y_k\}} - (1-\lambda_N) \cdot m^{-1}n^{-2} \cdot \sum_{\substack{1\le i\le m\\1\le k\le n}} u_N(X_i, Y_k) \tag{5.77}$$
$$- \frac{(1-\lambda_N) \cdot (n-1)}{n} \cdot U^3_{m,n}.$$

Now, the kernel function u_N is bounded:

$$\|u_N\| \le 2\|K'\| a_N^{-2}.$$

Which means for the sums in (5.77) we can write

$$\Big| - \lambda_N \cdot m^{-2}n^{-1} \cdot \sum_{\substack{1\le i\le m\\1\le k\le n}} u_N(X_i, Y_k) \cdot 1_{\{X_i \le Y_k\}} - (1-\lambda_N) \cdot m^{-1}n^{-2} \cdot \sum_{\substack{1\le i\le m\\1\le k\le n}} u_N(X_i, Y_k)\Big|$$
$$\le \lambda_N \cdot m^{-2}n^{-1} \cdot \sum_{\substack{1\le i\le m\\1\le k\le n}} \big|u_N(X_i, Y_k) \cdot 1_{\{X_i \le Y_k\}}\big| + (1-\lambda_N) \cdot m^{-1}n^{-2} \cdot \sum_{\substack{1\le i\le m\\1\le k\le n}} \big|u_N(X_i, Y_k)\big|$$
$$\le \lambda_N \cdot m^{-2}n^{-1} \cdot \sum_{\substack{1\le i\le m\\1\le k\le n}} \|u_N\| + (1-\lambda_N) \cdot m^{-1}n^{-2} \cdot \sum_{\substack{1\le i\le m\\1\le k\le n}} \|u_N\|$$
$$= \big[\lambda_N \cdot m^{-1} + (1-\lambda_N) \cdot n^{-1}\big] \cdot \|u_N\|$$
$$= O(a_N^{-2} \cdot N^{-1}). \tag{5.78}$$

Thus, we can partition (5.75) into the sum of three scaled U-statistics and a negligible rest:

$$U^1_{m,n} - \frac{\lambda_N \cdot (m-1)}{m} \cdot U^2_{m,n} - \frac{(1-\lambda_N)\cdot(n-1)}{n} \cdot U^3_{m,n} + O(a_N^{-2} \cdot N^{-1}). \tag{5.79}$$

In the following we will show that (5.79) is $O(a_N^{-2} \cdot N^{-1})$ as well, which will complete the proof. Begin by calculating each of the projections $\hat{U}^1_{m,n}$, $\hat{U}^2_{m,n}$ and $\hat{U}^3_{m,n}$. Firstly,

$$\begin{aligned}\hat{U}^1_{m,n} = m^{-1} \cdot \sum_{i=1}^m \int u_N(X_i, y) \cdot H_N(y)\, G(dy) + n^{-1} \cdot \sum_{k=1}^n \int u_N(x, Y_k) \cdot H_N(Y_k)\, F(dx) \\ - \iint u_N(x,y) \cdot H_N(y)\, F(dx)G(dy).\end{aligned}$$

Nextly, for $\lambda_N \cdot \hat{U}^2_{m,n}$ we have

$$\begin{aligned}&\lambda_N \cdot \hat{U}^2_{m,n} \\ &= \lambda_N \Big[m^{-1} \cdot \sum_{i=1}^m \Big[\iint u_N(X_i, y) \cdot 1_{\{z \le y\}}\, G(dy)F(dz) + \iint u_N(x,y) \cdot 1_{\{X_i \le y\}}\, F(dx)G(dz) \Big] \\ &\quad + n^{-1} \cdot \sum_{k=1}^n \iint u_N(x, Y_k) \cdot 1_{\{z \le Y_k\}}\, F(dx)F(dz) - 2 \iiint u_N(x,y) \cdot 1_{\{z \le y\}}\, F(dx)G(dy)F(dz) \Big] \\ &= m^{-1} \cdot \sum_{i=1}^m \int u_N(X_i, y) \cdot \lambda_N F(y)\, G(dy) + n^{-1} \cdot \sum_{k=1}^n \int u_N(x, Y_k) \cdot \lambda_N F(Y_k)\, F(dx) \\ &\quad - \iint u_N(x,y) \cdot \lambda_N F(y)\, F(dx)G(dy) \\ &\quad + \lambda_N \cdot m^{-1} \cdot \sum_{i=1}^m \Big[\iint u_N(x,y) \cdot 1_{\{X_i \le y\}}\, F(dx)G(dy) - \iint u_N(x,y) \cdot F(y)\, F(dx)G(dy) \Big].\end{aligned}$$

And lastly, for $(1-\lambda_N) \cdot \hat{U}^3_{m,n}$ we have

$$\begin{aligned}&(1-\lambda_N) \cdot \hat{U}^3_{m,n} \\ &= (1-\lambda_N) \Big[m^{-1} \cdot \sum_{i=1}^m \iint u_N(X_i, y) \cdot 1_{\{w \le y\}}\, G(dy)G(dw) \\ &\quad + n^{-1} \cdot \sum_{k=1}^n \Big[\iint u_N(x, Y_k) \cdot 1_{\{w \le Y_k\}}\, F(dx)G(dw) + \iint u_N(x,y) \cdot 1_{\{Y_k \le y\}}\, F(dx)G(dy) \Big] \\ &\quad - 2 \iiint u_N(x,y) \cdot 1_{\{w \le y\}}\, F(dx)G(dy)G(dw) \Big] \\ &= m^{-1} \cdot \sum_{i=1}^m \int u_N(X_i, y) \cdot (1-\lambda_N)G(y)\, G(dy) + n^{-1} \cdot \sum_{k=1}^n \int u_N(x, Y_k) \cdot (1-\lambda_N)G(Y_k)\, F(dx) \\ &\quad - \iint u_N(x,y) \cdot (1-\lambda_N)G(y)\, F(dx)G(dy) \\ &\quad + (1-\lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^n \Big[\iint u_N(x,y) \cdot 1_{\{Y_k \le y\}}\, F(dx)G(dy) - \iint u_N(x,y) \cdot G(y)\, F(dx)G(dy) \Big]\end{aligned}$$

Now, since

$$H_N = \lambda_N \cdot F + (1-\lambda_N) \cdot G,$$

we see that

$$
\begin{aligned}
&\hat{U}^1_{m,n} - \lambda_N \cdot \hat{U}^2_{m,n} - (1-\lambda_N) \cdot \hat{U}^3_{m,n} \\
&\quad = -\lambda_N \cdot m^{-1} \cdot \sum_{i=1}^{m} \left[\iint u_N(x,y) \cdot 1_{\{X_i \le y\}} \, F(dx)G(dy) - \iint u_N(x,y) \cdot F(y) \, F(dx)G(dy) \right] \\
&\qquad - (1-\lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^{n} \left[\iint u_N(x,y) \cdot 1_{\{Y_k \le y\}} \, F(dx)G(dy) - \iint u_N(x,y) \cdot G(y) \, F(dx)G(dy) \right] \\
&\quad = 0.
\end{aligned}
$$

due to

$$
\begin{aligned}
&\int u_N(x,y) \, F(dx) \\
&\quad = \int \left[a_N^{-2} \cdot \left[K'\big(a_N^{-1}(H_N(x) - H_N(y))\big) - \int K'\big(a_N^{-1}(H_N(z) - H_N(y))\big) \, F(dz) \right] \right] F(dx) \\
&\quad = a_N^{-2} \cdot \left[\int K'\big(a_N^{-1}(H_N(x) - H_N(y))\big) \, F(dx) - \int K'\big(a_N^{-1}(H_N(z) - H_N(y))\big) \, F(dz) \right] \\
&\quad = 0.
\end{aligned}
$$

Thus, for (5.75) we have

$$
\begin{aligned}
&U^1_{m,n} - \frac{\lambda_N \cdot (m-1)}{m} \cdot U^2_{m,n} - \frac{(1-\lambda_N) \cdot (n-1)}{n} \cdot U^3_{m,n} + O(a_N^{-2} \cdot N^{-1}) \\
&\quad = U^1_{m,n} - \lambda_N \cdot U^2_{m,n} + \frac{\lambda_N}{m} \cdot U^2_{m,n} - (1-\lambda_N) \cdot U^3_{m,n} + \frac{(1-\lambda_N)}{n} \cdot U^3_{m,n} + O(a_N^{-2} \cdot N^{-1}) \\
&\quad = U^1_{m,n} - \hat{U}^1_{m,n} - \lambda_N \cdot \Big[U^2_{m,n} - \hat{U}^2_{m,n} \Big] + \frac{\lambda_N}{m} \cdot U^2_{m,n} - (1-\lambda_N) \cdot \Big[U^3_{m,n} - \hat{U}^3_{m,n} \Big] \\
&\qquad + \frac{(1-\lambda_N)}{n} \cdot U^3_{m,n} + O(a_N^{-2} \cdot N^{-1}),
\end{aligned}
$$

and it remains only to bound $\Big[U^1_{m,n} - \hat{U}^1_{m,n}\Big]$, $\Big[U^2_{m,n} - \hat{U}^2_{m,n}\Big]$, $\Big[U^3_{m,n} - \hat{U}^3_{m,n}\Big]$, $\frac{\lambda_N}{m} \cdot U^2_{m,n}$ and $\frac{(1-\lambda_N)}{n} \cdot U^3_{m,n}$.

Firstly, using Lemma A.3 we have

$$
\mathbb{E}\Big[U^1_{m,n} - \hat{U}^1_{m,n}\Big]^2 = m^{-1} n^{-1} \cdot \mathbb{E}\big[u^*_{1N}(X_1, Y_1)\big]^2
$$

for u^*_{1N} defined as

$$
\begin{aligned}
u^*_{1N}(r,s) = u_N(r,s) \cdot H_N(s) &- \int u_N(r,y) \cdot H_N(y) \, G(dy) \\
&- \int u_N(x,s) \cdot H_N(s) \, F(dx) + \iint u_N(x,y) \cdot H_N(y) \, F(dx)G(dy)
\end{aligned}
$$

so that for the expectation we have

$$
\begin{aligned}
&\mathbb{E}\big[u^*_{1N}(X_1, Y_1)\big]^2 \\
&\quad \le 4 \cdot \mathbb{E}\Bigg[\big[u_N(X_1,Y_1) \cdot H_N(Y_1)\big]^2 + \left[\int u_N(X_1,y) \cdot H_N(y) \, G(dy) \right]^2 \\
&\qquad + \left[\int u_N(x,Y_1) \cdot H_N(Y_1) \, F(dx) \right]^2 + \left[\iint u_N(x,y) \cdot H_N(y) \, F(dx)G(dy) \right]^2 \Bigg]
\end{aligned}
$$

$$\leq 4\cdot\mathbb{E}\Big[\big[u_N(X_1,Y_1)\big]^2+\int\big[u_N(X_1,y)\big]^2\,G(dy)$$

$$+\Big[\int u_N(x,Y_1)\,F(dx)\Big]^2+\Big[\iint u_N(x,y)\,F(dx)\cdot H_N(y)\,G(dy)\Big]^2\Big]$$

$$\leq 4\cdot\mathbb{E}\Big[\big[2a_N^{-2}\cdot\|K'\|\big]^2+\int\big[2a_N^{-2}\cdot\|K'\|\big]^2\,F(dy)\Big]$$

$$\leq 32\|K'\|^2\cdot a_N^{-4}.$$

This yields

$$\begin{aligned}\mathbb{E}\Big[U^1_{m,n}-\hat{U}^1_{m,n}\Big]^2&=m^{-1}n^{-1}\cdot\mathbb{E}\big[u^*_{1N}(X_1,Y_1)\big]^2\\&\leq m^{-1}n^{-1}\cdot 32\|K'\|^2\cdot a_N^{-4}\\&=32\|K\|^2\cdot a_N^{-4}\cdot m^{-1}n^{-1}\\&=O(a_N^{-4}\cdot N^{-2}).\end{aligned}\tag{5.80}$$

Similarly, using Lemma A.5 we have

$$\mathbb{E}\big[U^2_{m,n}-\hat{U}^2_{m,n}\big]^2=\Big[O(m^{-1}n^{-1}+O(m^{-2})\Big]\cdot\|u^*_{2N}\|^2$$

and

$$\mathbb{E}\big[U^3_{m,n}-\hat{U}^3_{m,n}\big]^2=\Big[O(m^{-1}n^{-1})+O(n^{-2})\Big]\cdot\|u^*_{3N}\|^2$$

for u^*_{2N} and u^*_{3N} defined as

$$u^*_{2N}(r,s,t)=u_N(r,s)\cdot 1_{\{t\leq s\}}-\iint u_N(r,y)\cdot 1_{\{z\leq y\}}\,G(dy)F(dz)-\iint u_N(x,s)\cdot 1_{\{z\leq s\}}\,F(dx)F(dz)$$

$$-\iint u_N(x,y)\cdot 1_{\{t\leq y\}}\,F(dx)G(dy)+2\cdot\iiint u_N(x,y)\cdot 1_{\{z\leq y\}}\,F(dx)G(dy)F(dz),$$

and

$$u^*_{3N}(r,s,t)=u_N(r,s)\cdot 1_{\{t\leq s\}}-\iint u_N(r,y)\cdot 1_{\{z\leq y\}}\,G(dy)G(dz)-\iint u_N(x,s)\cdot 1_{\{z\leq s\}}\,F(dx)G(dz)$$

$$-\iint u_N(x,y)\cdot 1_{\{t\leq y\}}\,F(dx)G(dy)+2\cdot\iiint u_N(x,y)\cdot 1_{\{z\leq y\}}\,F(dx)G(dy)G(dz).$$

Bounding u^*_{2N} we obtain

$$\|u^*_{2N}(r,s,t)\|$$

$$=\Big\|u_N(r,s)\cdot 1_{\{t\leq s\}}-\iint u_N(r,y)\cdot 1_{\{z\leq y\}}\,G(dy)F(dz)-\iint u_N(x,s)\cdot 1_{\{z\leq s\}}\,F(dx)F(dz)$$

$$-\iint u_N(x,y)\cdot 1_{\{t\leq y\}}\,F(dx)G(dy)+2\cdot\iiint u_N(x,y)\cdot 1_{\{z\leq y\}}\,F(dx)G(dy)F(dz)\Big\|$$

$$=\Big\|u_N(r,s)\cdot 1_{\{t\leq s\}}-\iint u_N(r,y)\cdot 1_{\{z\leq y\}}\,F(dy)F(dz)-\iint u_N(x,s)\,F(dx)\cdot 1_{\{z\leq s\}}\,F(dz)$$

$$-\iint u_N(x,y)\,F(dx)\cdot 1_{\{t\leq y\}}\,F(dy)+2\cdot\iiint u_N(x,y)\,F(dx)\cdot 1_{\{z\leq y\}}\,F(dy)F(dz)\Big\|$$

$$\leq\|u_N(r,s)\|+\iint\|u_N(r,y)\|\,F(dy)F(dz)$$

$$\leq 2a_N^{-2} \cdot \|K'\| + 2a_N^{-2} \cdot \|K'\|$$
$$= 4a_N^{-2} \cdot \|K'\|.$$

Completely analogous arguments show that

$$\|u_{3N}^*(r,s,t)\| \leq 4a_N^{-2} \cdot \|K'\|$$

as well. This gives us

$$\mathbb{E}\big[U_{m,n}^2 - \hat{U}_{m,n}^2\big]^2 = \Big[O(m^{-1}n^{-1}) + O(m^{-2})\Big] \cdot \|u_{2N}^*\|^2 = O(a_N^{-4} \cdot N^{-2}) \tag{5.81}$$

and

$$\mathbb{E}\big[U_{m,n}^3 - \hat{U}_{m,n}^3\big]^2 = \Big[O(m^{-1}n^{-1}) + O(n^{-2})\Big] \cdot \|u_{3N}^*\|^2 = O(a_N^{-4} \cdot N^{-2}). \tag{5.82}$$

Lastly,

$$\begin{aligned}
\frac{\lambda_N}{m} \cdot U_{m,n}^2 &= \frac{\lambda_N}{m} \cdot m^{-1}(m-1)^{-1}n^{-1} \cdot \sum_{\substack{1\leq i\neq j\leq m\\ 1\leq k\leq n}} u_N(X_i, Y_k) \cdot 1_{\{X_j \leq Y_k\}} \\
&\leq \frac{\lambda_N}{m^2(m-1)n} \cdot \sum_{\substack{1\leq i\neq j\leq m\\ 1\leq k\leq n}} \|u_N\| \\
&= \frac{\lambda_N \cdot \|u_N\|}{m} \\
&\leq \frac{\lambda_N \cdot 2\|K'\|a_N^{-2}}{m} \\
&= O(a_N^{-2} \cdot N^{-1})
\end{aligned} \tag{5.83}$$

and

$$\begin{aligned}
\frac{(1-\lambda_N)}{n} \cdot U_{m,n}^3 &= \frac{(1-\lambda_N)}{n} \cdot m^{-1}n^{-1}(n-1)^{-1} \cdot \sum_{\substack{1\leq i\leq m\\ 1\leq k\neq q\leq n}} u_N(X_i, Y_k) \cdot 1_{\{Y_q \leq Y_k\}} \\
&\leq \frac{(1-\lambda_N)}{n^2(n-1)m} \cdot \sum_{\substack{1\leq i\leq m\\ 1\leq k\neq q\leq n}} \|u_N\| \\
&= \frac{(1-\lambda_N) \cdot \|u_N\|}{n} \\
&\leq \frac{(1-\lambda_N) \cdot 2\|K'\|a_N^{-2}}{n} \\
&= O(a_N^{-2} \cdot N^{-1}).
\end{aligned} \tag{5.84}$$

Combining (5.80), (5.81), (5.82), (5.83) and (5.84) we see that (5.75) is equal to

$$\begin{aligned}
&O_P(a_N^{-2} \cdot N^{-1}) - \lambda_N \cdot O_P(a_N^{-2} \cdot N^{-1}) + O(a_N^{-2} \cdot N^{-1}) - (1-\lambda_N) \cdot O_P(a_N^{-2} \cdot N^{-1}) \\
&\quad + O(a_N^{-2} \cdot N^{-1}) + O_P(a_N^{-2} \cdot N^{-1}) = O_P(a_N^{-2} \cdot N^{-1})
\end{aligned}$$

which completes the proof. □

LEMMA 5.16.

$$a_N^{-1} \cdot m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \Bigg[\int_{a_N^{-1}(H_N(X_i)-H_N(Y_k))}^{a_N^{-1}(H_N(X_i)-\hat{H}_N(Y_k))} \big(a_N^{-1}(H_N(X_i)-\hat{H}_N(Y_k))-t\big)\cdot K''(t)\,dt$$
$$- \iint_{a_N^{-1}(H_N(x)-H_N(Y_k))}^{a_N^{-1}(H_N(x)-\hat{H}_N(Y_k))} \big(a_N^{-1}(H_N(x)-\hat{H}_N(Y_k))-t\big)\cdot K''(t)\,dt\,F(dx)\Bigg]$$
$$= O_P(a_N^{-3}\cdot N^{-\frac{5}{4}}).$$

PROOF. As in the proof that (5.49) is $O_P(a_N^{-3}\cdot N^{-\frac{5}{4}})$ begin by defining

$$\hat{u}_N(r,s) = a_N^{-1}\cdot \Bigg[\int_{a_N^{-1}(H_N(r)-H_N(s))}^{a_N^{-1}(H_N(r)-\hat{H}_N(s))} \big(a_N^{-1}(H_N(r)-\hat{H}_N(s))-t\big)\cdot K''(t)\,dt$$
$$- \iint_{a_N^{-1}(H_N(x)-H_N(s))}^{a_N^{-1}(H_N(x)-\hat{H}_N(s))} \big(a_N^{-1}(H_N(x)-\hat{H}_N(s))-t\big)\cdot K''(t)\,dt\,F(dx)\Bigg].$$

Then we may write (5.76) as

$$m^{-1}n^{-1}\cdot\sum_{i=1}^{m}\sum_{k=1}^{n}\hat{u}_N(X_i,Y_k). \tag{5.85}$$

Looking at the second moment of (5.85) we find

$$\mathbb{E}\Bigg[m^{-1}n^{-1}\cdot\sum_{i=1}^{m}\sum_{k=1}^{n}\hat{u}_N(X_i,Y_k)\Bigg]^2$$
$$= m^{-2}n^{-2}\cdot\Big[mn\cdot\mathbb{E}\big[\hat{u}_N(X_1,Y_1)\big]^2 \tag{5.86}$$
$$+ m(m-1)n\cdot\mathbb{E}\big[\hat{u}_N(X_1,Y_1)\cdot\hat{u}_N(X_2,Y_1)\big] \tag{5.87}$$
$$+ mn(n-1)\cdot\mathbb{E}\big[\hat{u}_N(X_1,Y_1)\cdot\hat{u}_N(X_1,Y_2)\big] \tag{5.88}$$
$$+ m(m-1)n(n-1)\cdot\mathbb{E}\big[\hat{u}_N(X_1,Y_1)\cdot\hat{u}_N(X_2,Y_2)\big]\Big]. \tag{5.89}$$

In order to derive bounds for some of the expectations in (5.86) through (5.89) we define $\hat{H}_N^*$ to be equal to $\hat{H}_N$ with X_1 removed from the sample. That is,

$$\hat{H}_N^*(x) = N^{-1}\cdot\Bigg[\sum_{i=2}^{m}1_{\{X_i\le x\}}+\sum_{k=1}^{n}1_{\{Y_k\le x\}}\Bigg]. \tag{5.90}$$

Using $\hat{H}_N^*$ as defined above we use the same decomposition of $\hat{u}_N$ as in the proof of Lemma 5.13 with $\hat{u}_{1N}^*$, $\hat{u}_{2N}^*$ and $\hat{u}_{3N}^*$ defined as in (5.66), (5.67) and (5.68). Then, as before,

$$\hat{u}_N = \hat{u}_{1N}^* + \hat{u}_{2N}^* + \hat{u}_{3N}^*,$$

and

$$\mathbb{E}\big[\hat{u}_N(X_1,Y_1)\cdot\hat{u}_N(Y_1,Y_2)\big]$$
$$\le \mathbb{E}\big[\hat{u}_{1N}^*(X_1,Y_1)\cdot\hat{u}_{1N}^*(X_2,Y_2)\big] + 2\cdot\Big[\mathbb{E}\big[\hat{u}_{1N}^*(X_1,Y_1)\big]^2\Big]^{\frac{1}{2}}\cdot\Big[\mathbb{E}\big[\hat{u}_{2N}^*(X_1,Y_1)\big]^2\Big]^{\frac{1}{2}}$$
$$+ 2\cdot\Big[\mathbb{E}\big[\hat{u}_{1N}^*(X_1,Y_1)\big]^2\Big]^{\frac{1}{2}}\cdot\Big[\mathbb{E}\big[\hat{u}_{3N}^*(X_1,Y_1)\big]^2\Big]^{\frac{1}{2}} + \mathbb{E}\big[\hat{u}_{2N}^*(X_1,Y_1)\big]^2$$

$$+ 2 \cdot \Big[\mathbb{E}\big[\hat{u}^*_{2N}(X_1, Y_1)\big]^2\Big]^{\frac{1}{2}} \cdot \Big[\mathbb{E}\big[\hat{u}^*_{3N}(X_1, Y_1)\big]^2\Big]^{\frac{1}{2}} + \mathbb{E}\big[\hat{u}^*_{3N}(X_1, Y_1)\big]^2.$$

This means we only need to bound the four expectations $\mathbb{E}\big[\hat{u}^*_{1N}(X_1, Y_1) \cdot \hat{u}^*_{1N}(X_2, Y_2)\big]$, $\mathbb{E}\big[\hat{u}^*_{1N}(X_1, Y_1)\big]^2$, $\mathbb{E}\big[\hat{u}^*_{2N}(X_1, Y_1)\big]^2$ and $\mathbb{E}\big[\hat{u}^*_{3N}(X_1, Y_1)\big]^2$ in order to bound the expectation in (5.89). Firstly,

$$\begin{aligned}
&\mathbb{E}\big[\hat{u}^*_{1N}(X_1, X_2) \cdot \hat{u}^*_{1N}(X_2, Y_2)\big] \\
&\quad = \mathbb{E}\Big[\mathbb{E}\big[\hat{u}^*_{1N}(X_1, Y_1) \mid X_2, X_3, \cdots, X_m,\ Y_1, Y_2, \cdots, Y_n\big] \cdot \hat{u}^*_{1N}(X_2, Y_2)\Big] \\
&\quad = 0,
\end{aligned}$$

since for the inner expectation

$$\begin{aligned}
&\mathbb{E}\big[\hat{u}^*_{1N}(X_1, Y_1) \mid X_2, X_3, \cdots, X_m,\ Y_1, Y_2, \cdots, Y_n\big] \\
&= a_N^{-1} \cdot \Bigg[\iint_{a_N^{-1}(H_N(x)-H_N(Y_1))}^{a_N^{-1}(H_N(x)-\hat{H}^*_N(Y_1))} \big(a_N^{-1}(H_N(x) - \hat{H}^*_N(Y_1)) - t\big) \cdot K''(t)\, dt\, F(dx) \\
&\qquad - \iint_{a_N^{-1}(H_N(x)-H_N(Y_1))}^{a_N^{-1}(H_N(x)-\hat{H}^*_N(Y_1))} \big(a_N^{-1}(H_N(x) - \hat{H}^*_N(Y_1)) - t\big) \cdot K''(t)\, dt\, F(dx)\Bigg] \\
&= 0,
\end{aligned}$$

so that the first expectation vanishes completely.

The other three expectations do not vanish, but can be bound adequately. Altogether we have

$$\mathbb{E}\big[\hat{u}^*_{1N}(X_1, Y_1)\big]^2 = O(a_N^{-6} \cdot N^{-2}), \tag{5.91}$$

$$\mathbb{E}\big[\hat{u}^*_{2N}(X_1, Y_1)\big]^2 = O(a_N^{-6} \cdot N^{-3}), \tag{5.92}$$

and

$$\mathbb{E}\big[\hat{u}^*_{3N}(X_1, Y_1)\big]^2 = O(a_N^{-6} \cdot N^{-4}). \tag{5.93}$$

The proof of (5.91), (5.92) and (5.93) is completely analogous to the proof showing the rates in (5.69), (5.70) and (5.71) with Y_1 in place of X_2.

Combining (5.91), (5.92) and (5.93), we have shown for the expectation in (5.89) that

$$\begin{aligned}
&\mathbb{E}\big[\hat{u}_N(X_1, Y_1) \cdot \hat{u}_N(X_2, Y_2)\big] \\
&\quad = 2 \cdot \Big[O(a_N^{-6} \cdot N^{-2}\Big]^{\frac{1}{2}} \cdot \Big[O(a_N^{-6} \cdot N^{-3})\Big]^{\frac{1}{2}} \\
&\qquad + 2 \cdot \Big[O(a_N^{-6} \cdot N^{-2}\Big]^{\frac{1}{2}} \cdot \Big[O(a_N^{-6} \cdot N^{-4}\Big]^{\frac{1}{2}} + O(a_N^{-6} \cdot N^{-3}) \\
&\qquad + 2 \cdot \Big[O(a_N^{-6} \cdot N^{-3}\Big]^{\frac{1}{2}} \cdot \Big[O(a_N^{-6} \cdot N^{-4}\Big]^{\frac{1}{2}} + O(a_N^{-6} \cdot N^{-4}) \\
&\quad = O(a_N^{-6} \cdot N^{-\frac{5}{2}}) + O(a_N^{-6} \cdot N^{-3}) + O(a_N^{-6} \cdot N^{-\frac{7}{2}}) + O(a_N^{-6} \cdot N^{-4}) \\
&\quad = O(a_N^{-6} \cdot N^{-\frac{5}{2}}).
\end{aligned}$$

Using the Cauchy-Schwarz inequality, all of the other expectations in (5.86) through (5.88) are bound by the expectation $\mathbb{E}\big[\hat{u}_N(X_1, Y_1)\big]^2$. Bounding this expression we have

$$\mathbb{E}\big[\hat{u}_N(X_1, X_2)\big]^2 = O(a_N^{-6} \cdot N^{-2}), \tag{5.94}$$

(proof completely analogous to the proof of (5.72) with Y_1 in place of X_2) which means that the summands (5.86) through (5.88) are all $O(a_N^{-6} \cdot N^{-3})$. Combining this with the fact that (5.89) is $O(a_N^{-6} \cdot N^{-\frac{5}{2}})$ gives us a rate for (5.85), namely

$$m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \hat{u}_N(X_i, Y_k) = O_P(a_N^{-3} \cdot N^{-\frac{5}{4}}). \tag{5.95}$$

which completes the proof. □

Combining Lemmas 5.11, 5.12, 5.13, 5.14, 5.15 and 5.16 we have proven the following.

Lemma 5.17.

$$\int \Big[\hat{f}_N - \bar{f}_N\Big] \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big] = O_P(a_N^{-2} \cdot N^{-1}), \tag{5.96}$$

$$\int \Big[\hat{g}_N - \bar{g}_N\Big] \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big] = O_P(a_N^{-2} \cdot N^{-1}) \tag{5.97}$$

and thus

$$\int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big] \circ H_N(x)\Big[\hat{F}_m(dx) - F(dx)\Big] = O_P(a_N^{-2} \cdot N^{-1}). \tag{5.98}$$

5.2.3. Second bounded term. We continue our treatment of the asymptotically negligible terms of the expansion by showing that the second term (2.36) is negligible as well. For (2.36) we can write

$$\int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\Big[\hat{F}_m(dx) - F(dx)\Big]$$

$$= \int \Big[\hat{f}_N - \bar{f}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\Big[\hat{F}_m(dx) - F(dx)\Big] \tag{5.99}$$

$$- \int \Big[\hat{g}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\Big[\hat{F}_m(dx) - F(dx)\Big] \tag{5.100}$$

We will first work at bounding (5.99). The proof for (5.100) follows along similar lines.

Recalling the definitions (2.12) and (2.18) of $\hat{f}_N$ and $\bar{f}_N$, we can compute the first order derivatives of these as

$$\hat{f}_N'(t) = (a_N^{-2} \cdot m^{-1}) \sum_{j=1}^{m} K'\Big(\frac{t - \hat{H}_N(X_j)}{a_N}\Big) \quad \text{and} \tag{5.101}$$

$$\bar{f}_N'(t) = a_N^{-2} \int K'\Big(\frac{t - H_N(y)}{a_N}\Big) F(dy) \tag{5.102}$$

respectively, so that for (5.99) we may write

$$\int \Big[\hat{f}_N - \bar{f}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\Big[\hat{F}_m(dx) - F(dx)\Big]$$

$$
\begin{aligned}
&= m^{-1} \cdot \sum_{i=1}^{m} \Bigg[\Bigg[a_N^{-2} \cdot m^{-1} \cdot \sum_{j=1}^{m} K'\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(X_j))\big) \\
&\qquad - a_N^{-2} \int K'\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, F(dy)\Bigg] \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \Bigg[a_N^{-2} \cdot m^{-1} \int \sum_{j=1}^{m} K'\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_j))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\qquad - a_N^{-2} \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Bigg]\Bigg] \\
&= a_N^{-2} \cdot m^{-2} \cdot \sum_{i=1}^{m} \sum_{j=1}^{m} \Bigg[\Bigg[K'\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(X_j))\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, F(dy)\Bigg] \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_j))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Bigg].
\end{aligned}
$$

At this point we separate the summands with $i = j$ and use the Taylor expansion of the kernel function K about each of the $a_N^{-1}(H_N(X_i) - H_N(X_j))$ and $a_N^{-1}(H_N(x) - H_N(X_j))$ for the remaining summands with $i \neq j$, which yields

$$
\begin{aligned}
&\int \Big[\hat{f}_N - \bar{f}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\Big[\hat{F}_m(dx) - F(dx)\Big] \\
&= a_N^{-2} \cdot m^{-2} \cdot \sum_{i=1}^{m} \Bigg[K'(0) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, F(dy) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x) - H_N(X_i))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Bigg] && (5.103) \\
&\quad + a_N^{-2} \cdot m^{-2} \cdot \sum_{1 \le i \neq j \le m} \Bigg[K'\big(a_N^{-1}(H_N(X_i) - H_N(X_j))\big) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, F(dy) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x) - H_N(X_j))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Bigg] && (5.104) \\
&\quad + a_N^{-3} \cdot m^{-2} \cdot \sum_{1 \le i \neq j \le m} \Bigg[K''\big(a_N^{-1}(H_N(X_i) - H_N(X_j))\big) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_j))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Bigg] \cdot \big(H_N(X_j) - \hat{H}_N(X_j)\big) && (5.105)
\end{aligned}
$$

$$+\frac{1}{2}a_N^{-4}\cdot m^{-2}\cdot\sum_{1\le i\ne j\le m}\Big[K'''(\xi_{ij})\cdot\big(\hat{H}_N(X_i)-H_N(X_i)\big)$$
$$-\int K'''(\tau_j)\cdot\big(\hat{H}_N(x)-H_N(x)\big)\,F(dx)\Big]\cdot\big(H_N(X_j)-\hat{H}_N(X_j)\big)^2 \tag{5.106}$$

where ξ_{ij} and τ_j are appropriate values between the two ratios.

Since the Kernel function K is assumed to be bounded, it is easy to see that (5.103) is $O(a_N^{-2}\cdot N^{-\frac{3}{2}})$, since

$$\Big|a_N^{-2}\cdot m^{-2}\cdot\sum_{i=1}^{m}\Big[K'(0)\cdot\big(\hat{H}_N(X_i)-H_N(X_i)\big)$$
$$-\int K'\big(a_N^{-1}(H_N(X_i)-H_N(y))\big)\,F(dy)\cdot\big(\hat{H}_N(X_i)-H_N(X_i)\big)$$
$$-\int K'\big(a_N^{-1}(H_N(x)-H_N(X_i))\big)\cdot\big(\hat{H}_N(x)-H_N(x)\big)\,F(dx)$$
$$+\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,F(dy)\big(\hat{H}_N(x)-H_N(x)\big)\,F(dx)\Big]\Big|$$
$$\le a_N^{-2}\cdot m^{-1}\cdot 4\|K'\|\cdot\|\hat{H}_N-H_N\|$$
$$=O_P(a_N^{-2}\cdot N^{-\frac{3}{2}}) \tag{5.107}$$

due to the D-K-W bound on $\|\hat{H}_N-H_N\|$.

In the following lemmas we will derive bounds for the remaining three terms (5.104), (5.105) and (5.106).

LEMMA 5.18.

$$a_N^{-2}\cdot m^{-2}\cdot\sum_{1\le i\ne j\le m}\Big[K'\big(a_N^{-1}(H_N(X_i)-H_N(X_j))\big)\cdot\big(\hat{H}_N(X_i)-H_N(X_i)\big)$$
$$-\int K'\big(a_N^{-1}(H_N(X_i)-H_N(y))\big)\,F(dy)\cdot\big(\hat{H}_N(X_i)-H_N(X_i)\big)$$
$$-\int K'\big(a_N^{-1}(H_N(x)-H_N(X_j))\big)\cdot\big(\hat{H}_N(x)-H_N(x)\big)\,F(dx)$$
$$+\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,F(dy)\big(\hat{H}_N(x)-H_N(x)\big)\,F(dx)\Big]$$
$$=O_P(a_N^{-\frac{3}{2}}\cdot N^{-1}).$$

PROOF. Begin by defining

$$\hat{u}_N(r,s)=a_N^{-2}\cdot\Big[K'\big(a_N^{-1}(H_N(r)-H_N(s))\big)\cdot\big(\hat{H}_N(r)-H_N(r)\big)$$
$$-\int K'\big(a_N^{-1}(H_N(r)-H_N(y))\big)\,F(dy)\cdot\big(\hat{H}_N(r)-H_N(r)\big)$$
$$-\int K'\big(a_N^{-1}(H_N(x)-H_N(s))\big)\cdot\big(\hat{H}_N(x)-H_N(x)\big)\,F(dx)$$
$$+\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,F(dy)\big(\hat{H}_N(x)-H_N(x)\big)\,F(dx)\Big].$$

Then we may write (5.104) as

$$m^{-2} \cdot \sum_{1 \le i \ne j \le m} \hat{u}_N(X_i, X_j). \tag{5.108}$$

In order to derive bounds for (5.108) we look at the second moment and can use the standard expansion to write

$$\mathbb{E}\Big[m^{-2} \cdot \sum_{1 \le i \ne j \le m} \hat{u}_N(X_i, X_j)\Big]^2$$

$$= m^{-4} \cdot \Big[m(m-1) \cdot \mathbb{E}\big[\hat{u}_N(X_1, X_2)\big]^2 \tag{5.109}$$

$$+ m(m-1) \cdot \mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_2, X_1)\big] \tag{5.110}$$

$$+ 2m(m-1)(m-2) \cdot \mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_3, X_1)\big] \tag{5.111}$$

$$+ m(m-1)(m-2) \cdot \mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_1, X_3)\big] \tag{5.112}$$

$$+ m(m-1)(m-2) \cdot \mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_3, X_2)\big] \tag{5.113}$$

$$+ m(m-1)(m-2)(m-3) \cdot \mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_3, X_4)\big]\Big]. \tag{5.114}$$

In order to derive bounds for some of the expectations in (5.109) through (5.114) we again define $\hat{H}_N^*$ to be equal to $\hat{H}_N$ with X_1, X_2, X_3 and X_4 removed from the sample. That is,

$$\hat{H}_N^*(x) = N^{-1} \cdot \Big[\sum_{i=5}^{m} 1_{\{X_i \le x\}} + \sum_{k=1}^{n} 1_{\{Y_k \le x\}}\Big].$$

Also, define $\hat{u}_N^*$ as $\hat{u}_N$ with all occurrences of $\hat{H}_N$ replaced by $\hat{H}_N^*$, and recall that for any X_{i_1}, X_{i_2}, X_{i_3} and X_{i_4} from the sample, we have

$$\mathbb{E}\big[\hat{u}_N(X_{i_1}, X_{i_2}) \cdot \hat{u}_N(X_{i_3}, X_{i_4})\big]$$

$$= \mathbb{E}\big[\big((\hat{u}_N - \hat{u}_N^*)(X_{i_1}, X_{i_2}) + \hat{u}_N^*(X_{i_1}, X_{i_2})\big) \cdot \big((\hat{u}_N - \hat{u}_N^*)(X_{i_3}, X_{i_4}) + \hat{u}_N^*(X_{i_3}, X_{i_4})\big)\big]$$

$$= \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_{i_1}, X_{i_2}) \cdot (\hat{u}_N - \hat{u}_N^*)(X_{i_3}, X_{i_4})\big] + \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_{i_1}, X_{i_2}) \cdot \hat{u}_N^*(X_{i_3}, X_{i_4})\big]$$

$$+ \mathbb{E}\big[\hat{u}_N^*(X_{i_1}, X_{i_2}) \cdot (\hat{u}_N - \hat{u}_N^*)(X_{i_3}, X_{i_4})\big] + \mathbb{E}\big[\hat{u}_N^*(X_{i_1}, X_{i_2}) \cdot \hat{u}_N^*(X_{i_3}, X_{i_4})\big], \tag{5.115}$$

Since in our case $i_1 \ne i_2$ and $i_3 \ne i_4$, the Cauchy-inequality can be applied to the expectations on the right to get the bound

$$\big|\mathbb{E}\big[\hat{u}_N(X_{i_1}, X_{i_2}) \cdot \hat{u}_N(X_{i_3}, X_{i_4})\big]\big|$$

$$\le \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_{i_1}, X_{i_2})\big]^2 + 2\Big[\mathbb{E}\big[\hat{u}_N^*(X_{i_1}, X_{i_2})\big]^2 \cdot \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_{i_1}, X_{i_2})\big]^2\Big]^{\frac{1}{2}}$$

$$+ \big|\mathbb{E}\big[\hat{u}_N^*(X_{i_1}, X_{i_2}) \cdot \hat{u}_N^*(X_{i_3}, X_{i_4})\big]\big| \tag{5.116}$$

In the following, we will use the equation (5.115) and the inequality (5.116) to bound the expectations in (5.111) through (5.114).

We begin by applying the inequality (5.116) to the expectations in (5.111), (5.112) and (5.113) and show that for each of these the last expectation on the right hand side of the inequality vanishes. In the case of (5.111) we have

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\cdot \hat{u}_N^*(X_3,X_1)\big]\\
&\quad = \mathbb{E}\Big[\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\bigm| X_1,X_5,\dots,X_m,\ Y_1,\dots,Y_n\big]\cdot \mathbb{E}\big[\hat{u}_N^*(X_3,X_1)\bigm| X_1,X_5,\dots,X_m,\ Y_1,\dots,Y_n\big]\Big]\\
&\quad = 0,
\end{aligned}
$$

since for the first inner expectation

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\bigm| X_1,X_5,\dots,X_m,\ Y_1,\dots,Y_n\big]\\
&\quad = \mathbb{E}\Big[a_N^{-2}\cdot\Big[K'\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)\\
&\qquad\quad -\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,F(dy)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)\\
&\qquad\quad -\int K'\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\\
&\qquad\quad +\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,F(dy)\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\Big]\Bigm| X_1,X_5,\dots,X_m,\ Y_1,\dots,Y_n\Big]\\
&\quad = a_N^{-2}\cdot\Big[\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)\,F(dy)\\
&\qquad\quad -\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,F(dy)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)\\
&\qquad\quad -\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\,F(dy)\\
&\qquad\quad +\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,F(dy)\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\Big]\\
&\quad = 0.
\end{aligned}
$$

The expectation in (5.112) is quickly seen to vanish as well, due to

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\cdot \hat{u}_N^*(X_1,X_3)\big]\\
&\quad = \mathbb{E}\Big[\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\bigm| X_1,X_5,\dots,X_m,\ Y_1,\dots,Y_n\big]\cdot \mathbb{E}\big[\hat{u}_N^*(X_1,X_3)\bigm| X_1,X_5,\dots,X_m,\ Y_1,\dots,Y_n\big]\Big]\\
&\quad = 0,
\end{aligned}
$$

since we already know from the above that the first inner expectation vanishes.

Lastly, for the expectation in (5.113) we get

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\cdot \hat{u}_N^*(X_3,X_2)\big]\\
&\quad = \mathbb{E}\Big[\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\bigm| X_2,X_5,\dots,X_m,\ Y_1,\dots,Y_n\big]\cdot \mathbb{E}\big[\hat{u}_N^*(X_3,X_2)\bigm| X_2,X_5,\dots,X_m,\ Y_1,\dots,Y_n\big]\Big]\\
&\quad = 0,
\end{aligned}
$$

since for the first inner expectation

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\bigm| X_2,X_5,\dots,X_m,\ Y_1,\dots,Y_n\big]\\
&\quad = \mathbb{E}\Big[a_N^{-2}\cdot\Big[K'\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)
\end{aligned}
$$

$$
\begin{aligned}
&\quad - \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\quad - \int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx) \\
&\quad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big] \,\Big|\, X_2, X_5, \ldots, X_m,\ Y_1, \ldots, Y_n\Big] \\
&= a_N^{-2} \cdot \Big[\int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx) \\
&\quad - \int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx) \\
&\quad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big] \\
&= 0.
\end{aligned}
$$

Thus, using inequality (5.116) it remains only to bound the expectations $\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)\big]^2$ and $\mathbb{E}\big[\hat{u}_N^*(X_1, X_2)\big]^2$ in order to derive bounds for the summands (5.111), (5.112) and (5.113). For the first of these expectations note first that

$$
\begin{aligned}
&\hat{H}_N(x) - \hat{H}_N^*(x) \\
&\quad = N^{-1} \cdot \Big[\sum_{i=1}^{m} 1_{\{X_i \le x\}} + \sum_{k=1}^{n} 1_{\{Y_k \le x\}}\Big] - N^{-1} \cdot \Big[\sum_{i=5}^{m} 1_{\{X_i \le x\}} + \sum_{k=1}^{n} 1_{\{Y_k \le x\}}\Big] \\
&\quad = N^{-1} \cdot \sum_{i=1}^{4} 1_{\{X_i \le x\}},
\end{aligned}
$$

so that

$$
\begin{aligned}
&\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)\big]^2 \\
&\quad = \mathbb{E}\Big[a_N^{-2} \cdot \Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot N^{-1} \cdot \sum_{i=1}^{4} 1_{\{X_i \le X_1\}} \\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \cdot N^{-1} \cdot \sum_{i=1}^{4} 1_{\{X_i \le X_1\}} \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot N^{-1} \cdot \sum_{i=1}^{4} 1_{\{X_i \le x\}}\, F(dx) \\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) N^{-1} \cdot \sum_{i=1}^{4} 1_{\{X_i \le x\}}\, F(dx)\Big]\Big]^2 \\
&\quad = a_N^{-4} \cdot N^{-2} \cdot \mathbb{E}\Big[\sum_{i=1}^{4} \Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot 1_{\{X_i \le X_1\}} \\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_i \le X_1\}} \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot 1_{\{X_i \le x\}}\, F(dx)
\end{aligned}
$$

$$
\begin{aligned}
&+ \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_i \le x\}}\, F(dx)\Big]\Big]^2 \\
&\le 4a_N^{-4} \cdot N^{-2} \cdot \sum_{i=1}^{4} \Big[\mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot 1_{\{X_i \le X_1\}} \\
&- \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_i \le X_1\}} \\
&- \int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot 1_{\{X_i \le x\}}\, F(dx) \\
&+ \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_i \le x\}}\, F(dx)\Big]^2\Big] \\
&\le 4^2 a_N^{-4} \cdot N^{-2} \cdot \sum_{i=1}^{4} \Big[\mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot 1_{\{X_i \le X_1\}}\Big]^2 \\
&+ \mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_i \le X_1\}}\Big]^2 \\
&+ \mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot 1_{\{X_i \le x\}}\, F(dx)\Big]^2 \\
&+ \mathbb{E}\Big[\iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_i \le x\}}\, F(dx)\Big]^2\Big] \\
&\le 4^3 a_N^{-4} \cdot N^{-2} \cdot \Big[\mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big)\Big]^2 \\
&+ \mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy)\Big]^2 \\
&+ \mathbb{E}\Big[\int \Big(K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big)\Big)^2\, F(dx)\Big] \\
&+ \mathbb{E}\Big[\iint \Big(K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\Big)^2\, F(dy)\, F(dx)\Big]\Big] \\
&\le 4^4 a_N^{-4} \cdot N^{-2} \cdot \iint \Big(K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\Big)^2\, F(dx)\, F(dy)
\end{aligned}
$$

Using the bounds provided in (A.2) in Lemma A.1 this is less than or equal to

$$
4^4 a_N^{-4} \cdot N^{-2} \cdot 2\big\|K'\big\|^2 \cdot a_N\Big(1 + \frac{n}{m}\Big) = 2 \cdot 4^4\big\|K'\big\|^2 \cdot a_N^{-3} \cdot N^{-2}\Big(1 + \frac{n}{m}\Big).
$$

Thus for $\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)\big]^2$ we may write

$$
\begin{aligned}
\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)\big]^2 &\le 2 \cdot 4^4\big\|K'\big\|^2 \cdot a_N^{-3} \cdot N^{-2}\Big(1 + \frac{n}{m}\Big) \\
&= O(a_N^{-3} \cdot N^{-2}).
\end{aligned} \tag{5.117}
$$

Further, for the second expectation, $\mathbb{E}\big[\hat{u}_N^*(X_1, X_2)\big]^2$, we have

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_N^*(X_1, X_2)\big]^2 \\
&= \mathbb{E}\Big[a_N^{-2} \cdot \Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\quad - \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big)
\end{aligned}
$$

$$
\begin{aligned}
&- \int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx) \\
&+ \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big]\Big]^2 \\
\leq 4a_N^{-4} \cdot &\Big[\mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big)\Big]^2 \\
&+ \mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big)\Big]^2 \\
&+ \mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big]^2 \\
&+ \mathbb{E}\Big[\iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big]^2\Big] \\
\leq 4a_N^{-4} \cdot &\Big[\mathbb{E}\Big[\int \Big(K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\Big)^2 F(dy) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big)^2\Big] \\
&+ \mathbb{E}\Big[\int \Big(K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\Big)^2 F(dy) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big)^2\Big] \\
&+ \mathbb{E}\Big[\int \Big(K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big)\Big)^2 \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)^2 F(dx)\Big] \\
&+ \mathbb{E}\Big[\iint \Big(K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\Big)^2 F(dy)\big(\hat{H}_N^*(x) - H_N(x)\big)^2 F(dx)\Big]\Big].
\end{aligned}
$$

Using the bounds from (A.2) in Lemma A.1 this is less than or equal to

$$
\begin{aligned}
4a_N^{-4} \cdot 2\big\|K'\big\|^2 \cdot a_N\Big(1 + \frac{n}{m}\Big) \cdot &\Big[\mathbb{E}\Big[\hat{H}_N^*(X_1) - H_N(X_1)\Big]^2 \\
&+ \mathbb{E}\Big[\hat{H}_N^*(X_1) - H_N(X_1)\Big]^2 \\
&+ \mathbb{E}\Big[\big\|\hat{H}_N^* - H_N\big\|^2\Big] \\
&+ \mathbb{E}\Big[\big\|\hat{H}_N^* - H_N\big\|^2\Big]\Big] \\
\leq 2 \cdot 4^2 a_N^{-3} \cdot \big\|K'\big\|^2 \cdot \Big(1 + \frac{n}{m}\Big) &\cdot \mathbb{E}\Big[\big\|\hat{H}_N^* - H_N\big\|^2\Big].
\end{aligned}
$$

Thus for $\mathbb{E}\big[\hat{u}_N^*(X_1, X_2)\big]^2$ we may write

$$
\begin{aligned}
\mathbb{E}\big[\hat{u}_N^*(X_1, X_2)\big]^2 &\leq 2 \cdot 4^2 a_N^{-3} \cdot \big\|K'\big\|^2 \cdot (1 + \frac{n}{m}) \cdot \mathbb{E}\Big[\big\|\hat{H}_N^* - H_N\big\|^2\Big] \\
&= O(a_N^{-3} \cdot N^{-1}).
\end{aligned} \tag{5.118}
$$

Using these bounds in the inequality (5.116) we have shown for the expectations in (5.111), (5.112) and (5.113) that these are all less than or equal to

$$
\begin{aligned}
&\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)\big]^2 + 2\Big[\mathbb{E}\big[\hat{u}_N^*(X_1, X_2)\big]^2 \cdot \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)\big]^2\Big]^{\frac{1}{2}} \\
&= O(a_N^{-3} \cdot N^{-2}) + \big[O(a_N^{-3} \cdot N^{-1}) \cdot O(a_N^{-3} \cdot N^{-2})\big]^{\frac{1}{2}} \\
&= O(a_N^{-3} \cdot N^{-2}) + O(a_N^{-3} \cdot N^{-\frac{3}{2}}) \\
&= O(a_N^{-3} \cdot N^{-\frac{3}{2}}).
\end{aligned} \tag{5.119}
$$

The expectation in the summand (5.110) is less than or equal to the expectation $\mathbb{E}\big[\hat{u}_N(X_1,X_2)\big]^2$ in (5.109), so that we only need to bound this simpler expectation to bound both terms. Using (5.116), (5.118) and (5.117) once again, we obtain

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_N(X_1,X_2)\big]^2 \\
&\quad \le \mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,X_2)\big]^2 + 2\Big[\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\big]^2\cdot \mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,X_2)\big]^2\Big]^{\frac{1}{2}} \\
&\qquad + \big|\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\cdot \hat{u}_N^*(X_1,X_2)\big]\big| \\
&\quad \le \mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,X_2)\big]^2 + 2\Big[\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\big]^2\cdot \mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,X_2)\big]^2\Big]^{\frac{1}{2}} \\
&\qquad + \mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\big]^2 \\
&\quad = O(a_N^{-3}\cdot N^{-2}) + \big[O(a_N^{-3}\cdot N^{-1})\cdot O(a_N^{-3}\cdot N^{-2})\big]^{\frac{1}{2}} + O(a_N^{-3}\cdot N^{-1}) \\
&\quad = O(a_N^{-3}\cdot N^{-1}).
\end{aligned}
$$

Thus, the expectations in (5.109) and (5.110) are both of the order $O(a_N^{-3}\cdot N^{-1})$, and it remains only to bound the expectation in the last summand (5.114). In this case, we will use the equality (5.115), which tells us that

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_N(X_1,X_2)\cdot \hat{u}_N(X_3,X_4)\big] \\
&\quad = \mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,X_2)\cdot(\hat{u}_N-\hat{u}_N^*)(X_3,X_4)\big] + \mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,X_2)\cdot \hat{u}_N^*(X_3,X_4)\big] \\
&\qquad + \mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\cdot(\hat{u}_N-\hat{u}_N^*)(X_3,X_4)\big] + \mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\cdot \hat{u}_N^*(X_3,X_4)\big].
\end{aligned}
$$

The last expectation vanishes immediately:

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\cdot \hat{u}_N^*(X_3,X_4)\big] \\
&\quad = \mathbb{E}\Big[\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\bigm| X_5,\dots,X_m,\ Y_1,\dots,Y_n\big]\cdot \mathbb{E}\big[\hat{u}_N^*(X_3,X_4)\bigm| X_5,\dots,X_m,\ Y_1,\dots,Y_n\big]\Big] \\
&\quad = 0,
\end{aligned}
$$

since for the inner expectation

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\bigm| X_5,\dots,X_m,\ Y_1,\dots,Y_n\big] \\
&\quad = \mathbb{E}\Big[a_N^{-2}\cdot\Big[K'\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,F(dy)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx) \\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,F(dy)\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\Big]\Bigm| X_5,\dots,X_m,\ Y_1,\dots,Y_n\Big] \\
&\quad = a_N^{-2}\cdot\Big[\int K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\,F(dy) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,F(dy)\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)
\end{aligned}
$$

$$
\begin{aligned}
&\quad - \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\, F(dy) \\
&\quad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big] \\
&= 0.
\end{aligned}
$$

Interestingly, unlike in the case of the other expectations in (5.111) through (5.113), we will find that in the case of $\mathbb{E}\big[\hat{u}_N^*(X_1, X_2) \cdot \hat{u}_N^*(X_3, X_4)\big]$ the expectations $\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2) \cdot \hat{u}_N^*(X_3, X_4)\big]$ and $\mathbb{E}\big[\hat{u}_N^*(X_1, X_2) \cdot (\hat{u}_N - \hat{u}_N^*)(X_3, X_4)\big]$ on the right hand side of (5.115) vanish as well.

Recall that $\hat{H}_N(x) - \hat{H}_N^*(x) = N^{-1} \cdot \sum_{i=1}^4 1_{\{X_i \le x\}}$. Then we can note that the expression $(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)$ actually depends only on X_1, X_2, X_3 and X_4 and none of the rest of the sample, since

$$
\begin{aligned}
(\hat{u}_N - \hat{u}_N^*)(X_1, X_2) = a_N^{-2} \cdot \Big[&K'\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot N^{-1} \cdot \sum_{i=1}^4 1_{\{X_i \le X_1\}} \\
&- \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \cdot N^{-1} \cdot \sum_{i=1}^4 1_{\{X_i \le X_1\}} \\
&- \int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot N^{-1} \cdot \sum_{i=1}^4 1_{\{X_i \le x\}}\, F(dx) \\
&+ \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) N^{-1} \cdot \sum_{i=1}^4 1_{\{X_i \le x\}}\, F(dx)\Big].
\end{aligned}
$$

$\hat{u}_N^*(X_3, X_4)$, on the other hand, is equal to

$$
\begin{aligned}
\hat{u}_N^*(X_3, X_4) = a_N^{-2} \cdot \Big[&K'\big(a_N^{-1}(H_N(X_3) - H_N(X_4))\big) \cdot \big(\hat{H}_N^*(X_3) - H_N(X_3)\big) \\
&- \int K'\big(a_N^{-1}(H_N(X_3) - H_N(y))\big)\, F(dy) \cdot \big(\hat{H}_N^*(X_3) - H_N(X_3)\big) \\
&- \int K'\big(a_N^{-1}(H_N(x) - H_N(X_4))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx) \\
&+ \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big]
\end{aligned}
$$

which depends on $X_3, X_4, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$, but not on X_1 and X_2, so that for the expectation we may write

$$
\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2) \cdot \hat{u}_N^*(X_3, X_4)\big] = \mathbb{E}\Big[\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2) \mid X_3, X_4\big] \cdot \mathbb{E}\big[\hat{u}_N^*(X_3, X_4) \mid X_3, X_4\big]\Big].
$$

In the following we will show that the first inner expectation vanishes, since

$$
\begin{aligned}
&\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2) \mid X_3, X_4\big] \\
&\quad = \mathbb{E}\Big[a_N^{-2} \cdot \Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot N^{-1} \cdot \sum_{i=1}^4 1_{\{X_i \le X_1\}} \\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \cdot N^{-1} \cdot \sum_{i=1}^4 1_{\{X_i \le X_1\}}
\end{aligned}
$$

$$
\begin{aligned}
&- \int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot N^{-1} \cdot \sum_{i=1}^{4} 1_{\{X_i \le x\}}\, F(dx) \\
&+ \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) N^{-1} \cdot \sum_{i=1}^{4} 1_{\{X_i \le x\}}\, F(dx)\Big] \;\Big|\; X_3, X_4\Big] \\
= a_N^{-2} \cdot N^{-1} \cdot \mathbb{E}\Big[\sum_{i=1}^{4} \Big[&K'\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot 1_{\{X_i \le X_1\}} \\
&- \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_i \le X_1\}} \\
&- \int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot 1_{\{X_i \le x\}}\, F(dx) \\
&+ \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_i \le x\}}\, F(dx)\Big] \;\Big|\; X_3, X_4\Big] \\
= a_N^{-2} \cdot N^{-1} \cdot \mathbb{E}\Big[&K'\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot 1_{\{X_1 \le X_1\}} \\
&- \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_1 \le X_1\}} \\
&- \int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot 1_{\{X_1 \le x\}}\, F(dx) \\
&+ \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_1 \le x\}}\, F(dx) \;\Big|\; X_3, X_4\Big] \\
+ a_N^{-2} \cdot N^{-1} \cdot \mathbb{E}\Big[&K'\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot 1_{\{X_2 \le X_1\}} \\
&- \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_2 \le X_1\}} \\
&- \int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot 1_{\{X_2 \le x\}}\, F(dx) \\
&+ \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_2 \le x\}}\, F(dx) \;\Big|\; X_3, X_4\Big] \\
+ a_N^{-2} \cdot N^{-1} \cdot \sum_{i=3}^{4} \mathbb{E}\Big[&K'\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot 1_{\{X_i \le X_1\}} \\
&- \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_i \le X_1\}} \\
&- \int K'\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot 1_{\{X_i \le x\}}\, F(dx) \\
&+ \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_i \le x\}}\, F(dx) \;\Big|\; X_3, X_4\Big] \\
= a_N^{-2} \cdot N^{-1} \cdot \Big[&\iint K'\big(a_N^{-1}(H_N(w) - H_N(z))\big)\, F(dw)\, F(dz) \\
&- \iint K'\big(a_N^{-1}(H_N(w) - H_N(y))\big)\, F(dy)\, F(dw) \\
&- \iiint K'\big(a_N^{-1}(H_N(x) - H_N(z))\big) \cdot 1_{\{w \le x\}}\, F(dx)\, F(dw)\, F(dz) \\
&+ \iiint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot 1_{\{w \le x\}}\, F(dx)\, F(dw)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+ a_N^{-2} \cdot N^{-1} \cdot \iint \Big[K'\big(a_N^{-1}(H_N(w) - H_N(z))\big) \cdot 1_{\{z \le w\}}\, F(dw)\, F(dz) \\
&\quad - \iiint K'\big(a_N^{-1}(H_N(w) - H_N(y))\big)\, F(dy) \cdot 1_{\{z \le w\}}\, F(dw)\, F(dz) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(x) - H_N(z))\big) \cdot 1_{\{z \le x\}}\, F(dx)\, F(dz) \\
&\quad + \iiint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot 1_{\{z \le x\}}\, F(dx)\, F(dz) \Big] \\
&+ a_N^{-2} \cdot N^{-1} \cdot \sum_{i=3}^{4} \Big[\iint K'\big(a_N^{-1}(H_N(w) - H_N(z))\big) \cdot 1_{\{X_i \le w\}}\, F(dw)\, F(dz) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(w) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_i \le w\}}\, F(dw) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(x) - H_N(z))\big) \cdot 1_{\{X_i \le x\}}\, F(dx)\, F(dz) \\
&\quad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot 1_{\{X_i \le x\}}\, F(dx) \Big] \\
&= a_N^{-2} \cdot N^{-1} \cdot 0 + a_N^{-2} \cdot N^{-1} \cdot 0 + a_N^{-2} \cdot N^{-1} \cdot \sum_{i=3}^{4} 0 \\
&= 0.
\end{aligned}
$$

In the same manner we can show that the expectation $\mathbb{E}\big[\hat{u}_N^*(X_1, X_2) \cdot (\hat{u}_N - \hat{u}_N^*)(X_3, X_4)\big]$ vanishes as well, since

$$
\mathbb{E}\big[\hat{u}_N^*(X_1, X_2) \cdot (\hat{u}_N - \hat{u}_N^*)(X_3, X_4)\big] = \mathbb{E}\Big[\mathbb{E}\big[\hat{u}_N^*(X_1, X_2) \bigm| X_1, X_2\big] \cdot \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_3, X_4) \bigm| X_1, X_2\big]\Big],
$$

and $\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_3, X_4) \bigm| X_1, X_2\big] = 0$ (proof completely analogous to the proof above, that $\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2) \bigm| X_3, X_4\big] = 0$).

Thus, altogether for $\mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_3, X_4)\big]$ we have

$$
\begin{aligned}
\mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_3, X_4)\big] &= \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2) \cdot (\hat{u}_N - \hat{u}_N^*)(X_3, X_4)\big] \\
&\le \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)\big]^2 \\
&= O(a_N^{-3} \cdot N^{-2}).
\end{aligned}
$$

We now have bounds on all of the expectations in the summands (5.109) through (5.114), so that for the sum (5.108) we may write

$$
\begin{aligned}
&\mathbb{E}\Big[m^{-2} \cdot \sum_{1 \le i \ne j \le m} \hat{u}_N(X_i, X_j)\Big]^2 \\
&\quad = m^{-4} \cdot \Big[m(m-1) \cdot O(a_N^{-3} \cdot N^{-1}) \\
&\qquad + m(m-1) \cdot O(a_N^{-3} \cdot N^{-1}) \\
&\qquad + 2m(m-1)(m-2) \cdot O(a_N^{-3} \cdot N^{-\frac{3}{2}}) \\
&\qquad + m(m-1)(m-2) \cdot O(a_N^{-3} \cdot N^{-\frac{3}{2}})
\end{aligned}
$$

$$+ m(m-1)(m-2)\cdot O(a_N^{-3}\cdot N^{-\frac{3}{2}})$$

$$+ m(m-1)(m-2)(m-3)\cdot O(a_N^{-3}\cdot N^{-2})\Big]$$

$$= O(a_N^{-3}\cdot N^{-2}). \tag{5.120}$$

making (5.108) $O_P(a_N^{-\frac{3}{2}}\cdot N^{-1})$ as claimed. □

LEMMA 5.19.

$$a_N^{-3}\cdot m^{-2}\cdot\sum_{1\le i\ne j\le m}\Big[K''\big(a_N^{-1}(H_N(X_i)-H_N(X_j))\big)\cdot\big(\hat{H}_N(X_i)-H_N(X_i)\big)$$

$$-\int K''\big(a_N^{-1}(H_N(x)-H_N(X_j))\big)\cdot\big(\hat{H}_N(x)-H_N(x)\big)\,F(dx)\Big]\cdot\big(H_N(X_j)-\hat{H}_N(X_j)\big)$$

$$= O_P(a_N^{-\frac{5}{2}}\cdot N^{-\frac{5}{4}}).$$

PROOF. Begin by defining

$$\hat{u}_N(r,s) = a_N^{-3}\cdot\Big[K''\big(a_N^{-1}(H_N(r)-H_N(s))\big)\cdot\big(\hat{H}_N(r)-H_N(r)\big)$$

$$-\int K''\big(a_N^{-1}(H_N(x)-H_N(s))\big)\cdot\big(\hat{H}_N(x)-H_N(x)\big)\,F(dx)\Big]\cdot\big(H_N(s)-\hat{H}_N(s)\big).$$

Then we may write (5.105) as

$$m^{-2}\cdot\sum_{1\le i\ne j\le m}\hat{u}_N(X_i,X_j). \tag{5.121}$$

Looking at the second moment of (5.121) to derive an upper bound we can use the standard expansion to write

$$\mathbb{E}\Big[m^{-2}\cdot\sum_{1\le i\ne j\le m}\hat{u}_N(X_i,X_j)\Big]^2$$

$$= m^{-4}\cdot\Big[m(m-1)\cdot\mathbb{E}\big[\hat{u}_N(X_1,X_2)\big]^2 \tag{5.122}$$

$$+ m(m-1)\cdot\mathbb{E}\big[\hat{u}_N(X_1,X_2)\cdot\hat{u}_N(X_2,X_1)\big] \tag{5.123}$$

$$+ 2m(m-1)(m-2)\cdot\mathbb{E}\big[\hat{u}_N(X_1,X_2)\cdot\hat{u}_N(X_3,X_1)\big] \tag{5.124}$$

$$+ m(m-1)(m-2)\cdot\mathbb{E}\big[\hat{u}_N(X_1,X_2)\cdot\hat{u}_N(X_1,X_3)\big] \tag{5.125}$$

$$+ m(m-1)(m-2)\cdot\mathbb{E}\big[\hat{u}_N(X_1,X_2)\cdot\hat{u}_N(X_3,X_2)\big] \tag{5.126}$$

$$+ m(m-1)(m-2)(m-3)\cdot\mathbb{E}\big[\hat{u}_N(X_1,X_2)\cdot\hat{u}_N(X_3,X_4)\big]\Big]. \tag{5.127}$$

We construct a simple bound for $\|\hat{u}_N\|$, which will prove useful in deriving adequate bounds for the expectations (5.122) through (5.126). For all r and s we have

$$\big|\hat{u}_N(r,s)\big| = \Big|a_N^{-3}\cdot\Big[K''\big(a_N^{-1}(H_N(r)-H_N(s))\big)\cdot\big(\hat{H}_N(r)-H_N(r)\big)$$

$$-\int K''\big(a_N^{-1}(H_N(x)-H_N(s))\big)\cdot\big(\hat{H}_N(x)-H_N(x)\big)\,F(dx)\Big]\cdot\big(H_N(s)-\hat{H}_N(s)\big)\Big|$$

$$
\begin{aligned}
&= a_N^{-3} \cdot \Big| K''\big(a_N^{-1}(H_N(r) - H_N(s))\big) \cdot \big(\hat{H}_N(r) - H_N(r)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(s))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \Big| \cdot \big| H_N(s) - \hat{H}_N(s) \big| \\
&\le a_N^{-3} \cdot \Big[\Big| K''\big(a_N^{-1}(H_N(r) - H_N(s))\big) \cdot \big(\hat{H}_N(r) - H_N(r)\big) \Big| \\
&\qquad + \Big| \int K''\big(a_N^{-1}(H_N(x) - H_N(s))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \Big| \Big] \cdot \big\| H_N - \hat{H}_N \big\| \\
&\le a_N^{-3} \cdot 2 \big\| K'' \big\| \cdot \big\| \hat{H}_N - H_N \big\| \cdot \big\| H_N - \hat{H}_N \big\|,
\end{aligned}
$$

so that

$$
\big\| \hat{u}_N \big\| \le 2 \big\| K'' \big\| \cdot a_N^{-3} \cdot \big\| \hat{H}_N - H_N \big\|^2. \tag{5.128}
$$

Using (5.128) we obtain for the expectation in (5.122)

$$
\begin{aligned}
\mathbb{E}\big[\hat{u}_N(X_1, X_2)\big]^2 &\le 4 \big\| K'' \big\|^2 \cdot a_N^{-6} \cdot \mathbb{E}\big[\big\| \hat{H}_N - H_N \big\|^2\big]^2 \\
&= O(a_N^{-6} \cdot N^{-2}).
\end{aligned} \tag{5.129}
$$

Also, since the expectations in the following four summands (5.123) through (5.126) are bound by the expectation $\mathbb{E}\big[\hat{u}_N(X_1, X_2)\big]^2$ due to the Cauchy-inequality, these are all of the order $O(a_N^{-6} \cdot N^{-2})$ as well.

Thus, it remains only to bound the last expectation in (5.127). In order to derive bounds for this expectation we again define $\hat{H}_N^*$ to be equal to $\hat{H}_N$ with X_1, X_2, X_3 and X_4 removed from the sample. That is,

$$
\hat{H}_N^*(x) = N^{-1} \cdot \Big[\sum_{i=5}^{m} 1_{\{X_i \le x\}} + \sum_{k=1}^{n} 1_{\{Y_k \le x\}} \Big].
$$

Also, as in the proof of the previous lemma, define $\hat{u}_N^*$ as $\hat{u}_N$ with all occurrences of $\hat{H}_N$ replaced by $\hat{H}_N^*$.

To bound the expectation $\mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_3, X_4)\big]$ we will again use the inequality (5.116), which in this case gives us

$$
\begin{aligned}
&\big| \mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_3, X_4)\big] \big| \\
&\quad \le \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)\big]^2 + 2 \Big[\mathbb{E}\big[\hat{u}_N^*(X_1, X_2)\big]^2 \cdot \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)\big]^2 \Big]^{\frac{1}{2}} \\
&\qquad + \big| \mathbb{E}\big[\hat{u}_N^*(X_1, X_2) \cdot \hat{u}_N^*(X_3, X_4)\big] \big|.
\end{aligned}
$$

For the last expectation on the right hand side we have

$$
\mathbb{E}\big[\hat{u}_N^*(X_1, X_2) \cdot \hat{u}_N^*(X_3, X_4)\big] = \mathbb{E}\Big[\mathbb{E}\big[\hat{u}_N^*(X_1, X_2) \mid X_2, \ldots, X_m,\ Y_1, \ldots, Y_n\big] \cdot \hat{u}_N^*(X_3, X_4) \Big] = 0,
$$

since for the inner expectation

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_N^*(X_1, X_2) \mid X_2, \ldots, X_m,\ Y_1, \ldots, Y_n\big] \\
&\quad = \mathbb{E}\Big[a_N^{-3} \cdot \Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx) \Big]
\end{aligned}
$$

$$
\begin{aligned}
&\quad \times \big(H_N(X_2) - \hat{H}_N^*(X_2)\big) \,\bigg|\, X_2, \dots, X_m,\ Y_1, \dots, Y_n\bigg] \\
&= a_N^{-3} \cdot \bigg[\int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\bigg] \cdot \big(H_N(X_2) - \hat{H}_N^*(X_2)\big) \\
&= 0.
\end{aligned}
$$

Thus, using inequality (5.116) it remains only to bound the expectations $\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)\big]^2$ and $\mathbb{E}\big[\hat{u}_N^*(X_1, X_2)\big]^2$ in order to derive a bound for the summand (5.127). Now for the expression $(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)$ we have

$$
\begin{aligned}
&(\hat{u}_N - \hat{u}_N^*)(X_1, X_2) \\
&= a_N^{-3} \cdot \bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N(X_1) - H_N(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\bigg] \cdot \big(H_N(X_2) - \hat{H}_N(X_2)\big) \\
&\quad - a_N^{-3} \cdot \bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\bigg] \cdot \big(H_N(X_2) - \hat{H}_N^*(X_2)\big) \\
&= a_N^{-3} \cdot \bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \Big[\big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big) + \big(\hat{H}_N^*(X_1) - H_N(X_1)\big)\Big] \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \Big[\big(\hat{H}_N(x) - \hat{H}_N^*(x)\big) + \big(\hat{H}_N^*(x) - H_N(x)\big)\Big]\, F(dx)\bigg] \\
&\quad \times \Big[\big(H_N(X_2) - \hat{H}_N^*(X_2)\big) + \big(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)\big)\Big] \\
&\quad - a_N^{-3} \cdot \bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\bigg] \cdot \big(H_N(X_2) - \hat{H}_N^*(X_2)\big) \\
&= a_N^{-3} \cdot \bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx) \\
&\quad + K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\bigg] \\
&\quad \times \Big[\big(H_N(X_2) - \hat{H}_N^*(X_2)\big) + \big(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)\big)\Big] \\
&\quad - a_N^{-3} \cdot \bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\bigg] \cdot \big(H_N(X_2) - \hat{H}_N^*(X_2)\big)
\end{aligned}
$$

$$
\begin{aligned}
&= a_N^{-3} \cdot \Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big] \cdot \big(H_N(X_2) - \hat{H}_N^*(X_2)\big) \\
&\quad + a_N^{-3} \cdot \Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big] \cdot \big(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)\big) \\
&\quad + a_N^{-3} \cdot \Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big] \cdot \big(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)\big),
\end{aligned}
$$

so that for the expectation we have

$$
\begin{aligned}
&\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)\big]^2 \\
&= \mathbb{E}\Big[a_N^{-3} \cdot \Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big] \cdot \big(H_N(X_2) - \hat{H}_N^*(X_2)\big) \\
&\quad + a_N^{-3} \cdot \Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big] \cdot \big(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)\big) \\
&\quad + a_N^{-3} \cdot \Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big] \cdot \big(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)\big)\Big]^2 \\
&\le 4a_N^{-6} \cdot \Big[\mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big] \cdot \big(H_N(X_2) - \hat{H}_N^*(X_2)\big)\Big]^2 \\
&\quad + \mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big] \cdot \big(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)\big)\Big]^2 \\
&\quad + \mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big] \cdot \big(\hat{H}_N^*(X_2) - \hat{H}_N(X_2)\big)\Big]^2\Big] \\
&\le 4a_N^{-6} \cdot \Big[2 \cdot \mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big)^2 \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big)^2 \\
&\qquad + \Big[\int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big]^2\Big] \cdot \big(H_N(X_2) - \hat{H}_N^*(X_2)\big)^2\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+ 2\cdot\mathbb{E}\Bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2\cdot\big(\hat{H}_N(X_1)-\hat{H}_N^*(X_1)\big)^2\\
&\quad+\bigg[\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)\cdot\big(\hat{H}_N(x)-\hat{H}_N^*(x)\big)\,F(dx)\bigg]^2\bigg]\cdot\big(\hat{H}_N^*(X_2)-\hat{H}_N(X_2)\big)^2\Bigg]\\
&+ 2\cdot\mathbb{E}\Bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)^2\\
&\quad+\bigg[\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\bigg]^2\bigg]\cdot\big(\hat{H}_N^*(X_2)-\hat{H}_N(X_2)\big)^2\Bigg]\Bigg]\\
\le{}& 4a_N^{-6}\cdot\Bigg[2\cdot\mathbb{E}\Bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2\cdot 16N^{-2}\\
&\quad+\bigg[\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)\cdot\big(\hat{H}_N(x)-\hat{H}_N^*(x)\big)\,F(dx)\bigg]^2\bigg]\cdot\big\|H_N-\hat{H}_N^*\big\|^2\Bigg]\\
&+ 2\cdot\mathbb{E}\Bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2\cdot 16N^{-2}\\
&\quad+\bigg[\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)\cdot\big(\hat{H}_N(x)-\hat{H}_N^*(x)\big)\,F(dx)\bigg]^2\bigg]\cdot 16N^{-2}\Bigg]\\
&+ 2\cdot\mathbb{E}\Bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2\cdot\big\|\hat{H}_N^*-H_N\big\|^2\\
&\quad+\bigg[\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\bigg]^2\bigg]\cdot 16N^{-2}\Bigg]\Bigg]\\
\le{}& 4a_N^{-6}\cdot\Bigg[2\cdot\mathbb{E}\Bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2\cdot 16N^{-2}\\
&\quad+\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)^2\cdot\big(\hat{H}_N(x)-\hat{H}_N^*(x)\big)^2\,F(dx)\bigg]\cdot\big\|H_N-\hat{H}_N^*\big\|^2\Bigg]\\
&+ 2\cdot\mathbb{E}\Bigg[\bigg[\big\|K''\big\|^2\cdot 16N^{-2}\\
&\quad+\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)^2\cdot\big(\hat{H}_N(x)-\hat{H}_N^*(x)\big)^2\,F(dx)\bigg]\cdot 16N^{-2}\Bigg]\\
&+ 2\cdot\mathbb{E}\Bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2\cdot\big\|\hat{H}_N^*-H_N\big\|^2\\
&\quad+\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)^2\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)^2\,F(dx)\bigg]\cdot 16N^{-2}\Bigg]\Bigg]\\
\le{}& 8a_N^{-6}\cdot\Bigg[\mathbb{E}\Bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2\\
&\quad+\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)^2\cdot\ F(dx)\bigg]\cdot 16N^{-2}\cdot\big\|H_N-\hat{H}_N^*\big\|^2\Bigg]\\
&+\mathbb{E}\Bigg[\bigg[\big\|K''\big\|^2\cdot 16N^{-2}+\big\|K''\big\|^2\cdot 16N^{-2}\bigg]\cdot 16N^{-2}\Bigg]\\
&+\mathbb{E}\Bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2\\
&\quad+\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)^2\,F(dx)\bigg]\cdot\big\|\hat{H}_N^*-H_N\big\|^2\cdot 16N^{-2}\Bigg]\Bigg]\\
={}& 8a_N^{-6}\cdot\Bigg[2\cdot\mathbb{E}\Bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2
\end{aligned}
$$

$$+\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)^2\cdot\ F(dx)\bigg]\cdot 16N^{-2}\cdot\big\|H_N-\hat{H}_N^*\big\|^2\bigg]$$

$$+2\big\|K''\big\|^2\cdot 16^2N^{-4}\bigg]$$

$$=16^2a_N^{-6}\cdot N^{-2}\cdot\mathbb{E}\bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2$$

$$+\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)^2\cdot\ F(dx)\bigg]\cdot\big\|H_N-\hat{H}_N^*\big\|^2\bigg]$$

$$+16^3\big\|K''\big\|^2\cdot a_N^{-6}\cdot N^{-4}$$

$$=16^2a_N^{-6}\cdot N^{-2}\cdot\bigg[\mathbb{E}\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2$$

$$+\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)^2\cdot\ F(dx)\bigg]\cdot\mathbb{E}\Big[\big\|H_N-\hat{H}_N^*\big\|^2\Big]\bigg]$$

$$+16^3\big\|K''\big\|^2\cdot a_N^{-6}\cdot N^{-4}$$

$$=16^2a_N^{-6}\cdot N^{-2}\cdot\Big[2\cdot\mathbb{E}\Big[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2\Big]\cdot\mathbb{E}\Big[\big\|H_N-\hat{H}_N^*\big\|^2\Big]\Big]$$

$$+16^3\big\|K''\big\|^2\cdot a_N^{-6}\cdot N^{-4}$$

$$\le 16^2a_N^{-6}\cdot N^{-2}\cdot 4\big\|K''\big\|^2\cdot a_N\Big(1+\frac{n}{m}\Big)\cdot\mathbb{E}\Big[\big\|H_N-\hat{H}_N^*\big\|^2\Big]$$

$$+16^3\big\|K''\big\|^2\cdot a_N^{-6}\cdot N^{-4},$$

where we have used (A.2) to obtain the final inequality.

Thus, altogether for $\mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,X_2)\big]^2$ we have

$$\begin{aligned}\mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,X_2)\big]^2 &\le 16^2a_N^{-6}\cdot N^{-2}\cdot 4\big\|K''\big\|^2\cdot a_N\Big(1+\frac{n}{m}\Big)\cdot\mathbb{E}\Big[\big\|H_N-\hat{H}_N^*\big\|^2\Big]\\ &\quad+16^3\big\|K''\big\|^2\cdot a_N^{-6}\cdot N^{-4}\\ &=O(a_N^{-5}\cdot N^{-3})+O(a_N^{-6}\cdot N^{-4})\\ &=O(a_N^{-5}\cdot N^{-3}).\end{aligned}\tag{5.130}$$

Now, for the second expectation in the inequality (5.116) we have

$$\mathbb{E}\big[\hat{u}_N^*(X_1,X_2)\big]^2$$

$$=\mathbb{E}\bigg[a_N^{-3}\cdot\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)$$

$$-\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\bigg]\cdot\big(H_N(X_2)-\hat{H}_N^*(X_2)\big)\bigg]^2$$

$$\le a_N^{-6}\cdot\mathbb{E}\bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)$$

$$-\int K''\big(a_N^{-1}(H_N(x)-H_N(X_2))\big)\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\bigg]^2\cdot\big\|H_N-\hat{H}_N^*\big\|^2\bigg]$$

$$\le a_N^{-6}\cdot\mathbb{E}\bigg[2\cdot\bigg[K''\big(a_N^{-1}(H_N(X_1)-H_N(X_2))\big)^2\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)^2$$

$$+ \left[\int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\right]^2\Bigg] \cdot \big\|H_N - \hat{H}_N^*\big\|^2\Bigg]$$

$$\le a_N^{-6} \cdot \mathbb{E}\Bigg[2 \cdot \bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big)^2 \cdot \big\|\hat{H}_N^* - H_N\big\|^2$$

$$+ \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big)^2 \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)^2\, F(dx)\bigg] \cdot \big\|H_N - \hat{H}_N^*\big\|^2\Bigg]$$

$$\le 2a_N^{-6} \cdot \mathbb{E}\Bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big)^2$$

$$+ \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big)^2\, F(dx)\bigg] \cdot \big\|H_N - \hat{H}_N^*\big\|^4\Bigg]$$

$$= 2a_N^{-6} \cdot \mathbb{E}\bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big)^2$$

$$+ \int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big)^2\, F(dx)\bigg] \cdot \mathbb{E}\Big[\big\|H_N - \hat{H}_N^*\big\|^4\Big]$$

$$= 4a_N^{-6} \cdot \mathbb{E}\bigg[\int K''\big(a_N^{-1}(H_N(x) - H_N(X_2))\big)^2\, F(dx)\bigg] \cdot \mathbb{E}\Big[\big\|H_N - \hat{H}_N^*\big\|^4\Big]$$

$$\le 4a_N^{-6} \cdot 2\big\|K''\big\|^2 \cdot a_N\Big(1 + \frac{n}{m}\Big) \cdot \mathbb{E}\Big[\big\|H_N - \hat{H}_N^*\big\|^4\Big],$$

where we once again use (A.2) to obtain the last inequality.

Thus, altogether for $\mathbb{E}\big[\hat{u}_N^*(X_1, X_2)\big]^2$ we have

$$\begin{aligned}\mathbb{E}\big[\hat{u}_N^*(X_1, X_2)\big]^2 &\le 8\big\|K''\big\|^2 \cdot a_N^{-5}\Big(1 + \frac{n}{m}\Big) \cdot \mathbb{E}\Big[\big\|H_N - \hat{H}_N^*\big\|^4\Big] \\ &= O(a_N^{-5}) \cdot O(N^{-2}) \\ &= O(a_N^{-5} \cdot N^{-2}). \end{aligned} \tag{5.131}$$

Combining (5.130) and (5.131) gives us a bound for the expectation in the last summand (5.127), namely

$$\begin{aligned}&\big|\mathbb{E}\big[\hat{u}_N(X_1, X_2) \cdot \hat{u}_N(X_3, X_4)\big]\big| \\ &\quad\le \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)\big]^2 + 2\Big[\mathbb{E}\big[\hat{u}_N^*(X_1, X_2)\big]^2 \cdot \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)\big]^2\Big]^{\frac{1}{2}} \\ &\quad= O(a_N^{-5} \cdot N^{-3}) + 2\Big[O(a_N^{-5} \cdot N^{-2}) \cdot O(a_N^{-5} \cdot N^{-3})\Big]^{\frac{1}{2}} \\ &\quad= O(a_N^{-5} \cdot N^{-3}) + O(a_N^{-5} \cdot N^{-\frac{5}{2}}) \\ &\quad= O(a_N^{-5} \cdot N^{-\frac{5}{2}}). \end{aligned} \tag{5.132}$$

We now have bounds on all of the expectations in the summands (5.122) through (5.127), so that for the sum (5.121) we may write

$$\begin{aligned}&\mathbb{E}\Big[m^{-2} \cdot \sum_{1 \le i \ne j \le m} \hat{u}_N(X_i, X_j)\Big]^2 \\ &\quad= m^{-4} \cdot \Big[m(m-1) \cdot O(a_N^{-6} \cdot N^{-2}) \\ &\qquad+ m(m-1) \cdot O(a_N^{-6} \cdot N^{-2}) \\ &\qquad+ 2m(m-1)(m-2) \cdot O(a_N^{-6} \cdot N^{-2})\end{aligned}$$

$$
\begin{aligned}
&+ m(m-1)(m-2) \cdot O(a_N^{-6} \cdot N^{-2}) \\
&+ m(m-1)(m-2) \cdot O(a_N^{-6} \cdot N^{-2}) \\
&+ m(m-1)(m-2)(m-3) \cdot O(a_N^{-5} \cdot N^{-\frac{5}{2}})\Big] \\
= {} & O(a_N^{-5} \cdot N^{-\frac{5}{2}}).
\end{aligned} \tag{5.133}
$$

making (5.121) $O_P(a_N^{-\frac{5}{2}} \cdot N^{-\frac{5}{4}})$ as claimed. □

LEMMA 5.20.

$$
\begin{aligned}
&\frac{1}{2} a_N^{-4} \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \Big[K'''(\xi_{ij}) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'''(\tau_j) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \Big] \cdot \big(H_N(X_j) - \hat{H}_N(X_j)\big)^2 \\
&= O_P(a_N^{-4} \cdot N^{-\frac{3}{2}}).
\end{aligned}
$$

PROOF.

$$
\begin{aligned}
&\frac{1}{2} a_N^{-4} \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \Big[K'''(\xi_{ij}) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'''(\tau_j) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \Big] \cdot \big(H_N(X_j) - \hat{H}_N(X_j)\big)^2 \\
&\le \frac{1}{2} a_N^{-4} \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \Big[\big\|K'''\big\| \cdot \big\|\hat{H}_N - H_N\big\| + \int \big\|K'''\big\| \cdot \big\|\hat{H}_N - H_N\big\|\, F(dx) \Big] \cdot \big\|H_N - \hat{H}_N\big\|^2 \\
&= \frac{1}{2} a_N^{-4} \cdot \frac{m(m-1)}{m^2} \cdot 2\big\|K'''\big\| \cdot \big\|H_N - \hat{H}_N\big\|^3 \\
&= \big\|K'''\big\| \cdot a_N^{-4} \cdot \frac{m(m-1)}{m^2} \cdot \big\|H_N - \hat{H}_N\big\|^3 \\
&= O_P(a_N^{-4} \cdot N^{-\frac{3}{2}}).
\end{aligned}
$$

□

To bound (5.100) we will use very similar arguments to those which we used to show that (5.99) is $O_P(a_N^{-\frac{3}{2}} \cdot N^{-1})$. We begin by deriving a sum representation of (5.100).

$$
\begin{aligned}
&\Big[\hat{g}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big] \Big[\hat{F}_m(dx) - F(dx)\Big] \\
&= m^{-1} \cdot \sum_{i=1}^{m} \Bigg[\Bigg[a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} K'\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(Y_k))\big) \\
&\qquad - a_N^{-2} \int K'\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, G(dy) \Bigg] \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\quad - \Bigg[a_N^{-2} \cdot n^{-1} \int \sum_{k=1}^{m} K'\big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\qquad - a_N^{-2} \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \Bigg] \Bigg]
\end{aligned}
$$

$$
\begin{aligned}
&= a_N^{-2} \cdot m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} \Bigg[\bigg[K'\big(a_N^{-1}(H_N(X_i) - \hat{H}_N(Y_k))\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, G(dy) \bigg] \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \Bigg]
\end{aligned}
$$

Now using the Taylor expansion of K' about each of the $a_N^{-1}(H_N(X_i) - H_N(Y_k))$ and $a_N^{-1}(H_N(x) - H_N(Y_k))$ then yields

$$
\begin{aligned}
&\Big[\hat{g}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big] \Big[\hat{F}_m(dx) - F(dx)\Big] \\
&= a_N^{-2} \cdot m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} \Big[K'\big(a_N^{-1}(H_N(X_i) - H_N(Y_k))\big) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, G(dy) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \Big] && (5.134) \\
&\quad + a_N^{-3} \cdot m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} \Big[K''\big(a_N^{-1}(H_N(X_i) - H_N(Y_k))\big) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \Big] \cdot \big(H_N(Y_k) - \hat{H}_N(Y_k)\big) && (5.135) \\
&\quad + \frac{1}{2} a_N^{-4} \cdot m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} \Big[K'''(\xi_{ij}) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'''(\tau_j) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \Big] \cdot \big(H_N(Y_k) - \hat{H}_N(Y_k)\big)^2 && (5.136)
\end{aligned}
$$

where ξ_{ij} and τ_j are appropriate values between the two ratios.

Lemma 5.21.

$$
\begin{aligned}
&a_N^{-2} \cdot m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} \Big[K'\big(a_N^{-1}(H_N(X_i) - H_N(Y_k))\big) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, G(dy) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \Big] \\
&= O_P(a_N^{-\frac{3}{2}} \cdot N^{-1}).
\end{aligned}
$$

PROOF. Begin by defining

$$\begin{aligned}\hat{u}_N(r,s) = a_N^{-2} \cdot \Big[&K'\big(a_N^{-1}(H_N(r) - H_N(s))\big) \cdot \big(\hat{H}_N(r) - H_N(r)\big)\\ &- \int K'\big(a_N^{-1}(H_N(r) - H_N(y))\big)\, G(dy) \cdot \big(\hat{H}_N(r) - H_N(r)\big)\\ &- \int K'\big(a_N^{-1}(H_N(x) - H_N(s))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\\ &+ \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Big].\end{aligned}$$

Then we may write (5.134) as

$$m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \hat{u}_N(X_i, Y_k). \tag{5.137}$$

Looking at the second moment of (5.138) we can use the standard expansion to write

$$\mathbb{E}\Big[m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \hat{u}_N(X_i, Y_k)\Big]^2 \tag{5.138}$$

$$= m^{-2}n^{-2} \cdot \Big[mn \cdot \mathbb{E}\big[\hat{u}_N(X_1, Y_1)\big]^2 \tag{5.139}$$

$$+ m(m-1)n \cdot \mathbb{E}\big[\hat{u}_N(X_1, Y_1) \cdot \hat{u}_N(X_2, Y_1)\big] \tag{5.140}$$

$$+ mn(n-1) \cdot \mathbb{E}\big[\hat{u}_N(X_1, Y_1) \cdot \hat{u}_N(X_1, Y_2)\big] \tag{5.141}$$

$$+ m(m-1)n(n-1) \cdot \mathbb{E}\big[\hat{u}_N(X_1, Y_1) \cdot \hat{u}_N(X_2, Y_2)\big]\Big]. \tag{5.142}$$

To bound the expectation in (5.139) note that

$$\big\|\hat{u}_N\big\| \le 4\big\|K'\big\| \cdot a_N^{-2} \cdot \big\|\hat{H}_N - H_N\big\|,$$

so that

$$\begin{aligned}\mathbb{E}\big[\hat{u}_N(X_1, Y_1)\big]^2 &\le \mathbb{E}\big[\big\|\hat{u}_N\big\|\big]^2\\ &\le 16\big\|K'\big\|^2 \cdot a_N^{-4} \cdot \mathbb{E}\big[\big\|\hat{H}_N - H_N\big\|^2\big]\\ &= O(a_N^{-4} \cdot N^{-1}).\end{aligned} \tag{5.143}$$

In order to derive bounds for the expectations in (5.140) through (5.142) we define $\hat{H}_N^*$ to be equal to $\hat{H}_N$ with X_1, X_2, Y_1 and Y_2 removed from the sample. That is,

$$\hat{H}_N^*(x) = N^{-1} \cdot \Big[\sum_{i=3}^{m} 1_{\{X_i \le x\}} + \sum_{k=3}^{n} 1_{\{Y_k \le x\}}\Big].$$

Also, define $\hat{u}_N^*$ as $\hat{u}_N$ with all occurrences of $\hat{H}_N$ replaced by $\hat{H}_N^*$, and recall that for any X_{i_1}, X_{i_2}, Y_{k_1} and Y_{k_2} from the sample, we have

$$\begin{aligned}&\mathbb{E}\big[\hat{u}_N(X_{i_1}, Y_{k_1}) \cdot \hat{u}_N(X_{i_2}, Y_{k_2})\big]\\ &= \mathbb{E}\big[\big((\hat{u}_N - \hat{u}_N^*)(X_{i_1}, Y_{k_1}) + \hat{u}_N^*(X_{i_1}, Y_{k_1})\big) \cdot \big((\hat{u}_N - \hat{u}_N^*)(X_{i_2}, Y_{k_2}) + \hat{u}_N^*(X_{i_2}, Y_{k_2})\big)\big]\\ &= \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_{i_1}, Y_{k_1}) \cdot (\hat{u}_N - \hat{u}_N^*)(X_{i_2}, Y_{k_2})\big] + \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_{i_1}, Y_{k_1}) \cdot \hat{u}_N^*(X_{i_2}, Y_{k_2})\big]\end{aligned}$$

$$+ \mathbb{E}\big[\hat{u}_N^*(X_{i_1}, Y_{k_1}) \cdot (\hat{u}_N - \hat{u}_N^*)(X_{i_2}, Y_{k_2})\big] + \mathbb{E}\big[\hat{u}_N^*(X_{i_1}, Y_{k_1}) \cdot \hat{u}_N^*(X_{i_2}, Y_{k_2})\big], \tag{5.144}$$

and the Cauchy-inequality can be applied to the expectations on the right to get the bound

$$\begin{aligned}
&\big|\mathbb{E}\big[\hat{u}_N(X_{i_1}, Y_{k_1}) \cdot \hat{u}_N(X_{i_2}, Y_{k_2})\big]\big| \\
&\quad \le \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_{i_1}, Y_{k_1})\big]^2 + 2\Big[\mathbb{E}\big[\hat{u}_N^*(X_{i_1}, Y_{k_1})\big]^2 \cdot \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_{i_1}, Y_{k_1})\big]^2\Big]^{\frac{1}{2}} \\
&\qquad + \big|\mathbb{E}\big[\hat{u}_N^*(X_{i_1}, Y_{k_1}) \cdot \hat{u}_N^*(X_{i_2}, Y_{k_2})\big]\big|
\end{aligned} \tag{5.145}$$

In the following, we will use the equation (5.144) and the inequality (5.145) to bound the expectations in (5.140) through (5.142). We begin by applying (5.145) to the expectations in (5.140) and (5.141). In each of these cases, the last expectation on the right hand side of (5.145) vanishes. Beginning with (5.140) we see

$$\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1) \cdot \hat{u}_N^*(X_2, Y_1)\big] = \mathbb{E}\Big[\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1) \mid X_2, \ldots, X_m,\ Y_1, \ldots, Y_n\big] \cdot \hat{u}_N^*(X_2, Y_1)\Big] = 0,$$

since for the inner expectation

$$\begin{aligned}
&\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1) \mid X_2, \ldots, X_m,\ Y_1, \ldots, Y_n\big] \\
&= \mathbb{E}\Big[a_N^{-2} \cdot \Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, G(dy) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx) \\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big] \,\Big|\, X_2, \ldots, X_m,\ Y_1, \ldots, Y_n\Big] \\
&= a_N^{-2} \cdot \Big[\int K'\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx) \\
&\qquad - \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx) \\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big] \\
&= 0.
\end{aligned}$$

Further, in the case of (5.141) we have

$$\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1) \cdot \hat{u}_N^*(X_1, Y_2)\big] = \mathbb{E}\Big[\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1) \mid X_1, \ldots, X_m,\ Y_2, \ldots, Y_n\big] \cdot \hat{u}_N^*(X_1, Y_2)\Big] = 0,$$

since for the inner expectation

$$\begin{aligned}
&\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1) \mid X_1, \ldots, X_m,\ Y_2, \ldots, Y_n\big] \\
&= \mathbb{E}\Big[a_N^{-2} \cdot \Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, G(dy) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big)
\end{aligned}$$

$$
\begin{aligned}
&\qquad - \int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)\cdot\big(\hat H_N^*(x)-H_N(x)\big)\,F(dx)\\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,G(dy)\big(\hat H_N^*(x)-H_N(x)\big)\,F(dx)\Big]\;\Big|\;X_1,\dots,X_m,\;Y_2,\dots,Y_n\Big]\\
&= a_N^{-2}\cdot\Big[\int K'\big(a_N^{-1}(H_N(X_1)-H_N(z))\big)\cdot\big(\hat H_N^*(X_1)-H_N(X_1)\big)\,G(dz)\\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,G(dy)\cdot\big(\hat H_N^*(X_1)-H_N(X_1)\big)\\
&\qquad - \iint K'\big(a_N^{-1}(H_N(x)-H_N(z))\big)\cdot\big(\hat H_N^*(x)-H_N(x)\big)\,F(dx)\,G(dz)\\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,G(dy)\big(\hat H_N^*(x)-H_N(x)\big)\,F(dx)\Big]\\
&= 0.
\end{aligned}
$$

Thus, using inequality (5.145) it remains only to bound the expectations $\mathbb{E}\big[(\hat u_N-\hat u_N^*)(X_1,Y_1)\big]^2$ and $\mathbb{E}\big[\hat u_N^*(X_1,Y_1)\big]^2$ in order to derive bounds for the summands (5.140) and (5.141). For the first of these expectations note first that

$$
\begin{aligned}
&\hat H_N(x)-\hat H_N^*(x)\\
&\quad = N^{-1}\cdot\Big[\sum_{i=1}^m 1_{\{X_i\le x\}}+\sum_{k=1}^n 1_{\{Y_k\le x\}}\Big]-N^{-1}\cdot\Big[\sum_{i=3}^m 1_{\{X_i\le x\}}+\sum_{k=3}^n 1_{\{Y_k\le x\}}\Big]\\
&\quad = N^{-1}\cdot\Big[\sum_{i=1}^2 1_{\{X_i\le x\}}+\sum_{k=1}^2 1_{\{Y_k\le x\}}\Big],
\end{aligned}
$$

so that

$$
\begin{aligned}
&\mathbb{E}\big[(\hat u_N-\hat u_N^*)(X_1,Y_1)\big]^2\\
&\quad = \mathbb{E}\Big[a_N^{-2}\cdot\Big[K'\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)\cdot N^{-1}\cdot\Big[\sum_{i=1}^2 1_{\{X_i\le X_1\}}+\sum_{k=1}^2 1_{\{Y_k\le X_1\}}\Big]\\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,G(dy)\cdot N^{-1}\cdot\Big[\sum_{i=1}^2 1_{\{X_i\le X_1\}}+\sum_{k=1}^2 1_{\{Y_k\le X_1\}}\Big]\\
&\qquad - \int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)\cdot N^{-1}\cdot\Big[\sum_{i=1}^2 1_{\{X_i\le x\}}+\sum_{k=1}^2 1_{\{Y_k\le x\}}\Big]\,F(dx)\\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,G(dy)\cdot N^{-1}\cdot\Big[\sum_{i=1}^2 1_{\{X_i\le x\}}+\sum_{k=1}^2 1_{\{Y_k\le x\}}\Big]\,F(dx)\Big]\Big]^2\\
&\quad = a_N^{-4}\cdot N^{-2}\cdot\mathbb{E}\Big[\sum_{i=1}^2\Big[K'\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)\cdot 1_{\{X_i\le X_1\}}\\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,G(dy)\cdot 1_{\{X_i\le X_1\}}\\
&\qquad - \int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)\cdot 1_{\{X_i\le x\}}\,F(dx)\\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,G(dy)\cdot 1_{\{X_i\le x\}}\,F(dx)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+\sum_{k=1}^{2}\Big[K'\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)\cdot 1_{\{Y_k\le X_1\}}\\
&\quad-\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,G(dy)\cdot 1_{\{Y_k\le X_1\}}\\
&\quad-\int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)\cdot 1_{\{Y_k\le x\}}\,F(dx)\\
&\quad+\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,G(dy)\cdot 1_{\{Y_k\le x\}}\,F(dx)\Big]\Big]^2\\
\le{}& 4a_N^{-4}\cdot N^{-2}\cdot\Big[\sum_{i=1}^{2}\mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)\cdot 1_{\{X_i\le X_1\}}\\
&-\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,G(dy)\cdot 1_{\{X_i\le X_1\}}\\
&-\int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)\cdot 1_{\{X_i\le x\}}\,F(dx)\\
&+\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,G(dy)\cdot 1_{\{X_i\le x\}}\,F(dx)\Big]^2\\
&+\sum_{k=1}^{2}\mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)\cdot 1_{\{Y_k\le X_1\}}\\
&\quad-\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,G(dy)\cdot 1_{\{Y_k\le X_1\}}\\
&\quad-\int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)\cdot 1_{\{Y_k\le x\}}\,F(dx)\\
&\quad+\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,G(dy)\cdot 1_{\{Y_k\le x\}}\,F(dx)\Big]^2\Big]\\
\le{}& 16a_N^{-4}\cdot N^{-2}\cdot\Big[\sum_{i=1}^{2}\mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)\cdot 1_{\{X_i\le X_1\}}\Big]^2\\
&+\mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,G(dy)\cdot 1_{\{X_i\le X_1\}}\Big]^2\\
&+\mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)\cdot 1_{\{X_i\le x\}}\,F(dx)\Big]^2\\
&+\mathbb{E}\Big[\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,G(dy)\cdot 1_{\{X_i\le x\}}\,F(dx)\Big]^2\\
&+\sum_{k=1}^{2}\mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)\cdot 1_{\{Y_k\le X_1\}}\Big]^2\\
&\quad+\mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,G(dy)\cdot 1_{\{Y_k\le X_1\}}\Big]^2\\
&\quad+\mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)\cdot 1_{\{Y_k\le x\}}\,F(dx)\Big]^2\\
&\quad+\mathbb{E}\Big[\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,G(dy)\cdot 1_{\{Y_k\le x\}}\,F(dx)\Big]^2\Big]
\end{aligned}
$$

$$
\begin{aligned}
&\le 16a_N^{-4}\cdot N^{-2}\cdot\Bigg[\sum_{i=1}^{2}\mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)^2\Big]\\
&\qquad+\mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)^2\,G(dy)\Big]\\
&\qquad+\mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)^2\,F(dx)\Big]\\
&\qquad+\mathbb{E}\Big[\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)^2\,G(dy)\,F(dx)\Big]\\
&\quad+\sum_{k=1}^{2}\mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)^2\Big]\\
&\qquad+\mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)^2\,G(dy)\Big]\\
&\qquad+\mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)^2\,F(dx)\Big]\\
&\qquad+\mathbb{E}\Big[\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)^2\,G(dy)\,F(dx)\Big]\Bigg]\\
&=16^2a_N^{-4}\cdot N^{-2}\cdot\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)^2\,G(dy)\,F(dx)\\
&\le 16^2a_N^{-4}\cdot N^{-2}\cdot 2\big\|K'\big\|^2\cdot a_N\Big(1+\frac{m}{n}\Big)\\
&=512\big\|K'\big\|^2a_N^{-3}\cdot N^{-2}\cdot\Big(1+\frac{m}{n}\Big),
\end{aligned}
$$

where we have used (A.2) to obtain the last inequality.

Thus for $\mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,Y_1)\big]^2$ we may write

$$
\mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,Y_1)\big]^2=O(a_N^{-3}\cdot N^{-2}). \tag{5.146}
$$

For the second of these expectations we have

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_N^*(X_1,Y_1)\big]^2\\
&\quad=\mathbb{E}\Bigg[a_N^{-2}\cdot\Bigg[K'\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)\\
&\qquad-\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,G(dy)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)\\
&\qquad-\int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\\
&\qquad+\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,G(dy)\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\Bigg]\Bigg]^2\\
&\quad\le 4a_N^{-4}\cdot\Bigg[\mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)\Big]^2\\
&\qquad+\mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,G(dy)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)\Big]^2\\
&\qquad+\mathbb{E}\Big[\int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\Big]^2
\end{aligned}
$$

$$
\begin{aligned}
&+ \mathbb{E}\Big[\iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big]^2 \Big] \\
&\le 4a_N^{-4} \cdot \Big[\mathbb{E}\Big[\int \Big(K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\Big)^2 G(dy) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big)^2 \Big] \\
&+ \mathbb{E}\Big[\int \Big(K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\Big)^2 G(dy) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big)^2 \Big] \\
&+ \mathbb{E}\Big[\int \Big(K'\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)\Big)^2 \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)^2 F(dx) \Big] \\
&+ \mathbb{E}\Big[\iint \Big(K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\Big)^2 G(dy) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)^2 F(dx) \Big]\Big] \\
&\le 4a_N^{-4} \cdot 2\|K'\|^2 \cdot a_N\Big(1 + \frac{m}{n}\Big) \cdot \Big[\mathbb{E}\Big[\hat{H}_N^*(X_1) - H_N(X_1)\Big]^2 \\
&+ \mathbb{E}\Big[\hat{H}_N^*(X_1) - H_N(X_1)\Big]^2 \\
&+ \mathbb{E}\Big[\big\|\hat{H}_N^* - H_N\big\|^2\Big] \\
&+ \mathbb{E}\Big[\big\|\hat{H}_N^* - H_N\big\|^2\Big]\Big] \\
&\le 32a_N^{-3} \cdot \big\|K'\big\|^2 \cdot \Big(1 + \frac{m}{n}\Big) \cdot \mathbb{E}\Big[\big\|\hat{H}_N^* - H_N\big\|^2\Big],
\end{aligned}
$$

once again using (A.2) to obtain the penultimate inequality.

Thus for $\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1)\big]^2$ we may write

$$
\begin{aligned}
\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1)\big]^2 &\le 32a_N^{-3} \cdot \big\|K'\big\|^2 \cdot \Big(1 + \frac{m}{n}\Big) \cdot \mathbb{E}\Big[\big\|\hat{H}_N^* - H_N\big\|^2\Big] \\
&= O(a_N^{-3} \cdot N^{-1}). \qquad (5.147)
\end{aligned}
$$

Using the bounds (5.146) and (5.147) in the inequality (5.145) we have shown for the expectations in (5.140) and (5.141) that these are all less than or equal to

$$
\begin{aligned}
&\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1)\big]^2 + 2\Big[\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1)\big]^2 \cdot \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1)\big]^2\Big]^{\frac{1}{2}} \\
&= O(a_N^{-3} \cdot N^{-2}) + \big[O(a_N^{-3} \cdot N^{-1}) \cdot O(a_N^{-3} \cdot N^{-2})\big]^{\frac{1}{2}} \\
&= O(a_N^{-3} \cdot N^{-2}) + O(a_N^{-3} \cdot N^{-\frac{3}{2}}) \\
&= O(a_N^{-3} \cdot N^{-\frac{3}{2}}). \qquad (5.148)
\end{aligned}
$$

Thus, it remains only to bound the expectation in the last summand (5.142). In this case, we will use the equation (5.144), which tells us that

$$
\begin{aligned}
&\mathbb{E}\big[\hat{u}_N(X_1, Y_1) \cdot \hat{u}_N(X_2, Y_2)\big] \\
&= \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1) \cdot (\hat{u}_N - \hat{u}_N^*)(X_2, Y_2)\big] + \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1) \cdot \hat{u}_N^*(X_2, Y_2)\big] \\
&\quad + \mathbb{E}\big[\hat{u}_N^*(X_1, Y_1) \cdot (\hat{u}_N - \hat{u}_N^*)(X_2, Y_2)\big] + \mathbb{E}\big[\hat{u}_N^*(X_1, Y_1) \cdot \hat{u}_N^*(X_2, Y_2)\big].
\end{aligned}
$$

The last expectation vanishes immediately:

$$
\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1) \cdot \hat{u}_N^*(X_2, Y_2)\big] = \mathbb{E}\Big[\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1) \mid X_2, \ldots, X_m,\ Y_1, \ldots, Y_n\big] \cdot \hat{u}_N^*(X_2, Y_2)\Big] = 0,
$$

as we have already shown above that the inner expectation is 0.

Now, unlike in the case of the other expectations in (5.140) and (5.141), we will find that in the case of $\mathbb{E}\big[\hat{u}_N^*(X_1,Y_1)\cdot\hat{u}_N^*(X_2,Y_2)\big]$ the expectations $\mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,Y_1)\cdot\hat{u}_N^*(X_2,Y_2)\big]$ and $\mathbb{E}\big[\hat{u}_N^*(X_1,Y_1)\cdot(\hat{u}_N-\hat{u}_N^*)(X_2,Y_2)\big]$ on the right hand side of (5.144) vanish as well.

Recall that $\hat{H}_N(x)-\hat{H}_N^*(x) = N^{-1}\cdot\big[\sum_{i=1}^2 1_{\{X_i\le x\}}+\sum_{k=1}^2 1_{\{Y_k\le x\}}\big]$. Then we can note that the expression $(\hat{u}_N-\hat{u}_N^*)(X_1,Y_1)$ actually depends only on X_1, X_2, Y_1 and Y_2 and none of the rest of the sample, since

$$
\begin{aligned}
&(\hat{u}_N-\hat{u}_N^*)(X_1,Y_1)\\
&= a_N^{-2}\cdot\Bigg[K'\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)\cdot N^{-1}\cdot\Bigg[\sum_{i=1}^2 1_{\{X_i\le X_1\}}+\sum_{k=1}^2 1_{\{Y_k\le X_1\}}\Bigg]\\
&\quad-\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,G(dy)\cdot N^{-1}\cdot\Bigg[\sum_{i=1}^2 1_{\{X_i\le X_1\}}+\sum_{k=1}^2 1_{\{Y_k\le X_1\}}\Bigg]\\
&\quad-\int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)\cdot N^{-1}\cdot\Bigg[\sum_{i=1}^2 1_{\{X_i\le x\}}+\sum_{k=1}^2 1_{\{Y_k\le x\}}\Bigg]\,F(dx)\\
&\quad+\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,G(dy)\cdot N^{-1}\cdot\Bigg[\sum_{i=1}^2 1_{\{X_i\le x\}}+\sum_{k=1}^2 1_{\{Y_k\le x\}}\Bigg]\,F(dx)\Bigg].
\end{aligned}
$$

$\hat{u}_N^*(X_2,Y_2)$, on the other hand, is equal to

$$
\begin{aligned}
\hat{u}_N^*(X_2,Y_2) = a_N^{-2}\cdot\Bigg[&K'\big(a_N^{-1}(H_N(X_2)-H_N(Y_2))\big)\cdot\big(\hat{H}_N^*(X_2)-H_N(X_2)\big)\\
&-\int K'\big(a_N^{-1}(H_N(X_2)-H_N(y))\big)\,G(dy)\cdot\big(\hat{H}_N^*(X_2)-H_N(X_2)\big)\\
&-\int K'\big(a_N^{-1}(H_N(x)-H_N(Y_2))\big)\cdot\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\\
&+\iint K'\big(a_N^{-1}(H_N(x)-H_N(y))\big)\,G(dy)\big(\hat{H}_N^*(x)-H_N(x)\big)\,F(dx)\Bigg]
\end{aligned}
$$

which depends on $X_2,\dots,X_m$ and $Y_2,\dots,Y_n$, but not on X_1 and Y_1, so that for the expectation we may write

$$
\mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,Y_1)\cdot\hat{u}_N^*(X_2,Y_2)\big] = \mathbb{E}\Big[\mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,Y_1)\;\big|\;X_2,Y_2\big]\cdot\mathbb{E}\big[\hat{u}_N^*(X_2,Y_2)\;\big|\;X_2,Y_2\big]\Big].
$$

In the following we show that the first inner expectation vanishes:

$$
\begin{aligned}
&\mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,Y_1)\;\big|\;X_2,Y_2\big]\\
&=\mathbb{E}\Bigg[a_N^{-2}\cdot\Bigg[K'\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)\cdot N^{-1}\cdot\Bigg[\sum_{i=1}^2 1_{\{X_i\le X_1\}}+\sum_{k=1}^2 1_{\{Y_k\le X_1\}}\Bigg]\\
&\quad-\int K'\big(a_N^{-1}(H_N(X_1)-H_N(y))\big)\,G(dy)\cdot N^{-1}\cdot\Bigg[\sum_{i=1}^2 1_{\{X_i\le X_1\}}+\sum_{k=1}^2 1_{\{Y_k\le X_1\}}\Bigg]\\
&\quad-\int K'\big(a_N^{-1}(H_N(x)-H_N(Y_1))\big)\cdot N^{-1}\cdot\Bigg[\sum_{i=1}^2 1_{\{X_i\le x\}}+\sum_{k=1}^2 1_{\{Y_k\le x\}}\Bigg]\,F(dx)
\end{aligned}
$$

$$
\begin{aligned}
&\quad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot N^{-1} \cdot \Big[\sum_{i=1}^{2} 1_{\{X_i \le x\}} + \sum_{k=1}^{2} 1_{\{Y_k \le x\}}\Big]\, F(dx)\Big] \;\Big|\; X_2, Y_2\Big] \\
&= a_N^{-2} \cdot N^{-1} \cdot \mathbb{E}\Big[\sum_{i=1}^{2} \Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot 1_{\{X_i \le X_1\}} \\
&\quad - \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, G(dy) \cdot 1_{\{X_i \le X_1\}} \\
&\quad - \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot 1_{\{X_i \le x\}}\, F(dx) \\
&\quad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot 1_{\{X_i \le x\}}\, F(dx) \\
&\quad + \sum_{k=1}^{2} \Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot 1_{\{Y_k \le X_1\}} \\
&\qquad - \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, G(dy) \cdot 1_{\{Y_k \le X_1\}} \\
&\qquad - \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot 1_{\{Y_k \le x\}}\, F(dx) \\
&\qquad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot 1_{\{Y_k \le x\}}\, F(dx)\Big] \;\Big|\; X_2, Y_2\Big] \\
&= a_N^{-2} \cdot N^{-1} \cdot \mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot 1_{\{X_1 \le X_1\}} \\
&\quad - \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, G(dy) \cdot 1_{\{X_1 \le X_1\}} \\
&\quad - \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot 1_{\{X_1 \le x\}}\, F(dx) \\
&\quad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot 1_{\{X_1 \le x\}}\, F(dx) \;\Big|\; X_2, Y_2\Big] \\
&\quad + a_N^{-2} \cdot N^{-1} \cdot \mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot 1_{\{X_2 \le X_1\}} \\
&\quad - \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, G(dy) \cdot 1_{\{X_2 \le X_1\}} \\
&\quad - \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot 1_{\{X_2 \le x\}}\, F(dx) \\
&\quad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot 1_{\{X_2 \le x\}}\, F(dx) \;\Big|\; X_2, Y_2\Big] \\
&\quad + a_N^{-2} \cdot N^{-1} \cdot \mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot 1_{\{Y_1 \le X_1\}} \\
&\quad - \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, G(dy) \cdot 1_{\{Y_1 \le X_1\}} \\
&\quad - \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot 1_{\{Y_1 \le x\}}\, F(dx) \\
&\quad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot 1_{\{Y_1 \le x\}}\, F(dx) \;\Big|\; X_2, Y_2\Big] \\
&\quad + a_N^{-2} \cdot N^{-1} \cdot \mathbb{E}\Big[K'\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot 1_{\{Y_2 \le X_1\}}
\end{aligned}
$$

$$
\begin{aligned}
&\quad - \int K'\big(a_N^{-1}(H_N(X_1) - H_N(y))\big)\, G(dy) \cdot 1_{\{Y_2 \le X_1\}} \\
&\quad - \int K'\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot 1_{\{Y_2 \le x\}}\, F(dx) \\
&\quad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot 1_{\{Y_2 \le x\}}\, F(dx) \;\Big|\; X_2, Y_2\Big] \\
&= a_N^{-2} \cdot N^{-1} \cdot \Big[\iint K'\big(a_N^{-1}(H_N(w) - H_N(y))\big)\, F(dw)\, G(dy) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(w) - H_N(y))\big)\, G(dy)\, F(dw) \\
&\quad - \iiint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big) \cdot 1_{\{w \le x\}}\, F(dx)\, F(dw)\, G(dy) \\
&\quad + \iiint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot 1_{\{w \le x\}}\, F(dx)\, F(dw)\Big] \\
&\quad + a_N^{-2} \cdot N^{-1} \cdot \iint \Big[K'\big(a_N^{-1}(H_N(w) - H_N(y))\big) \cdot 1_{\{X_2 \le w\}}\, F(dw)\, G(dy) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(w) - H_N(y))\big)\, G(dy) \cdot 1_{\{X_2 \le w\}}\, F(dw) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big) \cdot 1_{\{X_2 \le x\}}\, F(dx)\, G(dy) \\
&\quad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot 1_{\{X_2 \le x\}}\, F(dx)\Big] \\
&\quad + a_N^{-2} \cdot N^{-1} \cdot \Big[\iint K'\big(a_N^{-1}(H_N(w) - H_N(y))\big) \cdot 1_{\{y \le w\}}\, F(dw)\, G(dy) \\
&\quad - \iiint K'\big(a_N^{-1}(H_N(w) - H_N(y))\big)\, G(dy) \cdot 1_{\{z \le w\}}\, F(dw)\, G(dz) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big) \cdot 1_{\{y \le x\}}\, F(dx)G(dy) \\
&\quad + \iiint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot 1_{\{z \le x\}}\, F(dx)\, G(dz)\Big] \\
&\quad + a_N^{-2} \cdot N^{-1} \cdot \Big[\iint K'\big(a_N^{-1}(H_N(w) - H_N(y))\big) \cdot 1_{\{Y_2 \le w\}}\, F(dw)\, G(dy) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(w) - H_N(y))\big)\, G(dy) \cdot 1_{\{Y_2 \le w\}}\, F(dw) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big) \cdot 1_{\{Y_2 \le x\}}\, F(dx)\, G(dy) \\
&\quad + \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot 1_{\{Y_2 \le x\}}\, F(dx)\Big] \\
&= 0 + 0 + 0 + 0.
\end{aligned}
$$

In the same manner we can show that the expectation $\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1) \cdot (\hat{u}_N - \hat{u}_N^*)(X_2, Y_2)\big]$ vanishes as well, since

$$
\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1) \cdot (\hat{u}_N - \hat{u}_N^*)(X_2, Y_2)\big] = \mathbb{E}\Big[\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1) \;\big|\; X_1, Y_1\big] \cdot \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_2, Y_2) \;\big|\; X_1, Y_1\big]\Big],
$$

and $\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_2, Y_2) \bigm| X_1, Y_1\big] = 0$ (proof completely analogous to the proof above, that $\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1) \bigm| X_2, Y_2\big] = 0$).

Thus, altogether for $\mathbb{E}\big[\hat{u}_N(X_1, Y_1) \cdot \hat{u}_N(X_2, Y_2)\big]$ we have

$$\begin{aligned}
\mathbb{E}\big[\hat{u}_N(X_1, Y_1) \cdot \hat{u}_N(X_2, Y_2)\big] &= \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1) \cdot (\hat{u}_N - \hat{u}_N^*)(X_2, Y_2)\big] \\
&\le \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1)\big]^2 \\
&= O(a_N^{-3} \cdot N^{-2}).
\end{aligned} \tag{5.149}$$

Using (5.143), (5.148) and (5.149) we now have bounds on all of the summands (5.139) through (5.142) so that for the sum (5.138) we may write

$$\begin{aligned}
&\mathbb{E}\left[m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \hat{u}_N(X_i, Y_k)\right]^2 \\
&= m^{-2}n^{-2} \cdot \Big[mn \cdot O(a_N^{-4} \cdot N^{-1}) \\
&\qquad + m(m-1)n \cdot O(a_N^{-3} \cdot N^{-\frac{3}{2}}) \\
&\qquad + mn(n-1) \cdot O(a_N^{-3} \cdot N^{-\frac{3}{2}}) \\
&\qquad + m(m-1)n(n-1) \cdot O(a_N^{-3} \cdot N^{-2})\Big] \\
&= O(a_N^{-3} \cdot N^{-2}).
\end{aligned}$$

This gives us

$$m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \hat{u}_N(X_i, Y_k) = O_P(a_N^{-\frac{3}{2}} \cdot N^{-1}),$$

which completes the proof. $\square$

Lemma 5.22.

$$\begin{aligned}
&a_N^{-3} \cdot m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \Big[K''\big(a_N^{-1}(H_N(X_i) - H_N(Y_k))\big) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Big] \cdot \big(H_N(Y_k) - \hat{H}_N(Y_k)\big) \\
&= O_P(a_N^{-\frac{5}{2}} \cdot N^{-\frac{5}{4}}).
\end{aligned}$$

Proof. As in the proof of Lemma 5.19 begin by defining

$$\begin{aligned}
\hat{u}_N(r, s) = a_N^{-3} \cdot \Big[&K''\big(a_N^{-1}(H_N(r) - H_N(s))\big) \cdot \big(\hat{H}_N(r) - H_N(r)\big) \\
&- \int K''\big(a_N^{-1}(H_N(x) - H_N(s))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Big] \cdot \big(H_N(s) - \hat{H}_N(s)\big).
\end{aligned}$$

Then we may write (5.135) as

$$m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \hat{u}_N(X_i, Y_k) \tag{5.150}$$

Looking at the second moment of (5.150) we can use the standard expansion to write

$$\mathbb{E}\Big[m^{-1}n^{-1}\cdot\sum_{i=1}^{m}\sum_{k=1}^{n}\hat{u}_N(X_i,Y_k)\Big]^2 \tag{5.151}$$

$$= m^{-2}n^{-2}\cdot\Big[mn\cdot\mathbb{E}\big[\hat{u}_N(X_1,Y_1)\big]^2 \tag{5.152}$$

$$+ m(m-1)n\cdot\mathbb{E}\big[\hat{u}_N(X_1,Y_1)\cdot\hat{u}_N(X_2,Y_1)\big] \tag{5.153}$$

$$+ mn(n-1)\cdot\mathbb{E}\big[\hat{u}_N(X_1,Y_1)\cdot\hat{u}_N(X_1,Y_2)\big] \tag{5.154}$$

$$+ m(m-1)n(n-1)\cdot\mathbb{E}\big[\hat{u}_N(X_1,Y_1)\cdot\hat{u}_N(X_2,Y_2)\big]\Big]. \tag{5.155}$$

In (5.128) (see proof of Lemma 5.19) we already constructed a simple bound for $\|\hat{u}_N\|$, which we can again use to quickly derive adequate bounds for the expectations (5.152) through (5.154), namely

$$\big\|\hat{u}_N\big\| \le 2\big\|K''\big\|\cdot a_N^{-3}\cdot\big\|\hat{H}_N - H_N\big\|^2.$$

Using (5.128) we obtain for the expectation in (5.152)

$$\begin{aligned}\mathbb{E}\big[\hat{u}_N(X_1,Y_1)\big]^2 &\le 4\big\|K''\big\|^2\cdot a_N^{-6}\cdot\mathbb{E}\big[\big\|\hat{H}_N - H_N\big\|^2\big]^2\\ &= O(a_N^{-6}\cdot N^{-2}).\end{aligned} \tag{5.156}$$

Also, since the expectations in the following two summands (5.153) and (5.154) are bound by the expectation $\mathbb{E}\big[\hat{u}_N(X_1,Y_1)\big]^2$ due to the Cauchy-inequality, these are all of the order $O(a_N^{-6}\cdot N^{-2})$ as well.

Thus, it remains only to bound the last expectation in (5.155). In order to derive bounds for this expectation we again define $\hat{H}_N^*$ to be equal to $\hat{H}_N$ with X_1, X_2, Y_1 and Y_2 removed from the sample. That is,

$$\hat{H}_N^*(x) = N^{-1}\cdot\Big[\sum_{i=3}^{m}1_{\{X_i\le x\}} + \sum_{k=3}^{n}1_{\{Y_k\le x\}}\Big].$$

Also, as in the proof of the previous lemma, define $\hat{u}_N^*$ as $\hat{u}_N$ with all occurrences of $\hat{H}_N$ replaced by $\hat{H}_N^*$.

To bound the expectation $\mathbb{E}\big[\hat{u}_N(X_1,Y_1)\cdot\hat{u}_N(X_2,Y_2)\big]$ we will again use the inequality (5.145), which in this case gives us

$$\begin{aligned}&\big|\mathbb{E}\big[\hat{u}_N(X_1,Y_1)\cdot\hat{u}_N(X_2,Y_2)\big]\big|\\ &\quad\le \mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,Y_1)\big]^2 + 2\Big[\mathbb{E}\big[\hat{u}_N^*(X_1,Y_1)\big]^2\cdot\mathbb{E}\big[(\hat{u}_N-\hat{u}_N^*)(X_1,Y_1)\big]^2\Big]^{\frac{1}{2}}\\ &\qquad+\big|\mathbb{E}\big[\hat{u}_N^*(X_1,Y_1)\cdot\hat{u}_N^*(X_2,Y_2)\big]\big|\end{aligned}$$

For the last expectation on the right hand side we have

$$\mathbb{E}\big[\hat{u}_N^*(X_1,Y_1)\cdot\hat{u}_N^*(X_2,Y_2)\big] = \mathbb{E}\Big[\mathbb{E}\big[\hat{u}_N^*(X_1,Y_1)\;\big|\;X_2,\ldots,X_m,\;Y_1,\ldots,Y_n\big]\cdot\hat{u}_N^*(X_2,Y_2)\Big] = 0,$$

since for the inner expectation

$$\begin{aligned}&\mathbb{E}\big[\hat{u}_N^*(X_1,Y_1)\;\big|\;X_2,\ldots,X_m,\;Y_1,\ldots,Y_n\big]\\ &\quad= \mathbb{E}\Big[a_N^{-3}\cdot\Big[K''\big(a_N^{-1}(H_N(X_1)-H_N(Y_1))\big)\cdot\big(\hat{H}_N^*(X_1)-H_N(X_1)\big)\end{aligned}$$

$$
\begin{aligned}
&- \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big] \\
&\quad \times \big(H_N(Y_1) - \hat{H}_N^*(Y_1)\big) \,\Big|\, X_2, \ldots, X_m,\ Y_1, \ldots, Y_n\Big] \\
&= a_N^{-3} \cdot \Big[\int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big] \cdot \big(H_N(Y_1) - \hat{H}_N^*(Y_1)\big) \\
&= 0.
\end{aligned}
$$

Thus, using inequality (5.145) it remains only to bound the expectations $\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1)\big]^2$ and $\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1)\big]^2$ in order to derive a bound for the summand (5.155). Now we can derive a similar representation of the expression $(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1)$ as the one derived for $(\hat{u}_N - \hat{u}_N^*)(X_1, X_2)$ in the proof of Lemma 5.19 (see proof for details), which gives us

$$
\begin{aligned}
&(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1) \\
&= a_N^{-3} \cdot \Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big] \cdot \big(H_N(Y_1) - \hat{H}_N^*(Y_1)\big) \\
&+ a_N^{-3} \cdot \Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big] \cdot \big(\hat{H}_N^*(Y_1) - \hat{H}_N(Y_1)\big) \\
&+ a_N^{-3} \cdot \Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big] \cdot \big(\hat{H}_N^*(Y_1) - \hat{H}_N(Y_1)\big),
\end{aligned}
$$

so that for the expectation we have

$$
\begin{aligned}
&\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1)\big]^2 \\
&= \mathbb{E}\Big[a_N^{-3} \cdot \Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big] \cdot \big(H_N(Y_1) - \hat{H}_N^*(Y_1)\big) \\
&+ a_N^{-3} \cdot \Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big] \cdot \big(\hat{H}_N^*(Y_1) - \hat{H}_N(Y_1)\big) \\
&+ a_N^{-3} \cdot \Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big] \cdot \big(\hat{H}_N^*(Y_1) - \hat{H}_N(Y_1)\big)\Big]^2 \\
&\le 4a_N^{-6} \cdot \Big[\mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big)
\end{aligned}
$$

$$
\begin{aligned}
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big] \cdot \big(H_N(Y_1) - \hat{H}_N^*(Y_1)\big)\Big]^2 \\
&+ \mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big] \cdot \big(\hat{H}_N^*(Y_1) - \hat{H}_N(Y_1)\big)\Big]^2 \\
&+ \mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\quad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big] \cdot \big(\hat{H}_N^*(Y_1) - \hat{H}_N(Y_1)\big)\Big]^2\Big] \\
\leq 4a_N^{-6} \cdot &\Big[2 \cdot \mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big)^2 \\
&\quad + \Big[\int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big]^2\Big] \cdot \big(H_N(Y_1) - \hat{H}_N^*(Y_1)\big)^2\Big] \\
&+ 2 \cdot \mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \cdot \big(\hat{H}_N(X_1) - \hat{H}_N^*(X_1)\big)^2 \\
&\quad + \Big[\int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big]^2\Big] \cdot \big(\hat{H}_N^*(Y_1) - \hat{H}_N(Y_1)\big)^2\Big] \\
&+ 2 \cdot \mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big)^2 \\
&\quad + \Big[\int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big]^2\Big] \cdot \big(\hat{H}_N^*(Y_1) - \hat{H}_N(Y_1)\big)^2\Big]\Big] \\
\leq 4a_N^{-6} \cdot &\Big[2 \cdot \mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \cdot 16N^{-2} \\
&\quad + \Big[\int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big]^2\Big] \cdot \big\|H_N - \hat{H}_N^*\big\|^2\Big] \\
&+ 2 \cdot \mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \cdot 16N^{-2} \\
&\quad + \Big[\int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)\, F(dx)\Big]^2\Big] \cdot 16N^{-2}\Big] \\
&+ 2 \cdot \mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \cdot \big\|\hat{H}_N^* - H_N\big\|^2 \\
&\quad + \Big[\int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\Big]^2\Big] \cdot 16N^{-2}\Big]\Big] \\
\leq 4a_N^{-6} \cdot &\Big[2 \cdot \mathbb{E}\Big[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \cdot 16N^{-2} \\
&\quad + \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)^2 \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)^2\, F(dx)\Big] \cdot \big\|H_N - \hat{H}_N^*\big\|^2\Big] \\
&+ 2 \cdot \mathbb{E}\Big[\Big[\big\|K''\big\|^2 \cdot 16N^{-2} \\
&\quad + \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)^2 \cdot \big(\hat{H}_N(x) - \hat{H}_N^*(x)\big)^2\, F(dx)\Big] \cdot 16N^{-2}\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+ 2 \cdot \mathbb{E}\Bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \cdot \big\|\hat{H}_N^* - H_N\big\|^2 \\
&\quad + \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)^2 \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)^2 \, F(dx)\bigg] \cdot 16N^{-2}\Bigg]\Bigg] \\
&\le 8a_N^{-6} \cdot \Bigg[\mathbb{E}\bigg[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \\
&\quad + \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)^2 \cdot \, F(dx)\Big] \cdot 16N^{-2} \cdot \big\|H_N - \hat{H}_N^*\big\|^2\bigg] \\
&\quad + \mathbb{E}\Big[\Big[\big\|K''\big\|^2 \cdot 16N^{-2} + \big\|K''\big\|^2 \cdot 16N^{-2}\Big] \cdot 16N^{-2}\Big] \\
&\quad + \mathbb{E}\bigg[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \\
&\quad + \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)^2 \, F(dx)\Big] \cdot \big\|\hat{H}_N^* - H_N\big\|^2 \cdot 16N^{-2}\bigg]\Bigg] \\
&= 8a_N^{-6} \cdot \Bigg[2 \cdot \mathbb{E}\bigg[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \\
&\quad + \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)^2 \cdot \, F(dx)\Big] \cdot 16N^{-2} \cdot \big\|H_N - \hat{H}_N^*\big\|^2\bigg] \\
&\quad + 2\big\|K''\big\|^2 \cdot 16^2 N^{-4}\Bigg] \\
&= 16^2 a_N^{-6} \cdot N^{-2} \cdot \mathbb{E}\bigg[\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \\
&\quad + \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)^2 \cdot \, F(dx)\Big] \cdot \big\|H_N - \hat{H}_N^*\big\|^2\bigg] \\
&\quad + 16^3\big\|K''\big\|^2 \cdot a_N^{-6} \cdot N^{-4} \\
&= 16^2 a_N^{-6} \cdot N^{-2} \cdot \bigg[\mathbb{E}\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \\
&\quad + \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)^2 \cdot \, F(dx)\Big] \cdot \mathbb{E}\Big[\big\|H_N - \hat{H}_N^*\big\|^2\Big]\bigg] \\
&\quad + 16^3\big\|K''\big\|^2 \cdot a_N^{-6} \cdot N^{-4} \\
&= 16^2 a_N^{-6} \cdot N^{-2} \cdot \Big[2 \cdot \mathbb{E}\Big[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2\Big] \cdot \mathbb{E}\Big[\big\|H_N - \hat{H}_N^*\big\|^2\Big]\Big] \\
&\quad + 16^3\big\|K''\big\|^2 \cdot a_N^{-6} \cdot N^{-4} \\
&\le 16^2 a_N^{-6} \cdot N^{-2} \cdot 4\big\|K''\big\|^2 \cdot a_N\Big(1 + \frac{n}{m}\Big) \cdot \mathbb{E}\Big[\big\|H_N - \hat{H}_N^*\big\|^2\Big] \\
&\quad + 16^3\big\|K''\big\|^2 \cdot a_N^{-6} \cdot N^{-4},
\end{aligned}
$$

where (A.2) was used as before to obtain the final inequality.

Thus, altogether for $\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1)\big]^2$ we have

$$
\begin{aligned}
\mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1)\big]^2 &\le 16^2 a_N^{-6} \cdot N^{-2} \cdot 4\big\|K''\big\|^2 \cdot a_N\Big(1 + \frac{n}{m}\Big) \cdot \mathbb{E}\Big[\big\|H_N - \hat{H}_N^*\big\|^2\Big] \\
&\quad + 16^3\big\|K''\big\|^2 \cdot a_N^{-6} \cdot N^{-4} \\
&= O(a_N^{-5} \cdot N^{-3}) + O(a_N^{-6} \cdot N^{-4})
\end{aligned}
$$

$$= O(a_N^{-5} \cdot N^{-3}). \tag{5.157}$$

Now, for the second expectation in the inequality (5.145) we have

$$\begin{aligned}
&\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1)\big]^2 \\
&= \mathbb{E}\bigg[a_N^{-3} \cdot \bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\bigg] \cdot \big(H_N(Y_1) - \hat{H}_N^*(Y_1)\big)\bigg]^2 \\
&\leq a_N^{-6} \cdot \mathbb{E}\bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big) \\
&\qquad - \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\bigg]^2 \cdot \big\|H_N - \hat{H}_N^*\big\|^2\bigg] \\
&\leq a_N^{-6} \cdot \mathbb{E}\bigg[2 \cdot \bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \cdot \big(\hat{H}_N^*(X_1) - H_N(X_1)\big)^2 \\
&\qquad + \bigg[\int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big) \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)\, F(dx)\bigg]^2\bigg] \cdot \big\|H_N - \hat{H}_N^*\big\|^2\bigg] \\
&\leq a_N^{-6} \cdot \mathbb{E}\bigg[2 \cdot \bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \cdot \big\|\hat{H}_N^* - H_N\big\|^2 \\
&\qquad + \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)^2 \cdot \big(\hat{H}_N^*(x) - H_N(x)\big)^2\, F(dx)\bigg] \cdot \big\|H_N - \hat{H}_N^*\big\|^2\bigg] \\
&\leq 2a_N^{-6} \cdot \mathbb{E}\bigg[\bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \\
&\qquad + \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)^2\, F(dx)\bigg] \cdot \big\|H_N - \hat{H}_N^*\big\|^4\bigg] \\
&= 2a_N^{-6} \cdot \mathbb{E}\bigg[K''\big(a_N^{-1}(H_N(X_1) - H_N(Y_1))\big)^2 \\
&\qquad + \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)^2\, F(dx)\bigg] \cdot \mathbb{E}\Big[\big\|H_N - \hat{H}_N^*\big\|^4\Big] \\
&= 4a_N^{-6} \cdot \mathbb{E}\bigg[\int K''\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)^2\, F(dx)\bigg] \cdot \mathbb{E}\Big[\big\|H_N - \hat{H}_N^*\big\|^4\Big] \\
&\leq 4a_N^{-6} \cdot 2\big\|K''\big\|^2 \cdot a_N\Big(1 + \frac{n}{m}\Big) \cdot \mathbb{E}\Big[\big\|H_N - \hat{H}_N^*\big\|^4\Big].
\end{aligned}$$

Thus, altogether for $\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1)\big]^2$ we have

$$\begin{aligned}
\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1)\big]^2 &\leq 8\big\|K''\big\|^2 \cdot a_N^{-5}\Big(1 + \frac{n}{m}\Big) \cdot \mathbb{E}\Big[\big\|H_N - \hat{H}_N^*\big\|^4\Big] \\
&= O(a_N^{-5}) \cdot O(N^{-2}) \\
&= O(a_N^{-5} \cdot N^{-2}).
\end{aligned} \tag{5.158}$$

Combining (5.157) and (5.158) gives us a bound for the expectation in the last summand (5.155), namely

$$\begin{aligned}
&\big|\mathbb{E}\big[\hat{u}_N(X_1, Y_1) \cdot \hat{u}_N(X_2, Y_2)\big]\big| \\
&\quad \leq \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1)\big]^2 + 2\Big[\mathbb{E}\big[\hat{u}_N^*(X_1, Y_1)\big]^2 \cdot \mathbb{E}\big[(\hat{u}_N - \hat{u}_N^*)(X_1, Y_1)\big]^2\Big]^{\frac{1}{2}}
\end{aligned}$$

$$
\begin{aligned}
&= O(a_N^{-5} \cdot N^{-3}) + 2\Big[O(a_N^{-5} \cdot N^{-2}) \cdot O(a_N^{-5} \cdot N^{-3})\Big]^{\frac{1}{2}} \\
&= O(a_N^{-5} \cdot N^{-3}) + O(a_N^{-5} \cdot N^{-\frac{5}{2}}) \\
&= O(a_N^{-5} \cdot N^{-\frac{5}{2}}). \qquad (5.159)
\end{aligned}
$$

We now have bounds on all of the expectations in the summands (5.152) through (5.155), so that for the sum (5.150) we may write

$$
\begin{aligned}
&\mathbb{E}\Big[m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \hat{u}_N(X_i, Y_k)\Big]^2 \\
&= m^{-2}n^{-2} \cdot \Big[mn \cdot O(a_N^{-6} \cdot N^{-2}) \\
&\qquad + m(m-1)n \cdot O(a_N^{-6} \cdot N^{-2}) \\
&\qquad + mn(n-1) \cdot O(a_N^{-6} \cdot N^{-2}) \\
&\qquad + m(m-1)n(n-1) \cdot O(a_N^{-5} \cdot N^{-\frac{5}{2}})\Big] \\
&= O(a_N^{-5} \cdot N^{-\frac{5}{2}}). \qquad (5.160)
\end{aligned}
$$

This gives us

$$
m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \hat{u}_N(X_i, Y_k) = O_P(a_N^{-\frac{5}{2}} \cdot N^{-\frac{5}{4}})
$$

as claimed. □

LEMMA 5.23.

$$
\begin{aligned}
&\frac{1}{2}a_N^{-4} \cdot m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \Big[K'''(\xi_{ij}) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'''(\tau_j) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Big] \cdot \big(H_N(Y_k) - \hat{H}_N(Y_k)\big)^2 \\
&= O_P(a_N^{-4} \cdot N^{-\frac{3}{2}}).
\end{aligned}
$$

PROOF.

$$
\begin{aligned}
&\frac{1}{2}a_N^{-4} \cdot m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} \Big[K'''(\xi_{ij}) \cdot \big(\hat{H}_N(X_i) - H_N(X_i)\big) \\
&\qquad - \int K'''(\tau_j) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Big] \cdot \big(H_N(Y_k) - \hat{H}_N(Y_k)\big)^2 \\
&\le \frac{1}{2}a_N^{-4} \cdot m^{-2} \cdot \sum_{i=1}^{m}\sum_{j=1}^{m} \Big[\big\|K'''\big\| \cdot \big\|\hat{H}_N - H_N\big\| + \int \big\|K'''\big\| \cdot \big\|\hat{H}_N - H_N\big\|\, F(dx)\Big] \cdot \big\|H_N - \hat{H}_N\big\|^2 \\
&= \frac{1}{2}a_N^{-4} \cdot 2\big\|K'''\big\| \cdot \big\|H_N - \hat{H}_N\big\|^3 \\
&= \big\|K'''\big\| \cdot a_N^{-4} \cdot \big\|H_N - \hat{H}_N\big\|^3 \\
&= O_P(a_N^{-4} \cdot N^{-\frac{3}{2}}).
\end{aligned}
$$

□

Combining Lemmas 5.18, 5.19, 5.20, 5.21, 5.22 and 5.23 we have proven the following.

Lemma 5.24.

$$\int \left[\hat{f}_N - \bar{f}_N\right]' \circ H_N(x) \cdot \left[\hat{H}_N(x) - H_N(x)\right]\left[\hat{F}_m(dx) - F(dx)\right] = O_P(a_N^{-\frac{3}{2}} \cdot N^{-1}), \tag{5.161}$$

$$\int \left[\hat{f}_N - \bar{f}_N\right]' \circ H_N(x) \cdot \left[\hat{H}_N(x) - H_N(x)\right]\left[\hat{F}_m(dx) - F(dx)\right] = O_P(a_N^{-\frac{3}{2}} \cdot N^{-1}) \tag{5.162}$$

and thus

$$\int \left[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\right]' \circ H_N(x) \cdot \left[\hat{H}_N(x) - H_N(x)\right]\left[\hat{F}_m(dx) - F(dx)\right] = O_P(a_N^{-\frac{3}{2}} \cdot N^{-1}). \tag{5.163}$$

5.2.4. Third bounded term. We continue our treatment of the asymptotically negligible terms of the expansion by showing that the term (2.39) is negligible as well. For (2.39) we can write

$$\begin{aligned}
&\int \left[\bar{f}_N - \bar{g}_N\right]' \circ H_N(x) \cdot \left[\hat{H}_N(x) - H_N(x)\right]\left[\hat{F}_m(dx) - F(dx)\right] \\
&= \int \bar{f}'_N \circ H_N(x) \cdot \left[\hat{H}_N(x) - H_N(x)\right]\left[\hat{F}_m(dx) - F(dx)\right] && (5.164) \\
&\quad - \int \bar{g}'_N \circ H_N(x) \cdot \left[\hat{H}_N(x) - H_N(x)\right]\left[\hat{F}_m(dx) - F(dx)\right] && (5.165)
\end{aligned}$$

We will first work at bounding (5.164). The proof for (5.165) follows along similar lines.

Recalling the first order derivatives $\bar{f}'_N$ and $\bar{g}'_N$ (see (5.102)), we have for (5.164):

$$\begin{aligned}
&\int \bar{f}'_N \circ H_N(x) \cdot \left[\hat{H}_N(x) - H_N(x)\right]\left[\hat{F}_m(dx) - F(dx)\right] \\
&= m^{-1} \cdot \sum_{i=1}^{m} \left[\bar{f}'_N \circ H_N(X_i) \cdot \left[\hat{H}_N(X_i) - H_N(X_i)\right] - \int \bar{f}'_N \circ H_N(x) \cdot \left[\hat{H}_N(x) - H_N(x)\right] F(dx)\right] \\
&= m^{-1} \cdot \sum_{i=1}^{m} \left[a_N^{-2} \cdot \int K'\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, F(dy) \cdot \left[\hat{H}_N(X_i) - H_N(X_i)\right]\right. \\
&\qquad \left. - a_N^{-2} \cdot \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot \left[\hat{H}_N(x) - H_N(x)\right] F(dx)\right]
\end{aligned}$$

In contrast to our development of terms (2.35) and (2.36) we won't need to use the Taylor expansion of the kernel function K' here, since there are no occurrences of the random empirical distribution function $\hat{H}_N$ in the arguments of K'. Instead, we can move directly to deriving a bound for (5.164) in the following lemma.

Lemma 5.25.

$$m^{-1} \cdot \sum_{i=1}^{m} \left[a_N^{-2} \cdot \int K'\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, F(dy) \cdot \left[\hat{H}_N(X_i) - H_N(X_i)\right]\right.$$

$$
\begin{aligned}
&- a_N^{-2} \cdot \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx)\Big] \\
&= O_P(a_N^{-2} \cdot N^{-1}).
\end{aligned}
$$

PROOF. Define

$$
u_N(s) = a_N^{-2} \cdot \int K'\big(a_N^{-1}(H_N(s) - H_N(y))\big)\, F(dy).
$$

Then

$$
\begin{aligned}
m^{-1} \cdot \sum_{i=1}^{m} \Big[& a_N^{-2} \cdot \int K'\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, F(dy) \cdot \Big[\hat{H}_N(X_i) - H_N(X_i)\Big] \\
&- a_N^{-2} \cdot \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx)\Big] \\
= m^{-1} \cdot \sum_{i=1}^{m} & \Big[u_N(X_i) \cdot \Big[\hat{H}_N(X_i) - H_N(X_i)\Big] - \int u_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx)\Big] \\
= m^{-1} \cdot \sum_{i=1}^{m} & \Big[u_N(X_i) \cdot \Big[N^{-1}\Big[\sum_{j=1}^{m} 1_{\{X_j \le X_i\}} + \sum_{k=1}^{n} 1_{\{Y_k \le X_i\}}\Big] - H_N(X_i)\Big] \\
&- \int u_N(x) \cdot \Big[N^{-1}\Big[\sum_{j=1}^{m} 1_{\{X_j \le x\}} + \sum_{k=1}^{n} 1_{\{Y_k \le x\}}\Big] - H_N(x)\Big]\, F(dx)\Big] \\
= m^{-1} \cdot \sum_{i=1}^{m} & u_N(X_i) \cdot N^{-1} \sum_{j=1}^{m} 1_{\{X_j \le X_i\}} + m^{-1} \cdot \sum_{i=1}^{m} u_N(X_i) \cdot N^{-1} \sum_{k=1}^{n} 1_{\{Y_k \le X_i\}} \\
&- m^{-1} \cdot \sum_{i=1}^{m} u_N(X_i) \cdot H_N(X_i) \\
&- \int u_N(x) \cdot N^{-1} \cdot \sum_{j=1}^{m} 1_{\{X_j \le x\}}\, F(dx) - \int u_N(x) \cdot N^{-1} \cdot \sum_{k=1}^{n} 1_{\{Y_k \le x\}}\, F(dx) \\
&+ \int u_N(x) \cdot H_N(x)\, F(dx) \\
= \lambda_N \cdot m^{-2} & \cdot \sum_{i=1}^{m} \sum_{j=1}^{m} u_N(X_i) \cdot 1_{\{X_j \le X_i\}} + (1 - \lambda_N) \cdot m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} u_N(X_i) \cdot 1_{\{Y_k \le X_i\}} \\
&- m^{-1} \cdot \sum_{i=1}^{m} u_N(X_i) \cdot H_N(X_i) \\
&- \lambda_N \cdot m^{-1} \cdot \sum_{j=1}^{m} \int u_N(x) \cdot 1_{\{X_j \le x\}}\, F(dx) - (1 - \lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x) \cdot 1_{\{Y_k \le x\}}\, F(dx) \\
&+ \int u_N(x) \cdot H_N(x)\, F(dx) \\
= \lambda_N \cdot m^{-2} & \cdot \sum_{1 \le i \ne j \le m} u_N(X_i) \cdot 1_{\{X_j \le X_i\}} + (1 - \lambda_N) \cdot m^{-1} n^{-1} \cdot \sum_{i=1}^{m} \sum_{k=1}^{n} u_N(X_i) \cdot 1_{\{Y_k \le X_i\}} \\
&+ \lambda_N \cdot m^{-2} \cdot \sum_{i=1}^{m} u_N(X_i) - m^{-1} \cdot \sum_{i=1}^{m} u_N(X_i) \cdot H_N(X_i) \\
&- \lambda_N \cdot m^{-1} \cdot \sum_{j=1}^{m} \int u_N(x) \cdot 1_{\{X_j \le x\}}\, F(dx) - (1 - \lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x) \cdot 1_{\{Y_k \le x\}}\, F(dx)
\end{aligned}
$$

$$+ \int u_N(x) \cdot H_N(x)\, F(dx)$$

Define the U-statistic U_m^1 and the generalized U-statistic $U_{m,n}^2$ as

$$U_m^1 = m^{-1}(m-1)^{-1} \cdot \sum_{1 \le i \neq j \le m} u_N(X_i) \cdot 1_{\{X_j \le X_i\}},$$

$$U_{m,n}^2 = m^{-1} n^{-1} \cdot \sum_{\substack{1 \le i \le m \\ 1 \le k \le n}} u_N(X_i) \cdot 1_{\{Y_k \le X_i\}},$$

and let $\hat{U}_m^1$ and $\hat{U}_{m,n}^2$ be the Hájek projections of U_m^1 and $U_{m,n}^2$ respectively as defined in Lemmas A.2 and A.3. Then (5.164) is equal to

$$\begin{aligned}
&\frac{\lambda_N(m-1)}{m} \cdot U_m^1 + (1-\lambda_N) \cdot U_{m,n}^2 + \lambda_N \cdot m^{-2} \cdot \sum_{i=1}^m u_N(X_i) - m^{-1} \cdot \sum_{i=1}^m u_N(X_i) \cdot H_N(X_i) \\
&- \lambda_N \cdot m^{-1} \cdot \sum_{j=1}^m \int u_N(x) \cdot 1_{\{X_j \le x\}}\, F(dx) - (1-\lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^n \int u_N(x) \cdot 1_{\{Y_k \le x\}}\, F(dx) \\
&+ \int u_N(x) \cdot H_N(x)\, F(dx).
\end{aligned}$$

Now, the kernel function u_N is bounded:

$$\|u_N\| \le 2\|K'\| a_N^{-2}.$$

Which means for the third sum $\lambda_N \cdot m^{-2} \cdot \sum_{i=1}^m u_N(X_i)$ we can write

$$\begin{aligned}
\left| \lambda_N \cdot m^{-2} \cdot \sum_{i=1}^m u_N(X_i) \right| &\le \lambda_N \cdot m^{-2} \cdot \sum_{i=1}^m |u_N(X_i)| \\
&\le \lambda_N \cdot m^{-2} \cdot \sum_{i=1}^m \|u_N\| \\
&= \lambda_N \cdot m^{-1} \cdot \|u_N\| \\
&= O(a_N^{-2} \cdot N^{-1}). \qquad (5.166)
\end{aligned}$$

Thus, we can partition (5.164) into the sum of two scaled U-statistics, some i.i.d. sequences and a negligible rest:

$$\begin{aligned}
&\frac{\lambda_N(m-1)}{m} \cdot U_m^1 + (1-\lambda_N) \cdot U_{m,n}^2 + O(a_N^{-2} \cdot N^{-1}) - m^{-1} \cdot \sum_{i=1}^m u_N(X_i) \cdot H_N(X_i) \\
&- \lambda_N \cdot m^{-1} \cdot \sum_{j=1}^m \int u_N(x) \cdot 1_{\{X_j \le x\}}\, F(dx) - (1-\lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^n \int u_N(x) \cdot 1_{\{Y_k \le x\}}\, F(dx) \\
&+ \int u_N(x) \cdot H_N(x)\, F(dx). \qquad (5.167)
\end{aligned}$$

In the following we will show that (5.167) is $O(a_N^{-2} \cdot N^{-1})$ as well, which will complete the proof. Begin by calculating each of the projections $\hat{U}_m^1$ and $\hat{U}_{m,n}^2$. Firstly, for $\lambda_N \cdot \hat{U}_m^1$

$$\lambda_N \cdot \hat{U}_m^1 = \lambda_N \cdot m^{-1} \cdot \sum_{i=1}^m \Big[\int u_N(X_i) \cdot 1_{\{x \le X_i\}}\, F(dx) + \int u_N(x) \cdot 1_{\{X_i \le x\}}\, F(dx)$$

$$
\begin{aligned}
&- \iint u_N(x) \cdot 1_{\{y \le x\}} \, F(dx)F(dy)\Big] \\
&= m^{-1} \cdot \sum_{i=1}^{m} u_N(X_i) \cdot \lambda_N \cdot F(X_i) + \lambda_N \cdot m^{-1} \cdot \sum_{i=1}^{m} \int u_N(x) \cdot 1_{\{X_i \le x\}} \, F(dx) \\
&\quad - \int u_N(x) \cdot \lambda_N \cdot F(x) \, F(dx).
\end{aligned}
$$

Nextly, for $(1-\lambda_N) \cdot \hat{U}^2_{m,n}$ we have

$$
\begin{aligned}
(1-\lambda_N) \cdot \hat{U}^2_m &= (1-\lambda_N) \cdot \Big[m^{-1} \cdot \sum_{i=1}^{m} \int u_N(X_i) \cdot 1_{\{x \le X_i\}} \, G(dx) + n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x) \cdot 1_{\{Y_k \le x\}} \, F(dx) \\
&\quad - \iint u_N(x) \cdot 1_{\{y \le x\}} \, F(dx)G(dy)\Big] \\
&= m^{-1} \cdot \sum_{i=1}^{m} u_N(X_i) \cdot (1-\lambda_N) \cdot G(X_i) + (1-\lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x) \cdot 1_{\{Y_k \le x\}} \, F(dx) \\
&\quad - \int u_N(x) \cdot (1-\lambda_N) \cdot G(x) \, F(dx).
\end{aligned}
$$

Now, since

$$H_N = \lambda_N \cdot F + (1-\lambda_N) \cdot G,$$

we see that

$$
\begin{aligned}
&\lambda_N \cdot \hat{U}^1_m + (1-\lambda_N) \cdot \hat{U}^2_{m,n} \\
&\quad = m^{-1} \cdot \sum_{i=1}^{m} u_N(X_i) \cdot H_N(X_i) \\
&\quad + \lambda_N \cdot m^{-1} \cdot \sum_{i=1}^{m} \int u_N(x) \cdot 1_{\{X_i \le x\}} \, F(dx) + (1-\lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x) \cdot 1_{\{Y_k \le x\}} \, F(dx) \\
&\quad - \int u_N(x) \cdot H_N(x) \, F(dx).
\end{aligned}
$$

so that

$$
\begin{aligned}
&\lambda_N \cdot \hat{U}^1_m + (1-\lambda_N) \cdot \hat{U}^2_{m,n} - m^{-1} \cdot \sum_{i=1}^{m} u_N(X_i) \cdot H_N(X_i) \\
&\quad - \lambda_N \cdot m^{-1} \cdot \sum_{j=1}^{m} \int u_N(x) \cdot 1_{\{X_j \le x\}} \, F(dx) - (1-\lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x) \cdot 1_{\{Y_k \le x\}} \, F(dx) \\
&\quad + \int u_N(x) \cdot H_N(x) \, F(dx) \\
&\quad = 0.
\end{aligned}
$$

Thus, for the representation (5.167) of (5.164) we have

$$
\frac{\lambda_N(m-1)}{m} \cdot U^1_m + (1-\lambda_N) \cdot U^2_{m,n} + O(a_N^{-2} \cdot N^{-1}) - m^{-1} \cdot \sum_{i=1}^{m} u_N(X_i) \cdot H_N(X_i)
$$

$$
\begin{aligned}
&- \lambda_N \cdot m^{-1} \cdot \sum_{j=1}^{m} \int u_N(x) \cdot 1_{\{X_j \le x\}} \, F(dx) - (1-\lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x) \cdot 1_{\{Y_k \le x\}} \, F(dx) \\
&+ \int u_N(x) \cdot H_N(x) \, F(dx) \\
= &\Bigg[\lambda_N \cdot U_m^1 + (1-\lambda_N) \cdot U_{m,n}^2 - m^{-1} \cdot \sum_{i=1}^{m} u_N(X_i) \cdot H_N(X_i) \\
&- \lambda_N \cdot m^{-1} \cdot \sum_{j=1}^{m} \int u_N(x) \cdot 1_{\{X_j \le x\}} \, F(dx) - (1-\lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x) \cdot 1_{\{Y_k \le x\}} \, F(dx) \\
&+ \int u_N(x) \cdot H_N(x) \, F(dx) + O(a_N^{-2} \cdot N^{-1}) \Bigg] - \frac{\lambda_N}{m} \cdot U_m^1 \\
= &\Big[\lambda_N \cdot \Big[U_m^1 - \hat{U}_m^1 \Big] + (1-\lambda_N) \cdot \Big[U_{m,n}^2 - \hat{U}_{m,n}^2 \Big] + O(a_N^{-2} \cdot N^{-1}) \Big] - \frac{\lambda_N}{m} \cdot U_m^1,
\end{aligned}
$$

and it remains only to bound $\big[U_m^1 - \hat{U}_m^1\big]$, $\big[U_{m,n}^2 - \hat{U}_{m,n}^2\big]$ and $\frac{\lambda_N}{m} \cdot U_m^1$. Firstly, using Lemma A.2 we have

$$
\mathbb{E}\Big[U_m^1 - \hat{U}_m^1\Big]^2 \le 2(m-1)m^{-3} \cdot \mathbb{E}\big[u_{1N}^*(X_1, X_2)\big]^2
$$

for u_{1N}^* defined as

$$
\begin{aligned}
u_{1N}^*(r,s) = u_N(r) \cdot 1_{\{s \le r\}} &- \int u_N(r) \cdot 1_{\{y \le r\}} \, F(dy) \\
&- \int u_N(x) \cdot 1_{\{s \le x\}} \, F(dx) + \iint u_N(x) \cdot 1_{\{y \le x\}} \, F(dx) F(dy)
\end{aligned}
$$

so that the expectation is easily bounded:

$$
\begin{aligned}
&\mathbb{E}\big[u_{1N}^*(X_1, X_2)\big]^2 \\
&\le 4 \cdot \mathbb{E}\Bigg[\big[u_N(X_1) \cdot 1_{\{X_2 \le X_1\}}\big]^2 + \Big[\int u_N(X_1) \cdot 1_{\{y \le X_1\}} \, F(dy) \Big]^2 \\
&\qquad + \Big[\int u_N(x) \cdot 1_{\{X_2 \le x\}} \, F(dx) \Big]^2 + \Big[\iint u_N(x) \cdot 1_{\{y \le x\}} \, F(dx) F(dy) \Big]^2 \Bigg] \\
&\le 4 \cdot \mathbb{E}\Bigg[\big[u_N(X_1)\big]^2 + \big[u_N(X_1)\big]^2 \\
&\qquad + \int \big[u_N(x)\big]^2 \, F(dx) + \Big[\int u_N(x) \cdot F(x) \, F(dy) \Big]^2 \Bigg] \\
&\le 4 \cdot \mathbb{E}\Big[4 \cdot \big[2a_N^{-2} \cdot \|K'\| \big]^2 \Big] \\
&\le 4^2 \cdot 2 \, \|K'\|^2 \cdot a_N^{-4}.
\end{aligned}
$$

Altogether this yields

$$
\begin{aligned}
\mathbb{E}\Big[U_m^1 - \hat{U}_m^1\Big]^2 &\le 2(m-1)m^{-3} \cdot \mathbb{E}\big[u_{1N}^*(X_1, X_2)\big]^2 \\
&\le 2(m-1)m^{-3} \cdot 4^2 \cdot 2 \, \|K'\|^2 \cdot a_N^{-4} \\
&= 4^3 \big\|K'\big\|^2 \cdot a_N^{-4} \cdot (m-1)m^{-3} \\
&= O(a_N^{-4} \cdot N^{-2}). \qquad (5.168)
\end{aligned}
$$

Similarly, using Lemma A.3 we have

$$\mathbb{E}\Big[U^2_{m,n} - \hat{U}^2_{m,n}\Big]^2 = m^{-1}n^{-1} \cdot \mathbb{E}\big[u^*_{2N}(X_1, Y_1)\big]^2$$

for u^*_{2N} defined as

$$\begin{aligned} u^*_{2N}(r,s) &= u_N(r) \cdot 1_{\{s \le r\}} - \int u_N(r) \cdot 1_{\{s \le y\}}\, G(dy) \\ &\quad - \int u_N(x) \cdot 1_{\{s \le x\}}\, F(dx) + \iint u_N(x) \cdot 1_{\{y \le x\}}\, F(dx)\, G(dy). \end{aligned}$$

Bounding $\mathbb{E}\big[u^*_{2N}(X_1, Y_1)\big]^2$ we obtain

$$\mathbb{E}\big[u^*_{2N}(X_1, Y_1)\big]^2 \le 4^2 \cdot 2 \, \|K'\|^2 \cdot a_N^{-4}$$

(proof completely analogous to the above showing that $\mathbb{E}\big[u^*_{2N}(X_1, X_2)\big]^2 \le 4^2 \cdot 2 \, \|K'\|^2 \cdot a_N^{-4}$). which gives us

$$\mathbb{E}\Big[U^2_{m,n} - \hat{U}^2_{m,n}\Big]^2 = O(m^{-1}n^{-1}) \cdot O(a_N^{-4}) = O(a_N^{-4} \cdot N^{-2}) \tag{5.169}$$

Lastly,

$$\begin{aligned} \frac{\lambda_N}{m} \cdot U^1_m &= \frac{\lambda_N}{m} \cdot m^{-1}(m-1)^{-1} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i) \cdot 1_{\{X_j \le X_i\}} \\ &\le \frac{\lambda_N}{m^2(m-1)} \cdot \sum_{1 \le i \ne j \le m} \big\|u_N\big\| \\ &= \frac{\lambda_N \cdot \big\|u_N\big\|}{m} \\ &\le \frac{2\lambda_N \cdot 2\big\|K'\big\| a_N^{-2}}{m} \\ &= O(a_N^{-2} \cdot N^{-1}). \end{aligned} \tag{5.170}$$

Combining (5.168), (5.169) and (5.170) we see that (5.164) is equal to

$$\begin{aligned} &\Big[\lambda_N \cdot \Big[U^1_m - \hat{U}^1_m\Big] + (1-\lambda_N) \cdot \Big[U^2_{m,n} - \hat{U}^2_{m,n}\Big] + O(a_N^{-2} \cdot N^{-1})\Big] - \frac{\lambda_N}{m} \cdot U^1_m \\ &= \Big[O_P(a_N^{-2} \cdot N^{-1}) + O_P(a_N^{-2} \cdot N^{-1}) + O(a_N^{-2} \cdot N^{-1})\Big] - O(a_N^{-2} \cdot N^{-1}) \\ &= O_P(a_N^{-2} \cdot N^{-1}) \end{aligned}$$

which completes the proof. □

To bound (5.165) we will use very similar arguments to those which we used to show that (5.164) is $O_P(a_N^{-2} \cdot N^{-1})$. We begin by deriving a sum representation of (5.165).Recalling the first order derivative $\bar{g}'_N$ (see (5.102)), we have for (5.165):

$$\begin{aligned} &\int \bar{g}'_N \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big] \Big[\hat{F}_m(dx) - F(dx)\Big] \\ &= m^{-1} \cdot \sum_{i=1}^{m} \Big[\bar{g}'_N \circ H_N(X_i) \cdot \Big[\hat{H}_N(X_i) - H_N(X_i)\Big] - \int \bar{g}'_N \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx)\Big] \end{aligned}$$

$$= m^{-1} \cdot \sum_{i=1}^{m} \Big[a_N^{-2} \cdot \int K'\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, G(dy) \cdot \Big[\hat{H}_N(X_i) - H_N(X_i)\Big]$$
$$- a_N^{-2} \cdot \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx)\Big]$$

Lemma 5.26.

$$m^{-1} \cdot \sum_{i=1}^{m} \Big[a_N^{-2} \cdot \int K'\big(a_N^{-1}(H_N(X_i) - H_N(y))\big)\, G(dy) \cdot \Big[\hat{H}_N(X_i) - H_N(X_i)\Big]$$
$$- a_N^{-2} \cdot \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx)\Big]$$
$$= O_P(a_N^{-2} \cdot N^{-1}).$$

Proof. Define

$$u_N(s) = a_N^{-2} \cdot \int K'\big(a_N^{-1}(H_N(s) - H_N(y))\big)\, G(dy).$$

Then the rest of the proof is identical to the proof of Lemma 5.25, which depends only on the fact that u_N is uniformly bounded:

$$\|u_N\| \le 2\|K'\|a_N^{-2}$$

which is the case for u_N defined here as well. □

Combining Lemmas 5.25 and 5.26 we have proven the following.

Lemma 5.27.

$$\int \bar{f}_N' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\Big[\hat{F}_m(dx) - F(dx)\Big] = O_P(a_N^{-2} \cdot N^{-1}), \tag{5.171}$$

$$\int \bar{g}_N' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\Big[\hat{F}_m(dx) - F(dx)\Big] = O_P(a_N^{-2} \cdot N^{-1}) \tag{5.172}$$

and thus

$$\int \Big[\bar{f}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\Big[\hat{F}_m(dx) - F(dx)\Big] = O_P(a_N^{-2} \cdot N^{-1}). \tag{5.173}$$

5.2.5. Fourth bounded term. We continue our treatment of the asymptotically negligible terms of the expansion by showing that the term (2.45) is negligible as well. For (2.45) we can write

$$\int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx)$$

$$= \int \Big[\hat{f}_N - \bar{f}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx) \tag{5.174}$$

$$- \int \Big[\hat{g}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx) \tag{5.175}$$

We will first work at bounding (5.174). The proof for (5.175) follows along similar lines.

Recalling (5.101) and (5.102) for (5.174) we may write

$$
\begin{aligned}
&\int \Big[\hat{f}_N - \bar{f}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx) \\
&= a_N^{-2} \cdot m^{-1} \int \sum_{j=1}^{m} K'\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_j))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\quad - a_N^{-2} \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&= a_N^{-2} \cdot m^{-1} \cdot \sum_{j=1}^{m} \Big[\int K'\big(a_N^{-1}(H_N(x) - \hat{H}_N(X_j))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Big].
\end{aligned}
$$

At this point we use the Taylor expansion of the kernel function K about each of the $a_N^{-1}(H_N(x) - \hat{H}_N(X_j))$ which yields

$$
\begin{aligned}
&\int \Big[\hat{f}_N - \bar{f}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx) \\
&= a_N^{-2} \cdot m^{-1} \cdot \sum_{j=1}^{m} \Big[\int K'\big(a_N^{-1}(H_N(x) - H_N(X_j))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Big] && (5.176) \\
&\quad + a_N^{-3} \cdot m^{-1} \cdot \sum_{j=1}^{m} \int K''\big(a_N^{-1}(H_N(x) - H_N(X_j))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \cdot \big(H_N(X_j) - \hat{H}_N(X_j)\big) && (5.177) \\
&\quad + \frac{1}{2} a_N^{-4} \cdot m^{-1} \cdot \sum_{j=1}^{m} \int K'''(\tau_j) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \cdot \big(H_N(X_j) - \hat{H}_N(X_j)\big)^2 && (5.178)
\end{aligned}
$$

where τ_j are appropriate values between the two ratios.

In the following lemmas we will derive bounds for the three terms (5.176), (5.177) and (5.178).

Lemma 5.28.

$$
\begin{aligned}
&a_N^{-2} \cdot m^{-1} \cdot \sum_{j=1}^{m} \Big[\int K'\big(a_N^{-1}(H_N(x) - H_N(X_j))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Big] \\
&\quad = O_P(a_N^{-2} \cdot N^{-1}).
\end{aligned}
$$

Proof. Define

$$
u_N(s,\, t) = a_N^{-2} \cdot \Big[K'\big(a_N^{-1}(H_N(s) - H_N(t))\big) - \int K'\big(a_N^{-1}(H_N(s) - H_N(y))\big)\, F(dy)\Big].
$$

Then

$$
\begin{aligned}
&a_N^{-2} \cdot m^{-1} \cdot \sum_{j=1}^{m} \Big[\int K'\big(a_N^{-1}(H_N(x) - H_N(X_j))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\quad - \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \Big] \\
&\quad = m^{-1} \cdot \sum_{j=1}^{m} \int u_N(x, X_j) \cdot \big[\hat{H}_N(x) - H_N(x)\big]\, F(dx) \\
&\quad = m^{-1} \cdot \sum_{j=1}^{m} \int u_N(x, X_j) \cdot \Big[N^{-1}\Big[\sum_{i=1}^{m} 1_{\{X_i \le x\}} + \sum_{k=1}^{n} 1_{\{Y_k \le x\}}\Big] - H_N(x)\Big]\, F(dx) \\
&\quad = m^{-1} \cdot \sum_{j=1}^{m} \int u_N(x, X_j) \cdot N^{-1} \sum_{i=1}^{m} 1_{\{X_i \le x\}}\, F(dx) \\
&\qquad + m^{-1} \cdot \sum_{j=1}^{m} \int u_N(x, X_j) \cdot N^{-1} \sum_{k=1}^{n} 1_{\{Y_k \le x\}}\, F(dx) \\
&\qquad - m^{-1} \cdot \sum_{j=1}^{m} \int u_N(x, X_j) \cdot H_N(x)\, F(dx) \\
&\quad = \lambda_N \cdot m^{-2} \cdot \sum_{1 \le i \ne j \le m} \int u_N(x, X_j) \cdot 1_{\{X_i \le x\}}\, F(dx) + \lambda_N \cdot m^{-2} \cdot \sum_{i=1}^{m} \int u_N(x, X_i) \cdot 1_{\{X_i \le x\}}\, F(dx) \\
&\qquad + (1 - \lambda_N) \cdot m^{-1} n^{-1} \cdot \sum_{j=1}^{m} \sum_{k=1}^{n} \int u_N(x, X_j) \cdot 1_{\{Y_k \le x\}}\, F(dx) \\
&\qquad - m^{-1} \cdot \sum_{j=1}^{m} \int u_N(x, X_j) \cdot H_N(x)\, F(dx).
\end{aligned}
$$

Define the U-statistic U_m^1 and the generalized U-statistic $U_{m,n}^2$ as

$$
\begin{aligned}
U_m^1 &= m^{-1}(m-1)^{-1} \cdot \sum_{1 \le i \ne j \le m} \int u_N(x, X_j) \cdot 1_{\{X_i \le x\}}\, F(dx), \\
U_{m,n}^2 &= m^{-1} n^{-1} \cdot \sum_{j=1}^{m} \sum_{k=1}^{n} \int u_N(x, X_j) \cdot 1_{\{Y_k \le x\}}\, F(dx),
\end{aligned}
$$

and let $\hat{U}_m^1$ and $\hat{U}_{m,n}^2$ be the Hájek projections of U_m^1 and $U_{m,n}^2$ respectively as defined in Lemmas A.2 and A.3. Then (5.176) is equal to

$$
\begin{aligned}
&\frac{\lambda_N \cdot (m-1)}{m} \cdot U_m^1 + \lambda_N \cdot m^{-2} \cdot \sum_{i=1}^{m} \int u_N(x, X_i) \cdot 1_{\{X_i \le x\}}\, F(dx) \\
&\quad + (1 - \lambda_N) \cdot U_{m,n}^2 - m^{-1} \cdot \sum_{j=1}^{m} \int u_N(x, X_j) \cdot H_N(x)\, F(dx).
\end{aligned} \tag{5.179}
$$

Now, the kernel function u_N is bounded:

$$
\|u_N\| \le 2\|K'\| a_N^{-2}.
$$

Which means for the second sum in (5.179) we can write

$$\begin{aligned}
&\left|\lambda_N \cdot m^{-2} \cdot \sum_{i=1}^{m} \int u_N(x, X_i) \cdot 1_{\{X_i \le x\}}\, F(dx)\right| \\
&\quad \le \lambda_N \cdot m^{-2} \cdot \sum_{i=1}^{m} \left| \int u_N(x, X_i) \cdot 1_{\{X_i \le x\}}\, F(dx)\right| \\
&\quad \le \lambda_N \cdot m^{-2} \cdot \sum_{i=1}^{m} \|u_N\| \\
&\quad = \lambda_N \cdot m^{-1} \|u_N\| \\
&\quad = O(a_N^{-2} \cdot N^{-1}). \qquad (5.180)
\end{aligned}$$

Thus, we can partition (5.176) into the sum of two scaled U-statistics, an i.i.d sum and a negligible rest:

$$\frac{\lambda_N \cdot (m-1)}{m} \cdot U_m^1 + (1-\lambda_N) \cdot U_{m,n}^2 - m^{-1} \cdot \sum_{j=1}^{m} \int u_N(x, X_j) \cdot H_N(x)\, F(dx) + O(a_N^{-2} \cdot N^{-1}). \quad (5.181)$$

In the following we will show that (5.181) is $O(a_N^{-2} \cdot N^{-1})$ as well, which will complete the proof. Begin by calculating each of the projections $\hat{U}_m^1$ and $\hat{U}_{m,n}^2$. Firstly,

$$\begin{aligned}
\lambda_N \cdot \hat{U}_m^1 = \lambda_N \cdot m^{-1} \cdot \sum_{i=1}^{m} \Big[&\iint u_N(x, X_i) \cdot 1_{\{y \le x\}}\, F(dx)\, F(dy) + \iint u_N(x, y) \cdot 1_{\{X_i \le x\}}\, F(dx)\, F(dy) \\
&- \iiint u_N(x, y) \cdot 1_{\{z \le x\}}\, F(dx)\, F(dy)\, F(dz) \Big] \\
= m^{-1} \cdot \sum_{i=1}^{m} \Big[&\int u_N(x, X_i) \cdot \lambda_N F(x)\, F(dx) + \lambda_N \cdot \iint u_N(x, y) \cdot 1_{\{X_i \le x\}}\, F(dx)\, F(dy) \\
&- \iint u_N(x, y) \cdot \lambda_N F(x)\, F(dx)\, F(dy) \Big].
\end{aligned}$$

Nextly, for $(1-\lambda_N) \cdot \hat{U}_{m,n}^2$ we have

$$\begin{aligned}
(1-\lambda_N) \cdot \hat{U}_{m,n}^2 = (1-\lambda_N) \cdot \Big[&m^{-1} \cdot \sum_{i=1}^{m} \iint u_N(x, X_i) \cdot 1_{\{y \le x\}}\, F(dx)\, G(dy) \\
&+ n^{-1} \cdot \sum_{k=1}^{n} \iint u_N(x, y) \cdot 1_{\{Y_k \le x\}}\, F(dx)\, F(dy) \\
&- \iiint u_N(x, z) \cdot 1_{\{y \le x\}}\, F(dx)\, G(dy)\, F(dz) \Big] \\
= m^{-1} \cdot \sum_{i=1}^{m} &\int u_N(x, X_i) \cdot (1-\lambda_N) G(x)\, F(dx) \\
&+ (1-\lambda_N) \cdot n^{-1} \cdot \sum_{k=1}^{n} \iint u_N(x, y) \cdot 1_{\{Y_k \le x\}}\, F(dx)\, F(dy) \\
&- \iint u_N(x, z) \cdot (1-\lambda_N) G(x)\, F(dx)\, F(dz).
\end{aligned}$$

Now, since

$$H_N = \lambda_N \cdot F + (1-\lambda_N) \cdot G,$$

we see that

$$\begin{aligned}
&\lambda_N \cdot \hat{U}^1_m + (1-\lambda_N) \cdot \hat{U}^2_{m,n} - m^{-1} \cdot \sum_{j=1}^m \int u_N(x, X_j) \cdot H_N(x)\, F(dx) \\
&= m^{-1} \cdot \sum_{i=1}^m \Big[\lambda_N \cdot \iint u_N(x,y) \cdot 1_{\{X_i \le x\}}\, F(dx)\, F(dy) - \iint u_N(x,y) \cdot \lambda_N F(x)\, F(dx)\, F(dy)\Big] \\
&\quad + (1-\lambda_N) \cdot \Big[n^{-1} \cdot \sum_{k=1}^n \iint u_N(x,y) \cdot 1_{\{Y_k \le x\}}\, F(dx)\, F(dy) - \iint u_N(x,z) \cdot G(x)\, F(dx)\, F(dz)\Big] \\
&= 0
\end{aligned}$$

due to

$$\begin{aligned}
&\int u_N(x,y)\, F(dy) \\
&\quad = \int a_N^{-2} \cdot \Big[K'\big(a_N^{-1}(H_N(x) - H_N(y))\big) - \int K'\big(a_N^{-1}(H_N(x) - H_N(z))\big)\, F(dz)\Big] F(dy) \\
&\quad = a_N^{-2} \cdot \Big[\int K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dy) - \int K'\big(a_N^{-1}(H_N(x) - H_N(z))\big)\, F(dz)\Big] \\
&\quad = 0.
\end{aligned}$$

Thus, for (5.176) we have

$$\begin{aligned}
&\frac{\lambda_N \cdot (m-1)}{m} \cdot U^1_m + (1-\lambda_N) \cdot U^2_{m,n} - m^{-1} \cdot \sum_{j=1}^m \int u_N(x, X_j) \cdot H_N(x)\, F(dx) + O(a_N^{-2} \cdot N^{-1}) \\
&= \lambda_N \cdot U^1_m + \frac{\lambda_N}{m} \cdot U^1_m + (1-\lambda_N) \cdot U^2_{m,n} - m^{-1} \cdot \sum_{j=1}^m \int u_N(x, X_j) \cdot H_N(x)\, F(dx) + O(a_N^{-2} \cdot N^{-1}) \\
&= \lambda_N \cdot \Big[U^1_m - \hat{U}^1_m\Big] + \frac{\lambda_N}{m} \cdot U^1_m + (1-\lambda_N) \cdot \Big[U^2_{m,n} - \hat{U}^2_{m,n}\Big] + O(a_N^{-2} \cdot N^{-1})
\end{aligned}$$

and it remains only to bound $\big[U^1_m - \hat{U}^1_m\big]$, $\big[U^2_{m,n} - \hat{U}^2_{m,n}\big]$ and $\frac{\lambda_N}{m} \cdot U^1_m$. Firstly, using Lemma A.2 we have

$$\mathbb{E}\Big[U^1_m - \hat{U}^1_m\Big]^2 \le 2(m-1)m^{-3} \cdot \mathbb{E}\big[u^*_{1N}(X_1, X_2)\big]^2$$

for u^*_{1N} defined as

$$\begin{aligned}
u^*_{1N}(r,s) = &\int u_N(x,s) \cdot 1_{\{r \le x\}}\, F(dx) - \iint u_N(x,y) \cdot 1_{\{r \le x\}}\, F(dx)\, F(dy) \\
&- \iint u_N(x,s) \cdot 1_{\{y \le x\}}\, F(dx)\, F(dy) + \iiint u_N(x,y) \cdot 1_{\{z \le x\}}\, F(dx)\, F(dy)\, F(dz)
\end{aligned}$$

so that the expectation is easily bounded:

$$\begin{aligned}
&\mathbb{E}\big[u^*_{1N}(X_1, X_2)\big]^2 \\
&\le 4 \cdot \mathbb{E}\Big[\Big[\int u_N(x, X_2) \cdot 1_{\{X_1 \le x\}}\, F(dx)\Big]^2 + \Big[\iint u_N(x,y) \cdot 1_{\{X_1 \le x\}}\, F(dx)\, F(dy)\Big]^2 \\
&\quad + \Big[\iint u_N(x, X_2) \cdot 1_{\{z \le x\}}\, F(dx)\, F(dz)\Big]^2 + \Big[\iiint u_N(x,y) \cdot 1_{\{z \le x\}}\, F(dx)\, F(dy)\, F(dz)\Big]^2\Big]
\end{aligned}$$

$$\le 4\cdot\mathbb{E}\Big[\int\big[u_N(x,X_2)\big]^2\,F(dx)+0+\iint\big[u_N(x,X_2)\big]^2\,F(dx)\,F(dz)+0\Big]$$
$$\le 4\cdot\mathbb{E}\Big[\int\big[2a_N^{-2}\cdot\|K'\|\big]^2\,F(dx)+\iint\big[2a_N^{-2}\cdot\|K'\|\big]^2\,F(dx)\,F(dz)\Big]$$
$$=32\,\|K'\|^2\cdot a_N^{-4}.$$

Altogether this yields

$$\begin{aligned}\mathbb{E}\Big[U_m^1-\hat U_m^1\Big]^2&\le 2(m-1)m^{-3}\cdot\mathbb{E}\big[u_{1N}^*(X_1,X_2)\big]^2\\&\le 2(m-1)m^{-3}\cdot 32\big\|K'\big\|^2\cdot a_N^{-4}\\&=64\big\|K'\big\|^2\cdot a_N^{-4}\cdot(m-1)m^{-3}\\&=O(a_N^{-4}\cdot N^{-2}).\end{aligned}\tag{5.182}$$

Similarly, since the kernel functions of U_m^1 and $U_{m,n}^2$ are equal, using Lemma A.3 for $U_{m,n}^2$ we can write

$$\mathbb{E}\Big[U_{m,n}^2-\hat U_{m,n}^2\Big]^2=m^{-1}n^{-1}\cdot\mathbb{E}\big[u_{2N}^*(X_1,Y_1)\big]^2$$

for u_{2N}^* defined as

$$\begin{aligned}u_{2N}^*(r,s)=&\int u_N(x,s)\cdot 1_{\{r\le x\}}\,F(dx)-\iint u_N(x,y)\cdot 1_{\{r\le x\}}\,F(dx)\,G(dy)\\&-\iint u_N(x,s)\cdot 1_{\{y\le x\}}\,F(dx)\,F(dy)+\iiint u_N(x,y)\cdot 1_{\{z\le x\}}\,F(dx)\,G(dy)\,F(dz)\end{aligned}$$

so that the expectation is easily bounded:

$$\begin{aligned}&\mathbb{E}\big[u_{2N}^*(X_1,Y_1)\big]^2\\&\le 4\cdot\mathbb{E}\Big[\Big[\int u_N(x,Y_1)\cdot 1_{\{X_1\le x\}}\,F(dx)\Big]^2+\Big[\iint u_N(x,y)\cdot 1_{\{X_1\le x\}}\,F(dx)\,G(dy)\Big]^2\\&\quad+\Big[\iint u_N(x,Y_1)\cdot 1_{\{z\le x\}}\,F(dx)\,F(dz)\Big]^2+\Big[\iiint u_N(x,y)\cdot 1_{\{z\le x\}}\,F(dx)\,G(dy)\,F(dz)\Big]^2\Big]\\&\le 4\cdot\mathbb{E}\Big[\int\big[u_N(x,Y_1)\big]^2\,F(dx)+\iint\big[u_N(x,y)\big]^2\,F(dx)\,G(dy)\\&\quad+\iint\big[u_N(x,Y_1)\big]^2\,F(dx)\,F(dz)+\iiint\big[u_N(x,y)\big]^2\,F(dx)\,G(dy)\,F(dz)\Big]\\&\le 4\cdot\mathbb{E}\big[4\cdot\big[2a_N^{-2}\cdot\big\|K'\big\|\big]^2\big]\\&=64\big\|K'\big\|^2\cdot a_N^{-4}.\end{aligned}$$

Altogether this yields

$$\begin{aligned}\mathbb{E}\Big[U_{m,n}^2-\hat U_{m,n}^2\Big]^2&=m^{-1}n^{-1}\cdot\mathbb{E}\big[u_{2N}^*(X_1,Y_1)\big]^2\\&\le m^{-1}n^{-1}\cdot 64\big\|K'\big\|^2\cdot a_N^{-4}\\&=64\big\|K'\big\|^2\cdot a_N^{-4}\cdot m^{-1}n^{-1}\\&=O(a_N^{-4}\cdot N^{-2}).\end{aligned}\tag{5.183}$$

Lastly,

$$\begin{aligned}\frac{\lambda_N}{m} \cdot U_m^1 &= \lambda_N \cdot m^{-2}(m-1)^{-1} \cdot \sum_{1 \le i \ne j \le m} \int u_N(x, X_j) \cdot 1_{\{X_i \le x\}} \, F(dx) \\ &\le \lambda_N \cdot m^{-1} \cdot \big\|u_N\big\| \\ &\le \lambda_N \cdot m^{-1} \cdot 2\big\|K'\big\| a_N^{-2} \\ &= O(a_N^{-2} \cdot N^{-1}). \end{aligned} \tag{5.184}$$

Combining (5.182), (5.183) and (5.184) we see that (5.176) is equal to

$$\lambda_N \cdot O_P(a_N^{-2} \cdot N^{-1}) + O(a_N^{-2} \cdot N^{-1}) + (1-\lambda_N) \cdot O_P(a_N^{-2} \cdot N^{-1}) + O(a_N^{-2} \cdot N^{-1}) = O_P(a_N^{-2} \cdot N^{-1})$$

which completes the proof. □

LEMMA 5.29.

$$a_N^{-3} \cdot m^{-1} \cdot \sum_{j=1}^{m} \int K''\big(a_N^{-1}(H_N(x) - H_N(X_j))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big) \, F(dx) \cdot \big(H_N(X_j) - \hat{H}_N(X_j)\big)$$
$$= O_P(a_N^{-2} \cdot N^{-1})$$

and

$$\frac{1}{2} a_N^{-4} \cdot m^{-1} \cdot \sum_{j=1}^{m} \int K'''(\tau_j) \cdot \big(\hat{H}_N(x) - H_N(x)\big) \, F(dx) \cdot \big(H_N(X_j) - \hat{H}_N(X_j)\big)^2 = O_P(a_N^{-2} \cdot N^{-1}).$$

PROOF. For the first expression we have

$$\begin{aligned} &\Big| a_N^{-3} \cdot m^{-1} \sum_{j=1}^{m} \int K''\big(a_N^{-1}(H_N(x) - H_N(X_j))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big) \, F(dx) \cdot \big(H_N(X_j) - \hat{H}_N(X_j)\big) \Big| \\ &\le a_N^{-3} \cdot m^{-1} \sum_{j=1}^{m} \int \big|K''\big(a_N^{-1}(H_N(x) - H_N(X_j))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\big| \, F(dx) \cdot \big|H_N(X_j) - \hat{H}_N(X_j)\big| \\ &\le a_N^{-3} \cdot m^{-1} \sum_{j=1}^{m} \int \big|K''\big(a_N^{-1}(H_N(x) - H_N(X_j))\big)\big| \cdot \big\|\hat{H}_N - H_N\big\| \, F(dx) \cdot \big|H_N(X_j) - \hat{H}_N(X_j)\big| \\ &\le a_N^{-3} \cdot m^{-1} \cdot \big\|\hat{H}_N - H_N\big\|^2 \cdot \sum_{j=1}^{m} \int \big|K''\big(a_N^{-1}(H_N(x) - H_N(X_j))\big)\big| \, F(dx). \end{aligned}$$

Since $\big|K''\big|$ is bounded and equal to zero outside of (-1, 1), we can apply the bound in Lemma A.1 to obtain

$$\int \big|K''\big(a_N^{-1}(H_N(x) - H_N(X_j))\big)\big| \, F(dx) \le 2\big\|K''\big\| \cdot a_N\Big(1 + \frac{n}{m}\Big).$$

For (5.177) this gives us a bound of

$$\begin{aligned} &a_N^{-3} \cdot m^{-1} \cdot \big\|\hat{H}_N - H_N\big\|^2 \cdot \sum_{j=1}^{m} 2\big\|K''\big\| \cdot a_N\Big(1 + \frac{n}{m}\Big) \\ &= 2\big\|K''\big\| \cdot a_N^{-2}\Big(1 + \frac{n}{m}\Big) \cdot \big\|\hat{H}_N - H_N\big\|^2 \end{aligned}$$

$$
\begin{aligned}
&= 2\big\|K''\big\| \cdot a_N^{-2}\Big(1+\frac{n}{m}\Big) \cdot O_P(N^{-1}) \\
&= O_P(a_N^{-2} \cdot N^{-1})
\end{aligned}
$$

using the D-K-W bound.

For the second expression we have

$$
\begin{aligned}
&\left|\frac{1}{2}a_N^{-4} \cdot m^{-1} \cdot \sum_{j=1}^{m} \int K'''(\tau_j) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \cdot \big(H_N(X_j) - \hat{H}_N(X_j)\big)^2\right| \\
&\le \frac{1}{2}a_N^{-4} \cdot m^{-1} \cdot \sum_{j=1}^{m} \int \big|K'''(\tau_j) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\big|\, F(dx) \cdot \big|H_N(X_j) - \hat{H}_N(X_j)\big|^2 \\
&\le \frac{1}{2}a_N^{-4} \cdot m^{-1} \cdot \sum_{j=1}^{m} \int \big\|K'''\big\| \cdot \big\|\hat{H}_N - H_N\big\|\, F(dx) \cdot \big\|H_N - \hat{H}_N\big\|^2 \\
&= \frac{1}{2}\big\|K'''\big\| \cdot a_N^{-4} \cdot \big\|H_N - \hat{H}_N\big\|^3 \\
&= O_P(a_N^{-4} \cdot N^{\frac{-3}{2}})
\end{aligned}
$$

using the D-K-W bound.

Since we require that $a_N \to 0$ slowly enough that $a_N^6 \cdot N \to \infty$, the second expression is $O_P(a_N^{-2} \cdot N^{-1})$, as $a_N^{-4} \cdot N^{-\frac{3}{2}} = \big(a_N^{-2} \cdot N^{-\frac{1}{2}}\big) \cdot \big(a_N^{-2} \cdot N^{-1}\big)$, and $a_N^{-2} \cdot N^{-\frac{1}{2}} \to 0$. $\square$

To bound (5.175) we will use very similar arguments to those which we used to show that (5.174) is $O_P(a_N^{-2} \cdot N^{-1})$. We begin by deriving a sum representation of (5.175):

$$
\begin{aligned}
&\int \Big[\hat{g}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx) \\
&= a_N^{-2} \cdot n^{-1} \int \sum_{k=1}^{n} K'\big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\qquad - a_N^{-2} \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&= a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \Big[\int K'\big(a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\
&\qquad - \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Big].
\end{aligned}
$$

At this point we use the Taylor expansion of the kernel function K about each of the $a_N^{-1}(H_N(x) - \hat{H}_N(Y_k))$ which yields

$$
\begin{aligned}
&\int \Big[\hat{g}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big]\, F(dx) \\
&= a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \Big[\int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)
\end{aligned}
$$

$$- \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Big] \tag{5.185}$$

$$+ a_N^{-3} \cdot n^{-1} \cdot \sum_{k=1}^{n} \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \cdot \big(H_N(Y_k) - \hat{H}_N(Y_k)\big) \tag{5.186}$$

$$+ \frac{1}{2} a_N^{-4} \cdot n^{-1} \cdot \sum_{k=1}^{n} \int K'''(\tau_k) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \cdot \big(H_N(Y_k) - \hat{H}_N(Y_k)\big)^2 \tag{5.187}$$

where τ_k are appropriate values between the two ratios.

In the following lemmas we will derive bounds for the three terms (5.185), (5.186) and (5.187).

Lemma 5.30.

$$\begin{aligned} &a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \Big[\int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\ &- \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Big] \\ &= O_P(a_N^{-2} \cdot N^{-1}). \end{aligned}$$

Proof. Define

$$u_N(s,\, t) = a_N^{-2} \cdot \Big[K'\big(a_N^{-1}(H_N(s) - H_N(t))\big) - \int K'\big(a_N^{-1}(H_N(s) - H_N(y))\big)\, G(dy)\Big].$$

Then

$$\begin{aligned} &a_N^{-2} \cdot n^{-1} \cdot \sum_{k=1}^{n} \Big[\int K'\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big)\, F(dx) \\ &- \iint K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy)\big(\hat{H}_N(x) - H_N(x)\big)\, F(dx)\Big] \\ &= n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x, Y_k) \cdot \big[\hat{H}_N(x) - H_N(x)\big]\, F(dx) \\ &= n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x, Y_k) \cdot \Big[N^{-1}\Big[\sum_{i=1}^{m} 1_{\{X_i \le x\}} + \sum_{l=1}^{n} 1_{\{Y_l \le x\}}\Big] - H_N(x)\Big]\, F(dx) \\ &= n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x, Y_k) \cdot N^{-1} \sum_{i=1}^{m} 1_{\{X_i \le x\}}\, F(dx) \\ &\quad + n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x, Y_k) \cdot N^{-1} \sum_{l=1}^{n} 1_{\{Y_l \le x\}}\, F(dx) \\ &\quad - n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x, Y_k) \cdot H_N(x)\, F(dx) \\ &= \lambda_N \cdot n^{-1} m^{-1} \cdot \sum_{k=1}^{n} \sum_{i=1}^{m} \int u_N(x, Y_k) \cdot 1_{\{X_i \le x\}}\, F(dx) \\ &\quad + (1 - \lambda_N) \cdot n^{-2} \cdot \sum_{1 \le k \ne l \le m} \int u_N(x, Y_k) \cdot 1_{\{Y_l \le x\}}\, F(dx) \end{aligned}$$

$$+ (1-\lambda_N)\cdot n^{-2}\cdot \sum_{k=1}^{n} \int u_N(x,Y_k)\cdot 1_{\{Y_k\le x\}}\, F(dx) - n^{-1}\cdot \sum_{k=1}^{n} \int u_N(x,Y_k)\cdot H_N(x)\, F(dx).$$

Define the U-statistic U_n^1 and the generalized U-statistic $U_{m,n}^2$ as

$$U_n^1 = n^{-1}(n-1)^{-1}\cdot \sum_{1\le k\ne l\le n} \int u_N(x,Y_k)\cdot 1_{\{Y_l\le x\}}\, F(dx),$$

$$U_{m,n}^2 = n^{-1}m^{-1}\cdot \sum_{k=1}^{n}\sum_{i=1}^{m} \int u_N(x,Y_k)\cdot 1_{\{X_i\le x\}}\, F(dx),$$

and let $\hat{U}_m^1$ and $\hat{U}_{m,n}^2$ be the Hájek projections of U_m^1 and $U_{m,n}^2$ respectively as defined in Lemmas A.2 and A.3. Then (5.185) is equal to

$$\lambda_N\cdot U_{m,n}^2 + \frac{(1-\lambda_N)\cdot(n-1)}{n}\cdot U_n^1 + (1-\lambda_N)\cdot n^{-2}\cdot \sum_{k=1}^{n} \int u_N(x,Y_k)\cdot 1_{\{Y_k\le x\}}\, F(dx)$$
$$- n^{-1}\cdot \sum_{k=1}^{n} \int u_N(x,Y_k)\cdot H_N(x)\, F(dx). \tag{5.188}$$

Now, the kernel function u_N is bounded:

$$\|u_N\| \le 2\|K'\| a_N^{-2}.$$

Which means for the third sum in (5.188) we can write

$$\begin{aligned}
&\left|(1-\lambda_N)\cdot n^{-2}\cdot \sum_{k=1}^{n} \int u_N(x,Y_k)\cdot 1_{\{Y_k\le x\}}\, F(dx)\right| \\
&\quad\le (1-\lambda_N)\cdot n^{-2}\cdot \sum_{k=1}^{n} \left| \int u_N(x,Y_k)\cdot 1_{\{Y_k\le x\}}\, F(dx)\right| \\
&\quad\le (1-\lambda_N)\cdot n^{-2}\cdot \sum_{k=1}^{n} \|u_N\| \\
&\quad= (1-\lambda_N)\cdot n^{-1}\|u_N\| \\
&\quad= O(a_N^{-2}\cdot N^{-1}).
\end{aligned} \tag{5.189}$$

Thus, we can partition (5.185) into the sum of two scaled U-statistics, an i.i.d sum and a negligible rest:

$$\lambda_N\cdot U_{m,n}^2 + \frac{(1-\lambda_N)\cdot(n-1)}{n}\cdot U_n^1 - n^{-1}\cdot \sum_{k=1}^{n} \int u_N(x,Y_k)\cdot H_N(x)\, F(dx) + O(a_N^{-2}\cdot N^{-1}). \tag{5.190}$$

In the following we will show that (5.190) is $O(a_N^{-2}\cdot N^{-1})$ as well, which will complete the proof. Begin by calculating each of the projections $\hat{U}_n^1$ and $\hat{U}_{m,n}^2$. Firstly,

$$\begin{aligned}
(1-\lambda_N)\cdot \hat{U}_n^1 &= (1-\lambda_N)\cdot n^{-1}\cdot \sum_{k=1}^{n} \Bigg[\iint u_N(x,Y_k)\cdot 1_{\{y\le x\}}\, F(dx)\, G(dy) \\
&\qquad + \iint u_N(x,y)\cdot 1_{\{Y_k\le x\}}\, F(dx)\, G(dy) - \iiint u_N(x,y)\cdot 1_{\{z\le x\}}\, F(dx)\, G(dy)\, G(dz)\Bigg] \\
&= n^{-1}\cdot \sum_{k=1}^{n} \Bigg[\int u_N(x,Y_k)\cdot (1-\lambda_N) G(x)\, F(dx)
\end{aligned}$$

$$+ (1-\lambda_N) \cdot \iint u_N(x,y) \cdot 1_{\{Y_k \le x\}}\, F(dx)\, G(dy)$$
$$- \iint u_N(x,y) \cdot (1-\lambda_N) G(x)\, F(dx)\, G(dy)\Big].$$

Nextly, for $\lambda_N \cdot \hat{U}^2_{m,n}$ we have

$$\begin{aligned}
\lambda_N \cdot \hat{U}^2_{m,n} = \lambda_N \cdot \Big[& n^{-1} \cdot \sum_{k=1}^{n} \iint u_N(x,Y_k) \cdot 1_{\{y \le x\}}\, F(dx)\, F(dy) \\
& + m^{-1} \cdot \sum_{i=1}^{m} \iint u_N(x,y) \cdot 1_{\{X_i \le x\}}\, F(dx)\, G(dy) \\
& - \iiint u_N(x,z) \cdot 1_{\{y \le x\}}\, F(dx)\, F(dy)\, G(dz)\Big] \\
= n^{-1} \cdot & \sum_{k=1}^{n} \int u_N(x,Y_k) \cdot \lambda_N F(x)\, F(dx) + \lambda_N \cdot m^{-1} \cdot \sum_{i=1}^{m} \iint u_N(x,y) \cdot 1_{\{X_i \le x\}}\, F(dx)\, G(dy) \\
& - \iint u_N(x,z) \cdot \lambda_N F(x)\, F(dx)\, G(dz).
\end{aligned}$$

Now, since

$$H_N = \lambda_N \cdot F + (1-\lambda_N) \cdot G,$$

we see that

$$\begin{aligned}
& \lambda_N \cdot \hat{U}^2_{m,n} + (1-\lambda_N) \cdot \hat{U}^1_n - n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x,Y_k) \cdot H_N(x)\, F(dx) \\
&= m^{-1} \cdot \sum_{i=1}^{m} \Big[\lambda_N \cdot \iint u_N(x,y) \cdot 1_{\{X_i \le x\}}\, F(dx)\, G(dy) - \iint u_N(x,z) \cdot \lambda_N F(x)\, F(dx)\, G(dz)\Big] \\
&\quad + (1-\lambda_N) \cdot \Big[n^{-1} \cdot \sum_{k=1}^{n} \iint u_N(x,y) \cdot 1_{\{Y_k \le x\}}\, F(dx)\, G(dy) - \iint u_N(x,y) \cdot G(x)\, F(dx)\, G(dy)\Big] \\
&= 0
\end{aligned}$$

due to

$$\begin{aligned}
& \int u_N(x,y)\, G(dy) \\
&\quad = \int a_N^{-2} \cdot \Big[K'\big(a_N^{-1}(H_N(x) - H_N(y))\big) - \int K'\big(a_N^{-1}(H_N(x) - H_N(z))\big)\, G(dz)\Big] G(dy) \\
&\quad = a_N^{-2} \cdot \Big[\int K'\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dy) - \int K'\big(a_N^{-1}(H_N(x) - H_N(z))\big)\, G(dz)\Big] \\
&\quad = 0.
\end{aligned}$$

Thus, for (5.185) we have

$$\begin{aligned}
& \lambda_N \cdot U^2_{m,n} + \frac{(1-\lambda_N) \cdot (n-1)}{n} \cdot U^1_n - n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x,Y_k) \cdot H_N(x)\, F(dx) + O(a_N^{-2} \cdot N^{-1}) \\
&= \lambda_N \cdot U^2_{m,n} + (1-\lambda_N) \cdot U^1_n - \frac{1-\lambda_N}{n} \cdot U^1_n - n^{-1} \cdot \sum_{k=1}^{n} \int u_N(x,Y_k) \cdot H_N(x)\, F(dx) + O(a_N^{-2} \cdot N^{-1})
\end{aligned}$$

$$= \lambda_N \cdot \Big[U^2_{m,n} - \hat{U}^2_{m,n}\Big] + (1-\lambda_N) \cdot \Big[U^1_n - \hat{U}^1_n\Big] + \frac{1-\lambda_N}{n} \cdot U^1_n + O(a_N^{-2} \cdot N^{-1})$$

and it remains only to bound $\big[U^1_n - \hat{U}^1_n\big]$, $\big[U^2_{m,n} - \hat{U}^2_{m,n}\big]$ and $\frac{1-\lambda_N}{n} \cdot U^1_n$.

Using Lemmas A.2 and A.3 we have

$$\mathbb{E}\Big[U^1_n - \hat{U}^1_n\Big]^2 \le 2(n-1)n^{-3} \cdot \mathbb{E}\big[u^*_{1N}(Y_1, Y_2)\big]^2,$$
$$\mathbb{E}\Big[U^2_{m,n} - \hat{U}^2_{m,n}\Big]^2 = m^{-1}n^{-1} \cdot \mathbb{E}\big[u^*_{2N}(X_1, Y_1)\big]^2$$

for u^*_{1N} and u^*_{2N} defined as

$$\begin{aligned} u^*_{1N}(r,s) = &\int u_N(x,s) \cdot 1_{\{r \le x\}}\, F(dx) - \iint u_N(x,y) \cdot 1_{\{r \le x\}}\, F(dx)\, G(dy) \\ &- \iint u_N(x,s) \cdot 1_{\{y \le x\}}\, F(dx)\, G(dy) + \iiint u_N(x,y) \cdot 1_{\{z \le x\}}\, F(dx)\, G(dy)\, G(dz) \\ u^*_{2N}(r,s) = &\int u_N(x,s) \cdot 1_{\{r \le x\}}\, F(dx) - \iint u_N(x,y) \cdot 1_{\{r \le x\}}\, F(dx)\, G(dy) \\ &- \iint u_N(x,s) \cdot 1_{\{y \le x\}}\, F(dx)\, F(dy) + \iiint u_N(x,y) \cdot 1_{\{z \le x\}}\, F(dx)\, G(dy)\, F(dz). \end{aligned}$$

From (5.183) we already know

$$\begin{aligned} \mathbb{E}\Big[U^2_{m,n} - \hat{U}^2_{m,n}\Big]^2 &= m^{-1}n^{-1} \cdot \mathbb{E}\big[u^*_{2N}(X_1, Y_1)\big]^2 \\ &\le m^{-1}n^{-1} \cdot 64\big\|K'\big\|^2 \cdot a_N^{-4} \\ &= 64\big\|K'\big\|^2 \cdot a_N^{-4} \cdot m^{-1}n^{-1} \\ &= O(a_N^{-4} \cdot N^{-2}). \end{aligned} \tag{5.191}$$

The other expectation can be bounded similarly as well:

$$\begin{aligned} &\mathbb{E}\big[u^*_{1N}(Y_1, Y_2)\big]^2 \\ &\le 4 \cdot \mathbb{E}\Bigg[\Big[\int u_N(x, Y_2) \cdot 1_{\{Y_1 \le x\}}\, F(dx)\Big]^2 + \Big[\iint u_N(x,y) \cdot 1_{\{Y_1 \le x\}}\, F(dx)\, G(dy)\Big]^2 \\ &\qquad + \Big[\iint u_N(x, Y_2) \cdot 1_{\{z \le x\}}\, F(dx)\, G(dz)\Big]^2 + \Big[\iiint u_N(x,y) \cdot 1_{\{z \le x\}}\, F(dx)\, G(dy)\, G(dz)\Big]^2\Bigg] \\ &\le 4 \cdot \mathbb{E}\Big[\int \big[u_N(x, Y_2)\big]^2\, F(dx) + 0 + \iint \big[u_N(x, Y_2)\big]^2\, F(dx)\, G(dz) + 0\Big] \\ &\le 4 \cdot \mathbb{E}\Big[\int \big[2a_N^{-2} \cdot \|K'\|\,\big]^2\, F(dx) + \iint \big[2a_N^{-2} \cdot \|K'\|\,\big]^2\, F(dx)\, G(dz)\Big] \\ &= 32\, \|K'\|^2 \cdot a_N^{-4}. \end{aligned}$$

Altogether this yields

$$\begin{aligned} \mathbb{E}\Big[U^1_n - \hat{U}^1_n\Big]^2 &\le 2(n-1)n^{-3} \cdot \mathbb{E}\big[u^*_{1N}(Y_1, Y_2)\big]^2 \\ &\le 2(n-1)n^{-3} \cdot 32\big\|K'\big\|^2 \cdot a_N^{-4} \\ &= 64\big\|K'\big\|^2 \cdot a_N^{-4} \cdot (n-1)n^{-3} \end{aligned}$$

$$= O(a_N^{-4} \cdot N^{-2}). \tag{5.192}$$

Lastly,

$$\begin{aligned}
\frac{1-\lambda_N}{n} \cdot U_n^1 &= (1-\lambda_N) \cdot n^{-2}(n-1)^{-1} \cdot \sum_{1 \le k \ne l \le n} \int u_N(x, Y_k) \cdot 1_{\{Y_l \le x\}} \, F(dx) \\
&\le (1-\lambda_N) \cdot n^{-1} \cdot \|u_N\| \\
&\le (1-\lambda_N) \cdot n^{-1} \cdot 2\|K'\| a_N^{-2} \\
&= O(a_N^{-2} \cdot N^{-1}). \qquad (5.193)
\end{aligned}$$

Combining (5.192), (5.191) and (5.193) we see that (5.185) is equal to

$$\lambda_N \cdot O_P(a_N^{-2} \cdot N^{-1}) + (1-\lambda_N) \cdot O_P(a_N^{-2} \cdot N^{-1}) + O(a_N^{-2} \cdot N^{-1}) + O(a_N^{-2} \cdot N^{-1}) = O_P(a_N^{-2} \cdot N^{-1})$$

which completes the proof. □

Lemma 5.31.

$$\begin{aligned}
&a_N^{-3} \cdot n^{-1} \cdot \sum_{k=1}^{n} \int K''\big(a_N^{-1}(H_N(x) - H_N(Y_k))\big) \cdot \big(\hat{H}_N(x) - H_N(x)\big) \, F(dx) \cdot \big(H_N(Y_k) - \hat{H}_N(Y_k)\big) \\
&= O_P(a_N^{-2} \cdot N^{-1})
\end{aligned}$$

and

$$\frac{1}{2} a_N^{-4} \cdot n^{-1} \cdot \sum_{k=1}^{n} \int K'''(\tau_k) \cdot \big(\hat{H}_N(x) - H_N(x)\big) \, F(dx) \cdot \big(H_N(Y_k) - \hat{H}_N(Y_k)\big)^2 = O_P(a_N^{-2} \cdot N^{-1}).$$

Proof. The proof is completely analogous to the proof of Lemma 5.29 with m, j and X_j replaced by n, k and Y_k repectively. □

Combining Lemmas 5.28, 5.29,5.30, and 5.31 we have proven the following.

Lemma 5.32.

$$\int \Big[\hat{f}_N - \bar{f}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big] \, F(dx) = O_P(a_N^{-2} \cdot N^{-1}), \tag{5.194}$$

$$\int \Big[\hat{g}_N - \bar{g}_N\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big] \, F(dx) = O_P(a_N^{-2} \cdot N^{-1}) \tag{5.195}$$

and thus

$$\int \Big[\hat{f}_N - \hat{g}_N - (\bar{f}_N - \bar{g}_N)\Big]' \circ H_N(x) \cdot \Big[\hat{H}_N(x) - H_N(x)\Big] \, F(dx) = O_P(a_N^{-2} \cdot N^{-1}). \tag{5.196}$$

5.3. Asymptotic variance under H_0

The lemmas in this section deal with the expectations that determine the asymptotic variance of the test statistic under H_0. The asymptotic variance under H_0 does not depend on the underlying distributions F and G, which is as we would expect when dealing with rank statistics. In this sense the test is distribution free. It is interesting to note, that the variance terms dealt with in the following lemmas do depend, however, on the bandwidth a_N and the choice of the kernel function K.

LEMMA 5.33. *Let F and G be continuous distribution functions and H_N be defined as in* (2.4). *Further, let K be a kernel on* $(-1, 1)$ *satisfying* (2.8) *through* (2.11) *then under* $\mathsf{H}_0 : F = G$

$$\mathbb{E}\left[a_N^{-1} \cdot \int K\left(a_N^{-1}(H_N(x) - H_N(X_1))\right) F(dx)\right]^2$$
$$= 1 + 2 \cdot a_N \int_{-1}^{0} \left[\int_{-1}^{x} K(v)\, dv\right]^2 dx - 4 \cdot a_N \int_0^1 vK(v)\, dv. \tag{5.197}$$

PROOF. Firstly, we define a funtion φ as the squared antiderivative of the Kernel K:

$$\varphi(x) = \left[\int_{-1}^{x} K(v)\, dv\right]^2. \tag{5.198}$$

Now, recall that under H_0: $F = G$ we have $H_N = \frac{m}{N}F + \frac{n}{N}G = F = G$ for all N, so that the $H_N(X_1)$ are uniformly distributed on the intervall $(0, 1)$, and have density $f_N = 1_{(0,1)}$. Then

$$\begin{aligned}
&\mathbb{E}\left[a_N^{-1} \cdot \int K\left(a_N^{-1}(H_N(x) - H_N(X_1))\right) F(dx)\right]^2 \\
&= \int \left[a_N^{-1} \cdot \int K\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx)\right]^2 F(dy) \\
&= \int \left[a_N^{-1} \cdot \int_0^1 K\left(a_N^{-1}(w - H_N(y))\right) dw\right]^2 F(dy) \\
&= \int \left[\int_{-a_N^{-1}H_N(y)}^{a_N^{-1}(1-H_N(y))} K(v)\, dv\right]^2 F(dy) \\
&= \int_0^1 \left[\int_{-a_N^{-1}w}^{a_N^{-1}(1-w)} K(v)\, dv\right]^2 dw \\
&= \int_0^1 \left[1 - 1_{\{a_N^{-1}(1-w)<1\}} \cdot \int_{a_N^{-1}(1-w)}^{1} K(v)\, dv - 1_{\{-a_N^{-1}w>-1\}} \cdot \int_{-1}^{-a_N^{-1}w} K(v)\, dv\right]^2 dw \\
&= \int_0^1 \Bigg[1 + 1_{\{a_N^{-1}(1-w)<1\}} \cdot \left[\int_{a_N^{-1}(1-w)}^{1} K(v)\, dv\right]^2 + 1_{\{-a_N^{-1}w>-1\}} \cdot \left[\int_{-1}^{-a_N^{-1}w} K(v)\, dv\right]^2 \\
&\quad - 2 \cdot 1_{\{a_N^{-1}(1-w)<1\}} \cdot \int_{a_N^{-1}(1-w)}^{1} K(v)\, dv - 2 \cdot 1_{\{-a_N^{-1}w>-1\}} \cdot \int_{-1}^{-a_N^{-1}w} K(v)\, dv \\
&\quad + 2 \cdot 1_{\{a_N^{-1}(1-w)<1\}} \cdot 1_{\{-a_N^{-1}w>-1\}} \cdot \int_{a_N^{-1}(1-w)}^{1} K(v)\, dv \cdot \int_{-1}^{-a_N^{-1}w} K(v)\, dv\Bigg] dw.
\end{aligned}$$

We require of our bandwidth sequence that $a_N < \frac{1}{2}\ \forall\ N$, so that $1_{\left\{a_N^{-1}(1-w)<1\right\}} \cdot 1_{\left\{-a_N^{-1}w>-1\right\}}$ vanishes, leaving

$$
\begin{aligned}
&\int_0^1 dw + \int_0^1 1_{\{w>1-a_N\}} \cdot \left[\int_{a_N^{-1}(1-w)}^1 K(v)\,dv\right]^2 dw + \int_0^1 1_{\{w<a_N\}} \cdot \left[\int_{-1}^{-a_N^{-1}w} K(v)\,dv\right]^2 dw \\
&\quad - 2\cdot \int_0^1 1_{\{w>1-a_N\}} \cdot \int_{a_N^{-1}(1-w)}^1 K(v)\,dv\,dw - 2\cdot \int_0^1 1_{\{w<a_N\}} \cdot \int_{-1}^{-a_N^{-1}w} K(v)\,dv\,dw \\
&= 1 + \int_{1-a_N}^1 \left[\int_{-1}^{a_N^{-1}(w-1)} K(v)\,dv\right]^2 dw + \int_0^{a_N} \left[\int_{-1}^{-a_N^{-1}w} K(v)\,dv\right]^2 dw \\
&\quad - 2\cdot \int_{-1}^1 K(v) \int_0^1 1_{\{w>1-a_N\}} \cdot 1_{\left\{a_N^{-1}(1-w)<v\right\}}\,dw\,dv \\
&\quad - 2\cdot \int_{-1}^1 K(v) \int_0^1 1_{\{w<a_N\}} \cdot 1_{\left\{v<-a_N^{-1}w\right\}}\,dw\,dv \\
&= 1 + \int_{1-a_N}^1 \varphi(a_N^{-1}(w-1))\,dw + \int_0^{a_N} \varphi(-a_N^{-1}w)\,dw \\
&\quad - 2\cdot \int_0^1 K(v) \int_0^1 1_{\{w>1-a_N v\}} dw\,dv - 2\cdot \int_0^1 K(v) \int_0^1 1_{\{w<a_N v\}} dw\,dv \\
&= 1 + \int_{-1}^0 a_N \cdot \varphi(x)\,dx - \int_0^{-1} a_N \cdot \varphi(x)\,dx \\
&\quad - 2\cdot a_N \cdot \int_0^1 vK(v)\,dv - 2\cdot a_N \cdot \int_0^1 vK(v)\,dv \\
&= 1 + 2\cdot a_N \int_{-1}^0 \varphi(x)\,dx - 4\cdot a_N \cdot \int_0^1 vK(v)\,dv \\
&= 1 + 2\cdot a_N \int_{-1}^0 \left[\int_{-1}^x K(v)\,dv\right]^2 dx - 4\cdot a_N \int_0^1 vK(v)\,dv.
\end{aligned}
$$

□

LEMMA 5.34. *Let F and G be continuous distribution functions and H_N be defined as in (2.4). Further, let K be a kernel on $(-1,\,1)$ satisfying (2.8) through (2.11) then under $\mathsf{H}_0\,:\,F=G$*

$$
\begin{aligned}
&\left[a_N^{-1} \cdot \iint K\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx)F(dy)\right]^2 \\
&\qquad = 1 - 4\cdot a_N \int_0^1 vK(v)\,dv + 4\cdot a_N^2 \left[\int_0^1 vK(v)\,dv\right]^2.
\end{aligned}
$$

PROOF.

$$
\begin{aligned}
&\left[a_N^{-1} \cdot \iint K\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx)F(dy)\right]^2 \\
&\quad = \left[a_N^{-1} \cdot \int \int_0^1 K\left(a_N^{-1}(w - H_N(y))\right) dw\,F(dy)\right]^2 \\
&\quad = \left[\int \int_{-a_N^{-1}H_N(y)}^{a_N^{-1}(1-H_N(y))} K(v)\,dv\,F(dy)\right]^2
\end{aligned}
$$

$$= \left[\int_0^1 \int_{-a_N^{-1}w}^{a_N^{-1}(1-w)} K(v)\,dv\,dw\right]^2$$

$$= \left[\int_0^1 \int_{-1}^1 1_{\left\{-a_N^{-1}w<v<a_N^{-1}(1-w)\right\}} K(v)\,dv\,dw\right]^2$$

$$= \left[\int_{-1}^1 K(v) \int_0^1 1_{\{-a_N v<w<v<1-a_N v\}}\,dw\,dv\right]^2$$

$$= \left[\int_{-1}^0 K(v) \int_0^1 1_{\{-a_N v<w\}}\,dw\,dv + \int_0^1 K(v) \int_0^1 1_{\{w<v<1-a_N v\}}\,dw\,dv\right]^2$$

$$= \left[\int_{-1}^0 K(v)(1+a_N v)\,dv + \int_0^1 K(v)(1-a_N v)\,dv\right]^2$$

$$= \left[\int_{-1}^0 K(v)\,dv + a_N \cdot \int_{-1}^0 vK(v)\,dv + \int_0^1 K(v)\,dv - a_N \cdot \int_0^1 vK(v)\,dv\right]^2$$

$$= \left[\int_{-1}^1 K(v)\,dv + a_N \cdot \left[\int_{-1}^0 vK(v)\,dv - \int_0^1 vK(v)\,dv\right]\right]^2$$

$$= \left[1 - 2 \cdot a_N \cdot \int_0^1 vK(v)\,dv\right]^2$$

$$= 1 - 4 \cdot a_N \int_0^1 vK(v)\,dv + 4 \cdot a_N^2 \left[\int_0^1 vK(v)\,dv\right]^2.$$

□

LEMMA 5.35. *Let F and G be continuous distribution functions and H_N be defined as in (2.4). Further, let K be a kernel on $(-1,\,1)$ satisfying (2.8) through (2.11) and let $0 < a_N < \frac{1}{2}$.*

Then for

$$u_N(s,t) = a_N^{-1} K\big(a_N^{-1}(H_N(s) - H_N(t))\big) - a_N^{-1} \int K\big(a_N^{-1}(H_N(x) - H_N(t))\big)\,F(dx),$$

under $\mathsf{H}_0 : F = G$ we have

$$\mathsf{Var}\left[m^{-1}(m-1)^{-1} \cdot \sum_{1\le i\ne j\le m} u_N(X_i, X_j) - m^{-1}n^{-1} \cdot \sum_{i=1}^m \sum_{k=1}^n u_N(X_i, Y_k)\right]$$

$$= m^{-1}(m-1)^{-1}\Bigg[\left[a_N^{-1} \int_{-1}^1 K^2(v)\,dv - 2\int_0^1 vK^2(v)\,dv\right]$$

$$+ (2n+m-1)n^{-1}\left[1 - 4 \cdot a_N \int_0^1 vK(v)\,dv + 4 \cdot a_N^2 \left[\int_0^1 vK(v)\,dv\right]^2\right]$$

$$- (1+2n^{-1})\left[1 + 2 \cdot a_N \int_{-1}^0 \left[\int_{-1}^x K(v)\,dv\right]^2 dx - 4 \cdot a_N \int_0^1 vK(v)\,dv\right]\Bigg].$$

PROOF.

$$\mathsf{Var}\left[m^{-1}(m-1)^{-1} \cdot \sum_{1\le i\ne j\le m} u_N(X_i, X_j) - m^{-1}n^{-1} \cdot \sum_{i=1}^m \sum_{k=1}^n u_N(X_i, Y_k)\right]$$

$$= \mathsf{Var}\left[m^{-1}(m-1)^{-1} \cdot \sum_{1\le i\ne j\le m} u_N(X_i, X_j)\right]$$

$$- 2 \cdot \mathsf{Cov}\Big[m^{-1}(m-1)^{-1} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j),\ m^{-1}n^{-1} \cdot \sum_{\substack{1 \le i \le m \\ 1 \le k \le n}} u_N(X_i, Y_k)\Big]$$

$$+ \mathsf{Var}\Big[m^{-1}n^{-1} \cdot \sum_{\substack{1 \le i \le m \\ 1 \le k \le n}} u_N(X_i, Y_k)\Big]$$

$$= \mathsf{E}\Big[m^{-1}(m-1)^{-1} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j)\Big]^2 \tag{5.199}$$

$$- 2 \cdot \mathsf{E}\Big[\Big[m^{-1}(m-1)^{-1} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j)\Big] \cdot \Big[m^{-1}n^{-1} \cdot \sum_{\substack{1 \le i \le m \\ 1 \le k \le n}} u_N(X_i, Y_k)\Big]\Big] \tag{5.200}$$

$$+ \mathsf{E}\Big[m^{-1}n^{-1} \cdot \sum_{\substack{1 \le i \le m \\ 1 \le k \le n}} u_N(X_i, Y_k)\Big]^2 \tag{5.201}$$

In the following we will expand each of the three expectations (5.199), (5.200) and (5.201) and combine the results to get a simpler expression for the overall variance. Beginning with (5.199) we get

$$\mathsf{E}\Big[m^{-1}(m-1)^{-1} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i, X_j)\Big]^2$$

$$= m^{-2}(m-1)^{-2}\Big[m(m-1)(m-2)(m-3) \cdot \mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_3, X_4)\big] \tag{5.202}$$

$$+ m(m-1)(m-2) \cdot \Big[\mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_1, X_3)\big] \tag{5.203}$$

$$+ 2 \cdot \mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_3, X_1)\big] \tag{5.204}$$

$$+ \mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_3, X_2)\big]\Big] \tag{5.205}$$

$$+ m(m-1) \cdot \Big[\mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_2, X_1)\big] \tag{5.206}$$

$$+ \mathbb{E}\big[u_N(X_1, X_2)\big]^2\Big]\Big]. \tag{5.207}$$

Clearly, $\mathbb{E}\big[u_N(X_1, X_2)\big] = 0$, so (5.202) vanishes immediately. Additionally, the expectation in (5.204) vanishes due to

$$\mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_3, X_1)\big] = \mathbb{E}\big[\mathbb{E}\big[u_N(X_1, X_2) \mid X_1\big] \cdot \mathbb{E}\big[u_N(X_3, X_1) \mid X_1\big]\big],$$

since for the inner expectation

$$\mathbb{E}\big[u_N(X_3, X_1) \,|\, X_1\big]$$

$$= a_N^{-1} \int K\big(a_N^{-1}(H_N(x) - H_N(X_1))\big)\, F(dx) - a_N^{-1} \int K\big(a_N^{-1}(H_N(x) - H_N(X_1))\big)\, F(dx)$$

$$= 0. \tag{5.208}$$

(5.208) implies directly that (5.205) vanishes as well, since

$$\mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_3, X_2)\big] = \mathbb{E}\big[\mathbb{E}\big[u_N(X_1, X_2) \mid X_2\big] \cdot \mathbb{E}\big[u_N(X_3, X_2) \mid X_2\big]\big],$$

so that we have

$$\begin{aligned}
&\mathsf{E}\Big[m^{-1}(m-1)^{-1}\cdot\sum_{1\le i\ne j\le m}u_N(X_i,X_j)\Big]^2\\
&= m^{-2}(m-1)^{-2}\Big[m(m-1)(m-2)\cdot\mathbb{E}\big[u_N(X_1,X_2)\cdot u_N(X_1,X_3)\big]\\
&\quad + m(m-1)\cdot\Big[\mathbb{E}\big[u_N(X_1,X_2)\cdot u_N(X_2,X_1)\big]\\
&\quad + \mathbb{E}\big[u_N(X_1,X_2)\big]^2\Big]\Big].
\end{aligned}$$

Now, for the second expectation (5.200) we have

$$\mathsf{E}\Big[\Big[m^{-1}(m-1)^{-1}\cdot\sum_{1\le i\ne j\le m}u_N(X_i,X_j)\Big]\cdot\Big[m^{-1}n^{-1}\cdot\sum_{\substack{1\le i\le m\\1\le k\le n}}u_N(X_i,Y_k)\Big]\Big]$$

$$= m^{-2}(m-1)^{-1}n^{-1}\Big[m(m-1)(m-2)n\cdot\mathbb{E}\big[u_N(X_1,X_2)\cdot u_N(X_3,Y_1)\big] \tag{5.209}$$

$$+ m(m-1)n\cdot\Big[\mathbb{E}\big[u_N(X_1,X_2)\cdot u_N(X_1,Y_1)\big] \tag{5.210}$$

$$+ \mathbb{E}\big[u_N(X_1,X_2)\cdot u_N(X_2,Y_1)\big]\Big]\Big] \tag{5.211}$$

Clearly, $\mathbb{E}\big[u_N(X_1,X_2)\big]=0$, so (5.209) vanishes immediately. Additionally, the expectation in (5.211) vanishes due to

$$\mathbb{E}\big[u_N(X_1,X_2)\cdot u_N(X_2,Y_1)\big]=\mathbb{E}\big[\mathbb{E}\big[u_N(X_1,X_2)\mid X_2\big]\cdot\mathbb{E}\big[u_N(X_2,Y_1)\mid X_2\big]\big],$$

since we know from (5.208) that the second inner expectation vanishes, so that we have

$$\begin{aligned}
&\mathsf{E}\Big[\Big[m^{-1}(m-1)^{-1}\cdot\sum_{1\le i\ne j\le m}u_N(X_i,X_j)\Big]\cdot\Big[m^{-1}n^{-1}\cdot\sum_{\substack{1\le i\le m\\1\le k\le n}}u_N(X_i,Y_k)\Big]\Big]\\
&= m^{-2}(m-1)^{-1}n^{-1}\cdot m(m-1)n\cdot\mathbb{E}\big[u_N(X_1,X_2)\cdot u_N(X_1,Y_1)\big]\\
&= m^{-1}\cdot\mathbb{E}\big[u_N(X_1,X_2)\cdot u_N(X_1,Y_1)\big].
\end{aligned}$$

Lastly, for the third expectation (5.201) we have

$$\mathsf{E}\Big[m^{-1}n^{-1}\cdot\sum_{\substack{1\le i\le m\\1\le k\le n}}u_N(X_i,Y_k)\Big]^2$$

$$= m^{-2}n^{-2}\Big[m(m-1)n(n-1)\cdot\mathbb{E}\big[u_N(X_1,Y_1)\cdot u_N(X_2,Y_2)\big] \tag{5.212}$$

$$+ mn(n-1)\cdot\mathbb{E}\big[u_N(X_1,Y_1)\cdot u_N(X_1,Y_2)\big] \tag{5.213}$$

$$+ m(m-1)n\cdot\mathbb{E}\big[u_N(X_1,Y_1)\cdot u_N(X_2,Y_1)\big] \tag{5.214}$$

$$+ mn\cdot\mathbb{E}\big[u_N(X_1,Y_1)\big]^2\Big]. \tag{5.215}$$

Clearly, $\mathbb{E}\big[u_N(X_1,Y_1)\big]=0$, so (5.212) vanishes immediately. Additionally, the expectation in (5.214) vanishes due to

$$\mathbb{E}\big[u_N(X_1,Y_1)\cdot u_N(X_2,Y_1)\big]=\mathbb{E}\big[\mathbb{E}\big[u_N(X_1,Y_1)\mid Y_1\big]\cdot\mathbb{E}\big[u_N(X_2,Y_1)\mid X_1\big]\big],$$

since for the inner expectation

$$\begin{aligned}
&\mathbb{E}\big[u_N(X_1, Y_1) \mid Y_1\big] \\
&= a_N^{-1} \int K\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)\, F(dx) - a_N^{-1} \int K\big(a_N^{-1}(H_N(x) - H_N(Y_1))\big)\, F(dx) \\
&= 0.
\end{aligned}$$

Then altogether for the expectation (5.201) we have

$$\begin{aligned}
\mathsf{E}\Big[m^{-1}n^{-1} \cdot \sum_{\substack{1\le i\le m\\ 1\le k\le n}} u_N(X_i, Y_k)\Big]^2 = m^{-2}n^{-2}\Big[&mn(n-1) \cdot \mathbb{E}\big[u_N(X_1, Y_1) \cdot u_N(X_1, Y_2)\big] \\
&+ mn \cdot \mathbb{E}\big[u_N(X_1, Y_1)\big]^2\Big].
\end{aligned}$$

At this point, we can combine our expressions for the expectations (5.199), (5.200) and (5.201) to get

$$\begin{aligned}
&\mathsf{Var}\Big[m^{-1}(m-1)^{-1} \cdot \sum_{1\le i\ne j\le m} u_N(X_i, X_j) - m^{-1}n^{-1} \cdot \sum_{i=1}^{m}\sum_{k=1}^{n} u_N(X_i, Y_k)\Big] \\
&= m^{-2}(m-1)^{-2}\Big[m(m-1)(m-2) \cdot \mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_1, X_3)\big] \\
&\quad + m(m-1) \cdot \Big[\mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_2, X_1)\big] \\
&\quad\quad + \mathbb{E}\big[u_N(X_1, X_2)\big]^2\Big]\Big] \\
&\quad - 2 \cdot m^{-1} \cdot \mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_1, Y_1)\big] \\
&\quad + m^{-2}n^{-2}\Big[mn(n-1) \cdot \mathbb{E}\big[u_N(X_1, Y_1) \cdot u_N(X_1, Y_2)\big] \\
&\quad\quad + mn \cdot \mathbb{E}\big[u_N(X_1, Y_1)\big]^2\Big],
\end{aligned}$$

which, under H_0, simplifies to

$$\begin{aligned}
&m^{-1}(m-1)^{-1}(m-2) \cdot \mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_1, X_3)\big] \\
&+ m^{-1}(m-1)^{-1} \cdot \mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_2, X_1)\big] \\
&+ m^{-1}(m-1)^{-1} \cdot \mathbb{E}\big[u_N(X_1, X_2)\big]^2 \\
&- 2 \cdot m^{-1} \cdot \mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_1, X_3)\big] \\
&+ m^{-1}n^{-1}(n-1) \cdot \mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_1, X_3)\big] \\
&+ m^{-1}n^{-1} \cdot \mathbb{E}\big[u_N(X_1, X_2)\big]^2 \\
&\quad = m^{-1}((m-1)^{-1}(m-2) + n^{-1}(n-1) - 2) \cdot \mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_1, X_3)\big] && (5.216) \\
&\quad\quad + m^{-1}((m-1)^{-1} + n^{-1}) \cdot \mathbb{E}\big[u_N(X_1, X_2)\big]^2 && (5.217) \\
&\quad\quad + m^{-1}(m-1)^{-1} \cdot \mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_2, X_1)\big]. && (5.218)
\end{aligned}$$

At this point introduce the function

$$v_N(s, t) = a_N^{-1} K\big(a_N^{-1}(H_N(s) - H_N(t))\big),$$

which is symmetric in its two arguments s and t, since we require K to be a symmetric kernel.

Then we may write the expectation in (5.216) as

$$\begin{aligned}
&\mathbb{E}\big[u_N(X_1,X_2)\cdot u_N(X_1,X_3)\big]\\
&\quad= \mathbb{E}\Big[\Big[v_N(X_1,X_2) - \mathbb{E}\big[v_N(X_1,X_2)\mid X_2\big]\Big]\cdot\Big[v_N(X_1,X_3) - \mathbb{E}\big[v_N(X_1,X_3)\mid X_3\big]\Big]\Big]\\
&\quad= \mathbb{E}\big[v_N(X_1,X_2)\cdot v_N(X_1,X_2)\big] - \mathbb{E}\Big[v_N(X_1,X_2)\cdot\mathbb{E}\big[v_N(X_1,X_3)\mid X_3\big]\Big]\\
&\qquad - \mathbb{E}\Big[v_N(X_1,X_3)\cdot\mathbb{E}\big[v_N(X_1,X_2)\mid X_2\big]\Big] + \mathbb{E}\Big[\mathbb{E}\big[v_N(X_1,X_2)\mid X_2\big]\cdot\mathbb{E}\big[v_N(X_1,X_3)\mid X_3\big]\Big]\\
&\quad= \mathbb{E}\big[v_N(X_1,X_2)\big]^2 - \Big[\mathbb{E}\big[v_N(X_1,X_2)\big]\Big]^2.
\end{aligned}$$

Using the symmetry in v_N the expectation in (5.217) becomes

$$\begin{aligned}
&\mathbb{E}\big[u_N(X_1,X_2)\big]^2\\
&\quad= \mathbb{E}\Big[v_N(X_1,X_2) - \mathbb{E}\big[v_N(X_1,X_2)\mid X_2\big]\Big]^2\\
&\quad= \mathbb{E}\big[v_N(X_1,X_2)\big]^2 - 2\cdot\mathbb{E}\Big[v_N(X_1,X_2)\cdot\mathbb{E}\big[v_N(X_1,X_2)\mid X_2\big]\Big] + \mathbb{E}\Big[\mathbb{E}\big[v_N(X_1,X_2)\mid X_2\big]\Big]^2\\
&\quad= \mathbb{E}\big[v_N(X_1,X_2)\big]^2 - \mathbb{E}\Big[\mathbb{E}\big[v_N(X_1,X_2)\mid X_2\big]\Big]^2\\
&\quad= \mathbb{E}\big[v_N(X_1,X_2)\big]^2 - \mathbb{E}\Big[\mathbb{E}\big[v_N(X_1,X_2)\mid X_2\big]\cdot\mathbb{E}\big[v_N(X_1,X_2)\mid X_2\big]\Big]\\
&\quad= \mathbb{E}\big[v_N(X_1,X_2)\big]^2 - \mathbb{E}\Big[\mathbb{E}\big[v_N(X_1,X_2)\mid X_2\big]\cdot\mathbb{E}\big[v_N(X_3,X_2)\mid X_2\big]\Big]\\
&\quad= \mathbb{E}\big[v_N(X_1,X_2)\big]^2 - \mathbb{E}\big[v_N(X_1,X_2)\cdot v_N(X_3,X_2)\big]\\
&\quad= \mathbb{E}\big[v_N(X_1,X_2)\big]^2 - \mathbb{E}\big[v_N(X_1,X_2)\cdot v_N(X_1,X_3)\big]
\end{aligned}$$

Finally, the expectation in (5.218) becomes

$$\begin{aligned}
&\mathbb{E}\big[u_N(X_1,X_2)\cdot u_N(X_2,X_1)\big]\\
&\quad= \mathbb{E}\Big[\Big[v_N(X_1,X_2) - \mathbb{E}\big[v_N(X_1,X_2)\mid X_2\big]\Big]\cdot\Big[v_N(X_2,X_1) - \mathbb{E}\big[v_N(X_2,X_1)\mid X_1\big]\Big]\Big]\\
&\quad= \mathbb{E}\big[v_N(X_1,X_2)\cdot v_N(X_2,X_1)\big] - \mathbb{E}\Big[v_N(X_1,X_2)\cdot\mathbb{E}\big[v_N(X_2,X_1)\mid X_1\big]\Big]\\
&\qquad - \mathbb{E}\Big[v_N(X_2,X_1)\cdot\mathbb{E}\big[v_N(X_1,X_2)\mid X_2\big]\Big] + \mathbb{E}\Big[\mathbb{E}\big[v_N(X_1,X_2)\mid X_2\big]\cdot\mathbb{E}\big[v_N(X_2,X_1)\mid X_1\big]\Big]\\
&\quad= \mathbb{E}\big[v_N(X_1,X_2)\big]^2 - 2\cdot\mathbb{E}\Big[\mathbb{E}\big[v_N(X_1,X_2)\mid X_1\big]\cdot\mathbb{E}\big[v_N(X_2,X_1)\mid X_1\big]\Big] + \Big[\mathbb{E}\big[v_N(X_1,X_2)\big]\Big]^2\\
&\quad= \mathbb{E}\big[v_N(X_1,X_2)\big]^2 - 2\cdot\mathbb{E}\big[v_N(X_1,X_2)\cdot v_N(X_1,X_3)\big] + \Big[\mathbb{E}\big[v_N(X_1,X_2)\big]\Big]^2
\end{aligned}$$

Taken together, this means that we can write the variance we are interested in as a function of the three fairly simple expectations $\mathbb{E}[v_N(X_1,X_2)]^2$, $\mathbb{E}[v_N(X_1,X_2)\cdot v_N(X_1,X_3)]$ and $[\mathbb{E}[v_N(X_1,X_2)]]^2$.

Under H_0 the integral

$$\mathbb{E}\big[v_N(X_1,X_2)\cdot v_N(X_1,X_3)\big] = \mathbb{E}\Big[\mathbb{E}\big[v_N(X_1,X_2)\mid X_1\big]\cdot\mathbb{E}\big[v_N(X_1,X_3)\mid X_1\big]\Big]$$

$$
\begin{aligned}
&= \mathbb{E}\Big[\mathbb{E}\big[v_N(X_1,X_2) \mid X_1\big] \cdot \mathbb{E}\big[v_N(X_1,X_2) \mid X_1\big]\Big] \\
&= \mathbb{E}\Big[\mathbb{E}\big[v_N(X_1,X_2) \mid X_1\big]\Big]^2 \\
&= \mathbb{E}\Big[\mathbb{E}\big[v_N(X_2,X_1) \mid X_1\big]\Big]^2 \\
&= \mathbb{E}\Big[a_N^{-1} \int K\left(a_N^{-1}(H_N(x) - H_N(X_2))\right) F(dx)\Big]^2
\end{aligned}
$$

has already been dealt with above in lemma 5.33.

Likewise, the integral

$$
\Big[\mathbb{E}\big[v_N(X_1,X_2)\big]\Big]^2 = \Big[a_N^{-1} \iint K\left(a_N^{-1}(H_N(x) - H_N(y))\right) F(dx)F(dy)\Big]^2
$$

has already been handled above in lemma 5.34.

Thus it remains only to evaluate the remaining expectation:

$$
\begin{aligned}
&\mathbb{E}\big[v_N(X_1,X_2)\big]^2 \\
&\quad = \mathbb{E}\Big[a_N^{-1} K\big(a_N^{-1}(H_N(X_1) - H_N(X_2))\big)\Big]^2 \\
&\quad = a_N^{-2} \iint K^2\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dx)F(dy) \\
&\quad = a_N^{-2} \int_0^1 \int_0^1 K^2\big(a_N^{-1}(H_N(v) - H_N(w))\big)\, dv\, dw \\
&\quad = a_N^{-1} \int_0^1 \int_{-a_N^{-1}w}^{a_N^{-1}(1-w)} K^2(u)\, du\, dw \\
&\quad = a_N^{-1} \int_{-1}^1 \int_0^1 1_{\{-a_N \cdot u < w < 1 - a_N \cdot u\}} K^2(u)\, dw\, du \\
&\quad = a_N^{-1} \cdot \Big[\int_{-1}^0 \int_0^1 1_{\{-a_N \cdot u < w\}}\, dw\, K^2(u)\, du + \int_0^1 \int_0^1 1_{\{w < 1 - a_N \cdot u\}}\, dw\, K^2(u)\, du\Big] \\
&\quad = a_N^{-1} \cdot \Big[\int_{-1}^0 (1 + a_N \cdot u) K^2(u)\, du + \int_0^1 (1 - a_N \cdot u) K^2(u)\, du\Big] \\
&\quad = a_N^{-1} \cdot \Big[\int_{-1}^0 K^2(u)\, du + a_N \int_{-1}^0 u K^2(u)\, du + \int_0^1 K^2(u)\, du - a_N \int_0^1 u K^2(u)\, du\Big] \\
&\quad = a_N^{-1} \cdot \Big[\int_{-1}^1 K^2(u)\, du - 2 \cdot a_N \int_0^1 u K^2(u)\, du\Big] \\
&\quad = a_N^{-1} \int_{-1}^1 K^2(u)\, du - 2 \int_0^1 u K^2(u)\, du.
\end{aligned}
$$

Combining the integrals we have calculated for the expectations in (5.216), (5.217) and (5.218), we get the simplified representation of the total variance as claimed.

$$
\begin{aligned}
&\mathsf{Var}\Big[m^{-1}(m-1)^{-1} \cdot \sum_{1 \le i \ne j \le m} u_N(X_i,X_j) - m^{-1}n^{-1} \cdot \sum_{i=1}^m \sum_{k=1}^n u_N(X_i,Y_k)\Big] \\
&\quad = m^{-1}((m-1)^{-1}(m-2) + n^{-1}(n-1) - 2) \cdot \mathbb{E}\big[u_N(X_1,X_2) \cdot u_N(X_1,X_3)\big]
\end{aligned}
$$

$$
\begin{aligned}
&+ m^{-1}((m-1)^{-1} + n^{-1}) \cdot \mathbb{E}\big[u_N(X_1, X_2)\big]^2 \\
&+ m^{-1}(m-1)^{-1} \cdot \mathbb{E}\big[u_N(X_1, X_2) \cdot u_N(X_2, X_1)\big] \\
= {}& m^{-1}((m-1)^{-1}(m-2) + n^{-1}(n-1) - 2) \cdot \Big[\mathbb{E}\big[v_N(X_1, X_2)\big]^2 - \Big[\mathbb{E}\big[v_N(X_1, X_2)\big]\Big]^2\Big] \\
&+ m^{-1}((m-1)^{-1} + n^{-1}) \cdot \Big[\mathbb{E}\big[v_N(X_1, X_2)\big]^2 - \mathbb{E}\big[v_N(X_1, X_2) \cdot v_N(X_1, X_3)\big]\Big] \\
&+ m^{-1}(m-1)^{-1} \cdot \Big[\mathbb{E}\big[v_N(X_1, X_2)\big]^2 - 2 \cdot \mathbb{E}\big[v_N(X_1, X_2) \cdot v_N(X_1, X_3)\big] + \Big[\mathbb{E}\big[v_N(X_1, X_2)\big]\Big]^2\Big] \\
= {}& m^{-1}\big((m-1)^{-1}(m-2) + n^{-1}(n-1) - 2 + (m-1)^{-1} + n^{-1} + (m-1)^{-1}\big) \cdot \mathbb{E}\big[v_N(X_1, X_2)\big]^2 \\
&+ m^{-1}\big((m-1)^{-1} - (m-1)^{-1}(m-2) - n^{-1}(n-1) + 2\big) \cdot \Big[\mathbb{E}\big[v_N(X_1, X_2)\big]\Big]^2 \\
&- m^{-1}\big((m-1)^{-1} + 2 \cdot n^{-1}(m-1)^{-1}\big) \cdot \mathbb{E}\big[v_N(X_1, X_2) \cdot v_N(X_1, X_3)\big] \\
= {}& m^{-1}(m-1)^{-1} \cdot \mathbb{E}\big[v_N(X_1, X_2)\big]^2 \\
&+ m^{-1}\big((m-1)^{-1}(3-m) - n^{-1}(n-1) + 2\big) \cdot \Big[\mathbb{E}\big[v_N(X_1, X_2)\big]\Big]^2 \\
&- m^{-1}\big((m-1)^{-1}(1 + 2 \cdot n^{-1})\big) \cdot \mathbb{E}\big[v_N(X_1, X_2) \cdot v_N(X_1, X_3)\big] \\
= {}& m^{-1}(m-1)^{-1} \cdot \Big[\mathbb{E}\big[v_N(X_1, X_2)\big]^2 \\
&\quad + \big((3-m) - (m-1)n^{-1}(n-1) + 2(m-1)\big) \cdot \Big[\mathbb{E}\big[v_N(X_1, X_2)\big]\Big]^2 \\
&\quad - (1 + 2 \cdot n^{-1}) \cdot \mathbb{E}\big[v_N(X_1, X_2) \cdot v_N(X_1, X_3)\big]\Big] \\
= {}& m^{-1}(m-1)^{-1}\Bigg[\Bigg[a_N^{-1}\int_{-1}^{1} K^2(v)\,dv - 2\int_0^1 vK^2(v)\,dv\Bigg] \\
&\quad + (2n + m - 1)n^{-1}\Bigg[1 - 4 \cdot a_N \int_0^1 vK(v)\,dv + 4 \cdot a_N^2 \Bigg[\int_0^1 vK(v)\,dv\Bigg]^2\Bigg] \\
&\quad - (1 + 2 \cdot n^{-1})\Bigg[1 + 2 \cdot a_N \int_{-1}^{0} \Bigg[\int_{-1}^{x} K(v)\,dv\Bigg]^2 dx - 4 \cdot a_N \int_0^1 vK(v)\,dv\Bigg]\Bigg].
\end{aligned}
$$

□

LEMMA 5.36. *Let F and G be continuous distribution functions and H_N be defined as in (2.4). Further, let K be a kernel on $(-1,\,1)$ satisfying (2.8) through (2.11) and let $0 < a_N < \frac{1}{2}$.*

Then for

$$
u_N(s,t) = a_N^{-1} K\big(a_N^{-1}(H_N(s) - H_N(t)\big)
$$

under $\mathsf{H}_0 \,:\, F = G$ we have

$$
\begin{aligned}
\mathsf{Covar}\Bigg[&m^{-1}\sum_{i=1}^{m}\Big[\mathsf{E}\big[u_N(X_1, X_i) \mid X_i\big] - \mathsf{E}\big[u_N(X_1, X_2)\big]\Big] \\
&- n^{-1}\sum_{k=1}^{n}\Big[\mathsf{E}\big[u_N(X_1, Y_k) \mid Y_k\big] - \mathsf{E}\big[u_N(X_1, Y_1)\big]\Big],
\end{aligned}
$$

$$m^{-1}(m-1)^{-1}\sum_{1\le i\ne j\le m}\Big[u_N(X_i,X_j)-\mathsf{E}\big[u_N(X_1,X_j)\mid X_j\big]\Big]$$
$$-\,m^{-1}n^{-1}\sum_{\substack{1\le i\le m\\1\le k\le n}}\Big[u_N(X_i,Y_k)-\mathsf{E}\big[u_N(X_1,Y_k)\mid Y_k\big]\Big]\Big]=0.$$

Proof. The individual summands are obviously centered so that the covariance becomes

$$\mathsf{E}\Big[m^{-1}\sum_{i=1}^{m}\Big[\mathsf{E}\big[u_N(X_1,X_i)\mid X_i\big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]$$
$$-\,n^{-1}\sum_{k=1}^{n}\Big[\mathsf{E}\big[u_N(X_1,Y_k)\mid Y_k\big]-\mathsf{E}\big[u_N(X_1,Y_1)\big]\Big]\Big]$$
$$\times\Big[m^{-1}(m-1)^{-1}\sum_{1\le i\ne j\le m}\Big[u_N(X_i,X_j)-\mathsf{E}\big[u_N(X_1,X_j)\mid X_j\big]\Big]$$
$$-\,m^{-1}n^{-1}\sum_{\substack{1\le i\le m\\1\le k\le n}}\Big[u_N(X_i,Y_k)-\mathsf{E}\big[u_N(X_1,Y_k)\mid Y_k\big]\Big]\Big]$$
$$=m^{-1}\cdot m^{-1}(m-1)^{-1}\sum_{i=1}^{m}\sum_{1\le j\ne l\le m}\mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,X_i)\mid X_i\big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]$$
$$\times\Big[u_N(X_j,X_l)-\mathsf{E}\big[u_N(X_1,X_l)\mid X_l\big]\Big]\Big] \tag{5.219}$$
$$-\,m^{-1}\cdot m^{-1}n^{-1}\sum_{i=1}^{m}\sum_{\substack{1\le l\le m\\1\le k\le n}}\mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,X_i)\mid X_i\big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]$$
$$\times\Big[u_N(X_l,Y_k)-\mathsf{E}\big[u_N(X_1,Y_k)\mid Y_k\big]\Big]\Big] \tag{5.220}$$
$$-\,n^{-1}\cdot m^{-1}(m-1)^{-1}\sum_{k=1}^{n}\sum_{1\le i\ne j\le m}\mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,Y_k)\mid Y_k\big]-\mathsf{E}\big[u_N(X_1,Y_1)\big]\Big]$$
$$\times\Big[u_N(X_i,X_j)-\mathsf{E}\big[u_N(X_1,X_j)\mid X_j\big]\Big]\Big] \tag{5.221}$$
$$+\,n^{-1}\cdot m^{-1}n^{-1}\sum_{k=1}^{n}\sum_{\substack{1\le i\le m\\1\le l\le n}}\mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,Y_k)\mid Y_k\big]-\mathsf{E}\big[u_N(X_1,Y_1)\big]\Big]$$
$$\times\Big[u_N(X_i,Y_l)-\mathsf{E}\big[u_N(X_1,Y_l)\mid Y_l\big]\Big]\Big]. \tag{5.222}$$

The expectation in (5.221) vanishes immediately, due to the independence of the X_i and Y_k, so that we are only concerned with the expectations in (5.219), (5.220) and (5.222).

Beginning with (5.219), we have

$$\sum_{i=1}^{m}\sum_{1\le j\ne l\le m}\mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,X_i)\mid X_i\big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]$$
$$\times\Big[u_N(X_j,X_l)-\mathsf{E}\big[u_N(X_1,X_l)\mid X_l\big]\Big]\Big]$$

$$
\begin{aligned}
&= m(m-1)(m-2)\mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]\\
&\qquad\times\Big[u_N(X_4,X_5)-\mathsf{E}\big[u_N(X_1,X_5)\mid X_5\big]\Big]\Big]\\
&\quad+m(m-1)\mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]\\
&\qquad\times\Big[u_N(X_3,X_4)-\mathsf{E}\big[u_N(X_1,X_4)\mid X_4\big]\Big]\Big]\\
&\quad+m(m-1)\mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]\\
&\qquad\times\Big[u_N(X_4,X_3)-\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]\Big]\Big]\\
&= m(m-1)(m-2)\cdot 0\\
&\quad+m(m-1)\mathsf{E}\Big[\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]\cdot u_N(X_3,X_4)\\
&\qquad-\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]\cdot\mathsf{E}\big[u_N(X_1,X_4)\mid X_4\big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\cdot u_N(X_3,X_4)\\
&\qquad+\mathsf{E}\big[u_N(X_1,X_2)\big]\cdot\mathsf{E}\big[u_N(X_1,X_4)\mid X_4\big]\Big]\\
&\quad+m(m-1)\mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]\\
&\qquad\times\mathsf{E}\Big[u_N(X_4,X_3)-\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]\;\Big|\;X_3\Big]\Big]\\
&= m(m-1)\Big[\mathsf{E}\Big[\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]\cdot u_N(X_4,X_3)\Big]-\Big[\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]^2\Big]\\
&\quad+m(m-1)\mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]\;\times\;0\Big]\\
&= m(m-1)\Big[\mathsf{E}\Big[\mathsf{E}\big[u_N(X_1,X_2)\mid X_2\big]\Big]^2-\Big[\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]^2\Big].
\end{aligned}
$$

Turning to (5.220) we find

$$
\begin{aligned}
&\sum_{i=1}^{m}\sum_{\substack{1\le l\le m\\1\le k\le n}}\mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,X_i)\mid X_i\big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]\cdot\Big[u_N(X_l,Y_k)-\mathsf{E}\big[u_N(X_1,Y_k)\mid Y_k\big]\Big]\Big]\\
&= m(m-1)n\cdot\mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]\cdot\Big[u_N(X_4,Y_1)-\mathsf{E}\big[u_N(X_1,Y_1)\mid Y_1\big]\Big]\Big]\\
&\quad+mn\cdot\mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]\cdot\Big[u_N(X_3,Y_1)-\mathsf{E}\big[u_N(X_1,Y_1)\mid Y_1\big]\Big]\Big]\\
&= m(m-1)n\cdot 0\\
&\quad+mn\cdot\mathsf{E}\Big[\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]\cdot u_N(X_3,Y_1)-\mathsf{E}\big[u_N(X_1,X_3)\mid X_3\big]\cdot\mathsf{E}\big[u_N(X_1,Y_1)\mid Y_1\big]\\
&\qquad-\mathsf{E}\big[u_N(X_1,X_2)\big]\cdot u_N(X_3,Y_1)+\mathsf{E}\big[u_N(X_1,X_2)\big]\cdot\mathsf{E}\big[u_N(X_1,Y_1)\mid Y_1\big]\Big]\\
&= mn\cdot\Big[\mathsf{E}\Big[\big[u_N(X_1,X_3)\mid X_3\big]\cdot u_N(Y_1,X_3)\Big]-\mathsf{E}\big[u_N(X_1,X_2)\big]\cdot\mathsf{E}\big[u_N(X_1,Y_1)\big]
\end{aligned}
$$

$$- \mathsf{E}\big[u_N(X_1,X_2)\big] \cdot \mathsf{E}\big[u_N(X_1,Y_1)\big] + \mathsf{E}\big[u_N(X_1,X_2)\big] \cdot \mathsf{E}\big[u_N(X_1,Y_1)\big]\Big]$$

$$= mn \cdot \Big[\mathsf{E}\Big[\big[u_N(X_1,X_3) \mid X_3\big] \cdot \mathsf{E}\big[u_N(Y_1,X_3) \mid X_3\big]\Big] - \mathsf{E}\big[u_N(X_1,X_2)\big] \cdot \mathsf{E}\big[u_N(X_1,Y_1)\big]\Big]$$

which under H_0 is equal to

$$mn \cdot \Big[\mathsf{E}\Big[\mathsf{E}\big[u_N(X_1,X_2) \mid X_2\big]\Big]^2 - \Big[\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]^2\Big].$$

Finally, for (5.222) we have

$$\sum_{k=1}^{n} \sum_{\substack{1\le i\le m\\ 1\le l\le n}} \mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,Y_k) \mid Y_k\big] - \mathsf{E}\big[u_N(X_1,Y_1)\big]\Big] \cdot \Big[u_N(X_i,Y_l) - \mathsf{E}\big[u_N(X_1,Y_l) \mid Y_l\big]\Big]\Big]$$

$$= nm(n-1) \cdot \mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,Y_2) \mid Y_2\big] - \mathsf{E}\big[u_N(X_1,Y_1)\big]\Big] \cdot \Big[u_N(X_1,Y_3) - \mathsf{E}\big[u_N(X_1,Y_3) \mid Y_3\big]\Big]\Big]$$

$$+ nm \cdot \mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,Y_2) \mid Y_2\big] - \mathsf{E}\big[u_N(X_1,Y_1)\big]\Big] \cdot \Big[u_N(X_1,Y_2) - \mathsf{E}\big[u_N(X_1,Y_2) \mid Y_2\big]\Big]\Big]$$

$$= nm(n-1) \cdot 0$$

$$+ nm \cdot \mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,Y_2) \mid Y_2\big] - \mathsf{E}\big[u_N(X_1,Y_1)\big]\Big] \cdot \mathsf{E}\Big[u_N(X_1,Y_2) - \mathsf{E}\big[u_N(X_1,Y_2) \mid Y_2\big] \,\Big|\, Y_2\Big]\Big]$$

$$= nm \cdot \mathsf{E}\Big[\Big[\mathsf{E}\big[u_N(X_1,Y_2) \mid Y_2\big] - \mathsf{E}\big[u_N(X_1,Y_1)\big]\Big] \cdot 0\Big]$$

$$= 0.$$

Combining our results for (5.219), (5.220) and (5.222) we have

$$\mathsf{E}\Big[m^{-1}\sum_{i=1}^{m}\Big[\mathsf{E}\big[u_N(X_1,X_i) \mid X_i\big] - \mathsf{E}\big[u_N(X_1,X_2)\big]\Big]$$

$$- n^{-1}\sum_{k=1}^{n}\Big[\mathsf{E}\big[u_N(X_1,Y_k) \mid Y_k\big] - \mathsf{E}\big[u_N(X_1,Y_1)\big]\Big]\Big]$$

$$\times \Big[m^{-1}(m-1)^{-1}\sum_{1\le i\ne j\le m}\Big[u_N(X_i,X_j) - \mathsf{E}\big[u_N(X_1,X_j) \mid X_j\big]\Big]$$

$$- m^{-1}n^{-1}\sum_{\substack{1\le i\le m\\ 1\le k\le n}}\Big[u_N(X_i,Y_k) - \mathsf{E}\big[u_N(X_1,Y_k) \mid Y_k\big]\Big]\Big]$$

$$= m^{-1} \cdot m^{-1}(m-1)^{-1} \cdot m(m-1)\Big[\mathsf{E}\Big[\mathsf{E}\big[u_N(X_1,X_2) \mid X_2\big]\Big]^2 - \Big[\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]^2\Big]$$

$$- m^{-1} \cdot m^{-1}n^{-1} \cdot mn\Big[\mathsf{E}\Big[\mathsf{E}\big[u_N(X_1,X_2) \mid X_2\big]\Big]^2 - \Big[\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]^2\Big]$$

$$= m^{-1} \cdot \Big[\mathsf{E}\Big[\mathsf{E}\big[u_N(X_1,X_2) \mid X_2\big]\Big]^2 - \Big[\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]^2\Big]$$

$$- m^{-1} \cdot \Big[\mathsf{E}\Big[\mathsf{E}\big[u_N(X_1,X_2) \mid X_2\big]\Big]^2 - \Big[\mathsf{E}\big[u_N(X_1,X_2)\big]\Big]^2\Big]$$

$$= 0.$$

which completes the proof. $\square$

APPENDIX A

Lemmata

LEMMA A.1. *Let $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ be independent real-valued random variables such that*

$$X_i \sim F, \quad 1 \leq i \leq m, \qquad\qquad \text{and} \qquad\qquad Y_k \sim G, \quad 1 \leq k \leq n$$

for continuous distribution functions F and G, and let

$$N = m + n$$

be the size of the pooled sample, and

$$H_N = \frac{m}{N}\, F + \frac{n}{N}\, G$$

be the pooled theoretical (not empirical!) distribution function. Let f_N be the Lebesgue-density of the random variables $H_N(X_i)$ and g_N be the Lebesgue-density of the $H_N(Y_k)$ and define

$$b_N = f_N - g_N.$$

Further, let ϕ be any integrable bounded function with $\phi = 0$ outside of the interval $(-1, 1)$.

Then the following inequalities hold for all sample sizes m and n:

$$\|f_N\| \leq 1 + \frac{n}{m}, \qquad \|g_N\| \leq 1 + \frac{m}{n}, \qquad \text{and} \qquad -\frac{N}{n} \leq b_N \leq \frac{N}{m}. \tag{A.1}$$

and

$$\begin{aligned}
&\left| \int \phi\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dx) \right| \leq 2\, \|\phi\| \cdot a_N \Big(1 + \frac{n}{m}\Big) \quad \text{and} \\
&\left| \int \phi\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, G(dx) \right| \leq 2\, \|\phi\| \cdot a_N \Big(1 + \frac{m}{n}\Big) \quad \text{for } x, y \in \mathbb{R}.
\end{aligned} \tag{A.2}$$

PROOF. We can derive the distribution functions of the $H_N(X_i)$ and $H_N(Y_k)$, since for all real y

$$\begin{aligned}
P(H_N(X_1) < y) &= 1 - P(H_N(X_1) \geq y) \\
&= 1 - P(X_1 \geq H_N^{-1}(y)) \\
&= 1 - \big[1 - F \circ H_N^{-1}(y)\big] \\
&= F \circ H_N^{-1}(y).
\end{aligned}$$

We can see immediately that $F \circ H_N^{-1}$ is left-continuous and admits limits from the right, since F is continuous and H_N^{-1} is left-continuous with limits from the right on $(0, 1)$ as the generalized inverse of the cumulative distribution function H_N.

We can also show that $F \circ H_N^{-1}$ is right-continuous as well, even when H_N^{-1} is not. To see this let $u \in (0,1)$ be a point where H_N^{-1} is not right-continuous, meaning there is a jump at u with

$$H_N^{-1}(u) < \lim_{v \downarrow u} H_N^{-1}(v),$$

where $\lim_{v\downarrow u} H_N^{-1}(v)$ is the right-hand limit of H_N^{-1} at u.

From the continuity of H_N we have

$$H_N \circ H_N^{-1}(u) = u = \lim_{v\downarrow u} v = \lim_{v\downarrow u} H_N \circ H_N^{-1}(v) = H_N(\lim_{v\downarrow u} H_N^{-1}(v)),$$

so that H_N must be constant equal to u on the interval $[H_N^{-1}(u), \lim_{v\downarrow u} H_N^{-1}(v)]$, which can only be the case if F and G are constant on this interval as well, which gives us the right-hand continuity of $F \circ H_N^{-1}$ in u, since

$$\lim_{v\downarrow u} F \circ H_N^{-1}(v) = F(\lim_{v\downarrow u} H_N^{-1}(v)) = F \circ H_N^{-1}(u).$$

Using the continuity of $F \circ H_N^{-1}$, we then have

$$P(H_N(X_1) \le y) = \lim_{x\downarrow y} P(H_N(X_1) < x) = \lim_{x\downarrow y} F \circ H_N^{-1}(x) = F \circ H_N^{-1}(y).$$

Analogously

$$P(H_N(Y_1) \le y) = G \circ H_N^{-1}(y).$$

Thus for the Lebesgue-densities f_N and g_N we may write,

$$f_N = \frac{d(F \circ H_N^{-1})}{d\mu} \qquad \text{and} \qquad g_N = \frac{d(G \circ H_N^{-1})}{d\mu}.$$

From the definition of H_N we get

$$Id = H_N \circ H_N^{-1} = \frac{m}{N}\, F \circ H_N^{-1} + \frac{n}{N}\, G \circ H_N^{-1},$$

which implies

$$1 = \frac{m}{N}\, f_N + \frac{n}{N}\, g_N.$$

This can be used to bound b_N:

$$\frac{m}{N} \cdot b_N = \frac{m}{N} \cdot \big[f_N - g_N\big] = 1 - \frac{n}{N} \cdot g_N - \frac{m}{N} \cdot g_N = 1 - g_N \le 1,$$

and

$$-\frac{n}{N} \cdot b_N = \frac{n}{N} \cdot \big[g_N - f_N\big] = 1 - \frac{m}{N} \cdot f_N - \frac{n}{N} \cdot f_N = 1 - f_N \le 1,$$

so that

$$-\frac{N}{n} \le b_N \le \frac{N}{m}.$$

Further,

$$0 \le f_N = \frac{m}{N} \cdot f_N + \frac{n}{N} \cdot f_N = \big[1 - \frac{n}{N} \cdot g_N\big] + \frac{n}{N} \cdot f_N = 1 + \frac{n}{N} \cdot b_N \le 1 + \frac{n}{m},$$

and

$$0 \le g_N = \frac{m}{N} \cdot g_N + \frac{n}{N} \cdot g_N = \frac{m}{N} \cdot g_N + \big[1 - \frac{m}{N} \cdot f_N\big] = 1 - \frac{m}{N} \cdot b_N \le 1 + \frac{m}{n},$$

so that

$$\|f_N\| \leq 1 + \frac{n}{m} \qquad \text{and} \qquad \|g_N\| \leq 1 + \frac{m}{n}$$

as claimed.

Regarding (A.2) it is easy to see that

$$\begin{aligned}\int \phi\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dx) &= \int_0^1 \phi\big(a_N^{-1}(w - H_N(y))\big) \cdot f_N(w)\, dw \\ &= a_N \int_{-a_N^{-1} H_N(y)}^{a_N^{-1}(1-H_N(y))} \phi(u) f_N(H_N(y) + a_N \cdot u)\, du,\end{aligned}$$

so that

$$\begin{aligned}\left| \int \phi\big(a_N^{-1}(H_N(x) - H_N(y))\big)\, F(dx)\right| &\leq a_N \cdot \|f_N\| \int_{-a_N^{-1} H_N(y)}^{a_N^{-1}(1-H_N(y))} \big|\phi(u)\big|\, du \\ &\leq 2\,\|\phi\| \cdot a_N(1 + \frac{n}{m}).\end{aligned}$$

The proof for the second inequality in (A.2) with $G(dx)$ in place of $F(dx)$ is completely analogous using the bound $\|g_N\| \leq 1 + \frac{m}{n}$ in the final inequality. □

Lemma A.2. *Let $X_1,\ X_2,\ \ldots,\ X_n \sim F$ be an i.i.d. sample and let*

$$U_n = n^{-1}(n-1)^{-1} \cdot \sum_{1 \leq i \neq j \leq n} u(X_i, X_j)$$

be a U-statistic of degree 2 with kernel u such that

$$\mathbb{E}\big[u(X_1, X_2)\big]^2 < \infty.$$

Further, define

$$\begin{aligned}\hat{U}_n &= n^{-1} \cdot \sum_{i=1}^{n} \left[\int u(X_i, y)\, F(dy) + \int u(x, X_i)\, F(dx) - \iint u(x, y)\, F(dx) F(dy)\right] \\ &= n^{-1}(n-1)^{-1} \cdot \sum_{1 \leq i \neq j \leq n} \left[\int u(X_i, y)\, F(dy) + \int u(x, X_j)\, F(dx) - \iint u(x, y)\, F(dx) F(dy)\right]\end{aligned}$$

to be the Hájek projection of U_n and u^* as

$$u^*(r, s) = u(r, s) - \int u(r, y)\, F(dy) - \int u(x, s)\, F(dx) + \iint u(x, y)\, F(dx) F(dy).$$

Then

$$\mathbb{E}\Big[U_n - \hat{U}_n\Big]^2 \leq 2n^{-1}(n-1)^{-1} \cdot \mathbb{E}\big[u^*(X_1, X_2)\big]^2.$$

Proof.

$$\begin{aligned}\mathbb{E}\Big[U_n - \hat{U}_n\Big]^2 &= \mathbb{E}\Big[n^{-1}(n-1)^{-1} \cdot \sum_{1 \leq i \neq j \leq n} u^*(X_i, X_j)\Big]^2 \\ &= n^{-2}(n-1)^{-2} \cdot \mathbb{E}\Big[\sum_{\substack{1 \leq i,\, j,\, k,\, l \leq n \\ i \neq j,\, k \neq l}} u^*(X_i, X_j) \cdot u^*(X_k, X_l)\Big].\end{aligned}$$

Expanding the expectation this is equal to

$$n^{-2}(n-1)^{-2}\Big[n(n-1)(n-2)(n-3)\cdot\mathbb{E}\big[u^*(X_1,X_2)\cdot u^*(X_3,X_4)\big] \tag{A.3}$$

$$+\,n(n-1)(n-2)\cdot\Big[\mathbb{E}\big[u^*(X_1,X_2)\cdot u^*(X_1,X_3)\big] \tag{A.4}$$

$$+\,2\cdot\mathbb{E}\big[u^*(X_1,X_2)\cdot u^*(X_2,X_3)\big] \tag{A.5}$$

$$+\,\mathbb{E}\big[u^*(X_1,X_2)\cdot u^*(X_3,X_2)\big]\Big] \tag{A.6}$$

$$+\,n(n-1)\cdot\Big[\mathbb{E}\big[u^*(X_1,X_2)\cdot u^*(X_2,X_1)\big] \tag{A.7}$$

$$+\,\mathbb{E}\big[u^*(X_1,X_2)\big]^2\Big]\Big]. \tag{A.8}$$

Clearly, $\mathbb{E}\big[u^*(X_1,X_2)\big]=0$, so (A.3) vanishes immediately. Additionally, the expectation in (A.4) vanishes due to

$$\mathbb{E}\big[u^*(X_1,X_2)\cdot u^*(X_1,X_3)\big]=\mathbb{E}\big[\mathbb{E}\big[u^*(X_1,X_2)\mid X_1\big]\cdot\mathbb{E}\big[u^*(X_1,X_3)\mid X_1\big]\big],$$

since for the inner expectation

$$\begin{aligned}&\mathbb{E}\big[u^*(X_1,X_2)\mid X_1\big]\\ &\quad=\int v(X_1,y)F(dy)-\int v(X_1,y)F(dy)-\iint v(y,z)F(dy)F(dz)+\iint v(y,z)F(dy)F(dz)\\ &\quad=0.\end{aligned} \tag{A.9}$$

(A.9) implies directly that (A.5) vanishes as well, since

$$\mathbb{E}\big[u^*(X_1,X_2)\cdot u^*(X_2,X_3)\big]=\mathbb{E}\big[\mathbb{E}\big[u^*(X_1,X_2)\mid X_2\big]\cdot\mathbb{E}\big[u^*(X_2,X_3)\mid X_2\big]\big].$$

Analogously, (A.6) vanishes as well, since

$$\mathbb{E}\big[u^*(X_1,X_2)\cdot u^*(X_3,X_2)\big]=\mathbb{E}\big[\mathbb{E}\big[u^*(X_1,X_2)\mid X_2\big]\cdot\mathbb{E}\big[u^*(X_3,X_2)\mid X_2\big]\big]$$

and the inner expectation is

$$\begin{aligned}&\mathbb{E}\big[u^*(X_1,X_2)\mid X_2\big]\\ &\quad=\int v(y,X_2)F(dy)-\iint v(z,y)F(dy)F(dz)-\int v(y,X_2)F(dy)+\iint v(y,z)F(dy)F(dz)\\ &\quad=0.\end{aligned} \tag{A.10}$$

The expectation in (A.7) is bounded by the expectation in (A.8):

$$\begin{aligned}\big|\mathbb{E}\big[u^*(X_1,X_2)\cdot u^*(X_2,X_1)\big]\big|&\le\Big[\mathbb{E}\big[u^*(X_1,X_2)\big]^2\Big]^{\frac{1}{2}}\cdot\Big[\mathbb{E}\big[u^*(X_2,X_1)\big]^2\Big]^{\frac{1}{2}}\\ &=\mathbb{E}\big[u^*(X_1,X_2)\big]^2,\end{aligned}$$

so that we have

$$n^{-2}(n-1)^{-2}\cdot\mathbb{E}\Big[\sum_{\substack{1\le i,\,j,\,k,\,l\le n\\ i\ne j,\,k\ne l}}u^*(X_i,X_j)\cdot u^*(X_k,X_l)\Big]$$

$$\leq 2n^{-1}(n-1)^{-1} \cdot \mathbb{E}\big[u^*(X_1, X_2)\big]^2.$$

□

LEMMA A.3. *Let $X_1,\ X_2,\ \ldots,\ X_m \sim F$ and $Y_1,\ Y_2,\ \ldots,\ Y_n \sim G$ be two independent i.i.d. samples and let*

$$U_{m,n} = m^{-1}n^{-1} \cdot \sum_{\substack{1\leq i\leq m \\ 1\leq k\leq n}} u(X_i, Y_k)$$

be a generalized U-statistic of degree 2 with kernel u such that

$$\mathbb{E}\big[u(X_1, Y_1)\big]^2 < \infty.$$

Further, define

$$\begin{aligned}\hat{U}_{m,n} &= m^{-1} \cdot \sum_{i=1}^{m} \int u(X_i, y)\, G(dy) + n^{-1} \cdot \sum_{k=1}^{n} \int u(x, Y_k)\, F(dx) - \iint u(x, y)\, F(dx)G(dy) \\ &= m^{-1}n^{-1} \cdot \sum_{\substack{1\leq i\leq m \\ 1\leq k\leq n}} \Big[\int u(X_i, y)\, G(dy) + \int u(x, Y_k)\, F(dx) - \iint u(x, y)\, F(dx)G(dy) \Big]\end{aligned}$$

to be the Hájek projection of $U_{m,n}$ and u^* as

$$u^*(r, s) = u(r, s) - \int u(r, y)\, G(dy) - \int u(x, s)\, F(dx) + \iint u(x, y)\, F(dx)\, G(dy).$$

Then

$$\mathbb{E}\Big[U_{m,n} - \hat{U}_{m,n}\Big]^2 = m^{-1}n^{-1} \cdot \mathbb{E}\big[u^*(X_1, Y_1)\big]^2.$$

PROOF.

$$\begin{aligned}\mathbb{E}\Big[U_{m,n} - \hat{U}_{m,n}\Big]^2 &= \mathbb{E}\Big[m^{-1}n^{-1} \cdot \sum_{\substack{1\leq i\leq m \\ 1\leq k\leq n}} u^*(X_i, Y_k)\Big]^2 \\ &= m^{-2}n^{-2} \cdot \mathbb{E}\Big[\sum_{\substack{1\leq i\leq m \\ 1\leq k\leq n}} \sum_{\substack{1\leq j\leq m \\ 1\leq l\leq n}} u^*(X_i, Y_k) \cdot u^*(X_j, Y_l)\Big].\end{aligned}$$

Expanding the expectation this is equal to

$$m^{-2}n^{-2}\Big[m(m-1)n(n-1) \cdot \mathbb{E}\big[u^*(X_1, Y_1) \cdot u^*(X_2, Y_2)\big] \tag{A.11}$$

$$+\, mn(n-1) \cdot \mathbb{E}\big[u^*(X_1, Y_1) \cdot u^*(X_1, Y_2)\big] \tag{A.12}$$

$$+\, m(m-1)n \cdot \mathbb{E}\big[u^*(X_1, Y_1) \cdot u^*(X_2, Y_1)\big] \tag{A.13}$$

$$+\, mn \cdot \mathbb{E}\big[u^*(X_1, Y_1)\big]^2\Big]. \tag{A.14}$$

Clearly, $\mathbb{E}\left[u^*(X_1, Y_1)\right] = 0$, so the expectation in (A.11) vanishes immediately due to the independence of $u^*(X_1, Y_1)$ and $u^*(X_2, Y_2)$. Furthermore, the expectation in (A.12) vanishes due to

$$\mathbb{E}\big[u^*(X_1, Y_1) \cdot u^*(X_1, Y_2)\big] = \mathbb{E}\Big[\mathbb{E}\big[u^*(X_1, Y_1) \mid X_1\big] \cdot \mathbb{E}\big[u^*(X_1, Y_2) \mid X_1\big]\Big],$$

since for the inner expectation

$$\begin{aligned}
&\mathbb{E}\big[u^*(X_1, Y_1) \mid X_1\big] \\
&\quad = \int v(X_1, y) G(dy) - \int v(X_1, y) G(dy) - \iint v(x, y) F(dx) G(dy) + \iint v(x, y) F(dx) G(dy) \\
&\quad = 0.
\end{aligned}$$

Analogously, (A.13) vanishes as well, since

$$\mathbb{E}\big[u^*(X_1, Y_1) \cdot u^*(X_2, Y_1)\big] = \mathbb{E}\Big[\mathbb{E}\big[u^*(X_1, Y_1) \mid Y_1\big] \cdot \mathbb{E}\big[u^*(X_2, Y_1) \mid Y_1\big]\Big],$$

and the inner expectation is

$$\begin{aligned}
&\mathbb{E}\big[u^*(X_1, Y_1) \mid Y_1\big] \\
&\quad = \int v(x, Y_1)\, F(dx) - \iint v(x, y)\, F(dx) G(dy) - \int v(x, Y_1)\, F(dx) + \iint v(x, y)\, F(dx) G(dy) \\
&\quad = 0.
\end{aligned}$$

This leaves us with

$$m^{-2} n^{-2} \cdot \mathbb{E}\Bigg[\sum_{\substack{1 \le i \le m \\ 1 \le k \le n}} \sum_{\substack{1 \le j \le m \\ 1 \le l \le n}} u^*(X_i, Y_k) \cdot u^*(X_j, Y_l)\Bigg] = m^{-1} n^{-1} \cdot \mathbb{E}\big[u^*(X_1, Y_1)\big]^2.$$

□

LEMMA A.4. *Let $X_1, X_2, \ldots, X_n \sim F$ be an i.i.d. sample and let*

$$U_n = n^{-1}(n-1)^{-1}(n-2)^{-1} \cdot \sum_{\substack{1 \le i,\, j,\, k \le n \\ i \ne j,\, j \ne k,\, i \ne k}} u(X_i, X_j, X_k)$$

be a U-statistic of degree 3 with kernel u such that $\mathbb{E}\big[u(X_1, X_2, X_3)\big]^2 < \infty$. Further, define

$$\begin{aligned}
\hat{U}_n &= n^{-1} \cdot \sum_{i=1}^{n} \Bigg[\iint u(X_i, y, z)\, F(dy) F(dz) + \iint u(x, X_i, z)\, F(dx) F(dz) \\
&\qquad + \iint u(x, y, X_i)\, F(dx) F(dy) - 2 \cdot \iiint u(x, y, z)\, F(dx) F(dy) F(dz)\Bigg] \\
&= n^{-1}(n-1)^{-1}(n-2)^{-1} \cdot \sum_{\substack{1 \le i,\, j,\, k \le n \\ i \ne j,\, j \ne k,\, i \ne k}} \Bigg[\iint u(X_i, y, z)\, F(dy) F(dz) + \iint u(x, X_j, z)\, F(dx) F(dz) \\
&\qquad + \iint u(x, y, X_k)\, F(dx) F(dy) - 2 \cdot \iiint u(x, y, z)\, F(dx) F(dy) F(dz)\Bigg]
\end{aligned}$$

to be the Hájek projection of U_n and u^* as

$$\begin{aligned}
u^*(r, s, t) = u(r, s, t) &- \iint u(r, y, z)\, F(dy) F(dz) - \iint u(x, s, z)\, F(dx) F(dz) \\
&- \iint u(x, y, t)\, F(dx) F(dy) + 2 \cdot \iiint u(x, y, z)\, F(dx) F(dy) F(dz).
\end{aligned}$$

Then

$$\mathbb{E}\Big[U_n - \hat{U}_n\Big]^2 \le \left[\frac{18(n-3)+6}{n(n-1)(n-2)}\right] \cdot \mathbb{E}\big[u^*(X_1, X_2, X_3)\big]^2$$
$$= O(n^{-2}) \cdot \mathbb{E}\big[u^*(X_1, X_2, X_3)\big]^2.$$

Proof.

$$\mathbb{E}\Big[U_n - \hat{U}_n\Big]^2 \tag{A.15}$$
$$= \quad \mathbb{E}\Bigg[n^{-1}(n-1)^{-1}(n-2)^{-1} \cdot \sum_{\substack{1\le i,\, j,\, k\le n\\ i\ne j,\, j\ne k,\, i\ne k}} u^*(X_i, X_j, X_k)\Bigg]^2$$
$$= \mathbb{E}\Bigg[n^{-2}(n-1)^{-2}(n-2)^{-2} \cdot \sum_{\substack{1\le i_1,\, i_2,\, i_3\le n\\ i_1\ne i_2,\, i_2\ne i_3,\, i_1\ne i_3}} \quad \sum_{\substack{1\le i_4,\, i_5,\, i_6\le n\\ i_4\ne i_5,\, i_5\ne i_6,\, i_4\ne i_6}} u^*(X_{i_1}, X_{i_2}, X_{i_3}) \cdot u^*(X_{i_4}, X_{i_5}, X_{i_6})\Bigg]$$
$$= n^{-2}(n-1)^{-2}(n-2)^{-2} \cdot \Bigg[\sum_{\substack{1\le i_1,\, i_2,\, i_3,\, i_4,\, i_5,\, i_6\le n\\ i_1\ne i_2,\, i_2\ne i_3,\, i_1\ne i_3 \text{ and } i_4\ne i_5,\, i_5\ne i_6,\, i_4\ne i_6\\ |\{i_1,i_2,i_3\}\cap\{i_4,i_5,i_6\}|=0}} \mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, X_{i_3}) \cdot u^*(X_{i_4}, X_{i_5}, X_{i_6})\big] \tag{A.16}$$
$$+ \sum_{\substack{1\le i_1,\, i_2,\, i_3,\, i_4,\, i_5,\, i_6\le n\\ i_1\ne i_2,\, i_2\ne i_3,\, i_1\ne i_3 \text{ and } i_4\ne i_5,\, i_5\ne i_6,\, i_4\ne i_6\\ |\{i_1,i_2,i_3\}\cap\{i_4,i_5,i_6\}|=1}} \mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, X_{i_3}) \cdot u^*(X_{i_4}, X_{i_5}, X_{i_6})\big] \tag{A.17}$$
$$+ \sum_{\substack{1\le i_1,\, i_2,\, i_3,\, i_4,\, i_5,\, i_6\le n\\ i_1\ne i_2,\, i_2\ne i_3,\, i_1\ne i_3 \text{ and } i_4\ne i_5,\, i_5\ne i_6,\, i_4\ne i_6\\ |\{i_1,i_2,i_3\}\cap\{i_4,i_5,i_6\}|=2}} \mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, X_{i_3}) \cdot u^*(X_{i_4}, X_{i_5}, X_{i_6})\big] \tag{A.18}$$
$$+ \sum_{\substack{1\le i_1,\, i_2,\, i_3,\, i_4,\, i_5,\, i_6\le n\\ i_1\ne i_2,\, i_2\ne i_3,\, i_1\ne i_3 \text{ and } i_4\ne i_5,\, i_5\ne i_6,\, i_4\ne i_6\\ |\{i_1,i_2,i_3\}\cap\{i_4,i_5,i_6\}|=3}} \mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, X_{i_3}) \cdot u^*(X_{i_4}, X_{i_5}, X_{i_6})\big]\Bigg]. \tag{A.19}$$

(A.16) is made up of $n(n-1)(n-2)(n-3)(n-4)(n-5)$ summands which are all equal to zero, due to the independence of the $X_1, \dots, X_n$ and since

$$\begin{aligned}
&\mathbb{E}\big[u^*(X_1, X_2, X_3)\big]\\
&\quad = \mathbb{E}\Big[u(X_1, X_2, X_3) - \iint u(X_1, y, z)\, F(dy)F(dz) - \iint u(x, X_2, z)\, F(dx)F(dz)\\
&\qquad\quad - \iint u(x, y, X_3)\, F(dx)F(dy) + 2 \cdot \iiint u(x, y, z)\, F(dx)F(dy)F(dz)\Big]\\
&\quad = \iiint u(x, y, z)\, F(dx)F(dy)F(dz) - \iiint u(x, y, z)\, F(dy)F(dz)F(dx)\\
&\qquad\quad - \iiint u(x, y, z)\, F(dx)F(dz)F(dy) - \iiint u(x, y, z)\, F(dx)F(dy)F(dz)\\
&\qquad\quad + 2 \cdot \iiint u(x, y, z)\, F(dx)F(dy)F(dz)\\
&\quad = 0.
\end{aligned}$$

Each of the expectations in (A.17) has the form

$$\mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, X_{i_3}) \cdot u^*(X_{i_4}, X_{i_5}, X_{i_6})\big]$$

such that exactly one of the $\{X_{i_1}, X_{i_2}, X_{i_3}\}$ is equal to exactly one of the $\{X_{i_4}, X_{i_5}, X_{i_6}\}$. In the case that $X_{i_1} \in \{X_{i_4}, X_{i_5}, X_{i_6}\}$ we can write

$$\mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, X_{i_3}) \cdot u^*(X_{i_4}, X_{i_5}, X_{i_6})\big] = \mathbb{E}\big[\mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, X_{i_3}) \mid X_{i_1}\big] \cdot u^*(X_{i_4}, X_{i_5}, X_{i_6})\big] = 0,$$

since for the inner expectation

$$\begin{aligned}
&\mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, X_{i_3}) \mid X_{i_1}\big] \\
&= \mathbb{E}\Big[u(X_{i_1}, X_{i_2}, X_{i_3}) - \iint u(X_{i_1}, y, z)\, F(dy)F(dz) - \iint u(x, X_{i_2}, z)\, F(dx)F(dz) \\
&\qquad - \iint u(x, y, X_{i_3})\, F(dx)F(dy) + 2 \cdot \iiint u(x, y, z)\, F(dx)F(dy)F(dz) \mid X_{i_1}\Big] \\
&= \iint u(X_{i_1}, y, z)\, F(dy)F(dz) - \iint u(X_{i_1}, y, z)\, F(dy)F(dz) \\
&\qquad - \iiint u(x, y, z)\, F(dx)F(dz)F(dy) - \iiint u(x, y, z)\, F(dx)F(dy)F(dz) \\
&\qquad + 2 \cdot \iiint u(x, y, z)\, F(dx)F(dy)F(dz) \\
&= 0.
\end{aligned}$$

In the case that $X_{i_2} \in \{X_{i_4}, X_{i_5}, X_{i_6}\}$ or $X_{i_3} \in \{X_{i_4}, X_{i_5}, X_{i_6}\}$ completely analogous arguments show that the expectations are equal to zero as well, so that we have shown that the sums (A.16) and (A.17) both vanish completely.

There are $\frac{18 \cdot n!}{(n-4)!}$ summands in (A.18) and $\frac{6n!}{(n-3)!}$ summands in (A.19) all of which are bounded by $\mathbb{E}\big[u^*(X_1, X_2, X_3)\big]^2$ due to the Cauchy-Schwarz inequality, so that altogether

$$\begin{aligned}
\mathbb{E}\Big[U_n - \hat{U}_n\Big]^2 &\le n^{-2}(n-1)^{-2}(n-2)^{-2} \cdot \left[\frac{18 \cdot n!}{(n-4)!} + \frac{6n!}{(n-3)!}\right] \cdot \mathbb{E}\big[u^*(X_1, X_2, X_3)\big]^2 \\
&= \left[\frac{18n(n-1)(n-2)(n-3) + 6 \cdot n(n-1)(n-2)}{n^2(n-1)^2(n-2)^2}\right] \cdot \mathbb{E}\big[u^*(X_1, X_2, X_3)\big]^2 \\
&= \left[\frac{18(n-3) + 6}{n(n-1)(n-2)}\right] \cdot \mathbb{E}\big[u^*(X_1, X_2, X_3)\big]^2 \\
&= O(n^{-2}) \cdot \mathbb{E}\big[u^*(X_1, X_2, X_3)\big]^2.
\end{aligned}$$

□

LEMMA A.5. *Let $X_1, X_2, \ldots, X_m \sim F$ and $Y_1, Y_2, \ldots, Y_n \sim G$ be two independent i.i.d. samples and let*

$$U_{m,n} = m^{-1}(m-1)^{-1}n^{-1} \cdot \sum_{\substack{1 \le i \ne j \le m \\ 1 \le k \le n}} u(X_i, X_j, Y_k)$$

be a generalized U-statistic of degree 3 with kernel u such that $\mathbb{E}\big[u(X_1,X_2,Y_1)\big]^2<\infty$. *Further, define*

$$\hat{U}_{m,n}=m^{-1}\cdot\sum_{i=1}^{m}\bigg[\iint u(X_i,y,z)\,F(dy)G(dz)+\iint u(x,X_i,z)\,F(dx)G(dz)\bigg]$$
$$+\,n^{-1}\sum_{k=1}^{n}\iint u(x,y,Y_k)\,F(dx)F(dy)-2\cdot\iiint u(x,y,z)\,F(dx)F(dy)G(dz)$$
$$=m^{-1}(m-1)^{-1}n^{-1}\cdot\sum_{\substack{1\le i\neq j\le m\\1\le k\le n}}\bigg[\iint u(X_i,y,z)\,F(dy)G(dz)+\iint u(x,X_j,z)\,F(dx)G(dz)$$
$$+\iint u(x,y,Y_k)\,F(dx)F(dy)-2\cdot\iiint u(x,y,z)\,F(dx)F(dy)G(dz)\bigg]$$

to be the Hájek projection of $U_{m,n}$ *and* u^* *as*

$$u^*(r,s,t)=u(r,s,t)-\iint u(r,y,z)\,F(dy)G(dz)-\iint u(x,s,z)\,F(dx)G(dz)$$
$$-\iint u(x,y,t)\,F(dx)F(dy)+2\cdot\iiint u(x,y,z)\,F(dx)F(dy)G(dz).$$

Then

$$\mathbb{E}\Big[U_{m,n}-\hat{U}_{m,n}\Big]^2\le\bigg[\frac{4(m-2)+2n}{m^{-1}(m-1)^{-1}n^{-1}}\bigg]\cdot\mathbb{E}\big[u^*(X_1,X_2,Y_1)\big]^2$$
$$=\Big[O(m^{-1}n^{-1})+O(m^{-2})\Big]\cdot\mathbb{E}\big[u^*(X_1,X_2,Y_1)\big]^2.$$

PROOF.

$$\mathbb{E}\Big[U_{m,n}-\hat{U}_{m,n}\Big]^2$$
$$=\mathbb{E}\bigg[m^{-1}(m-1)^{-1}n^{-1}\cdot\sum_{\substack{1\le i\neq j\le m\\1\le k\le n}}u^*(X_i,X_j,Y_k)\bigg]^2$$
$$=\mathbb{E}\bigg[m^{-2}(m-1)^{-2}n^{-2}\cdot\sum_{\substack{1\le i_1\neq i_2\le m\\1\le i_3\le n}}\ \sum_{\substack{1\le i_4\neq i_5\le m\\1\le i_6\le n}}u^*(X_{i_1},X_{i_2},Y_{i_3})\cdot u^*(X_{i_4},X_{i_5},Y_{i_6})\bigg]$$
$$=m^{-2}(m-1)^{-2}n^{-2}\cdot\bigg[\sum_{\substack{1\le i_1,\,i_2,\,i_4,\,i_5\le m\text{ and }1\le i_3,\,i_6\le n\\ i_1\neq i_2,\,i_4\neq i_5\text{ and }i_3=i_6\\ |\{i_1,i_2\}\cap\{i_4,i_5\}|=0}}\mathbb{E}\big[u^*(X_{i_1},X_{i_2},Y_{i_3})\cdot u^*(X_{i_4},X_{i_5},Y_{i_6})\big]\tag{A.20}$$
$$+\sum_{\substack{1\le i_1,\,i_2,\,i_4,\,i_5\le m\text{ and }1\le i_3,\,i_6\le n\\ i_1\neq i_2,\,i_4\neq i_5\text{ and }i_3=i_6\\ |\{i_1,i_2\}\cap\{i_4,i_5\}|=1}}\mathbb{E}\big[u^*(X_{i_1},X_{i_2},Y_{i_3})\cdot u^*(X_{i_4},X_{i_5},Y_{i_6})\big]\tag{A.21}$$
$$+\sum_{\substack{1\le i_1,\,i_2,\,i_4,\,i_5\le m\text{ and }1\le i_3,\,i_6\le n\\ i_1\neq i_2,\,i_4\neq i_5\text{ and }i_3=i_6\\ |\{i_1,i_2\}\cap\{i_4,i_5\}|=2}}\mathbb{E}\big[u^*(X_{i_1},X_{i_2},Y_{i_3})\cdot u^*(X_{i_4},X_{i_5},Y_{i_6})\big]\tag{A.22}$$
$$+\sum_{\substack{1\le i_1,\,i_2,\,i_4,\,i_5\le m\text{ and }1\le i_3,\,i_6\le n\\ i_1\neq i_2,\,i_4\neq i_5\text{ and }i_3\neq i_6\\ |\{i_1,i_2\}\cap\{i_4,i_5\}|=0}}\mathbb{E}\big[u^*(X_{i_1},X_{i_2},Y_{i_3})\cdot u^*(X_{i_4},X_{i_5},Y_{i_6})\big]\tag{A.23}$$

$$+ \sum_{\substack{1\le i_1,\, i_2,\, i_4,\, i_5\le m \text{ and } 1\le i_3,\, i_6\le n \\ i_1\ne i_2,\, i_4\ne i_5 \text{ and } i_3\ne i_6 \\ |\{i_1,i_2\}\cap\{i_4,i_5\}|=1}} \mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, Y_{i_3}) \cdot u^*(X_{i_4}, X_{i_5}, Y_{i_6})\big] \tag{A.24}$$

$$+ \sum_{\substack{1\le i_1,\, i_2,\, i_4,\, i_5\le m \text{ and } 1\le i_3,\, i_6\le n \\ i_1\ne i_2,\, i_4\ne i_5 \text{ and } i_3\ne i_6 \\ |\{i_1,i_2\}\cap\{i_4,i_5\}|=2}} \mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, Y_{i_3}) \cdot u^*(X_{i_4}, X_{i_5}, Y_{i_6})\big]\Bigg]. \tag{A.25}$$

(A.23) is made up of $m(m-1)(m-2)(m-3)n(n-1)$ summands which are all equal to zero, due to the independence of the $X_1, \dots, X_m$ and $Y_1, \dots, Y_n$ and since

$$\begin{aligned}
&\mathbb{E}\big[u^*(X_1, X_2, Y_3)\big] \\
&\quad = \mathbb{E}\Big[u(X_1, X_2, Y_3) - \iint u(X_1, y, z)\, F(dy)G(dz) - \iint u(x, X_2, z)\, F(dx)G(dz) \\
&\qquad\quad - \iint u(x, y, Y_3)\, F(dx)F(dy) + 2 \cdot \iiint u(x, y, z)\, F(dx)F(dy)G(dz)\Big] \\
&\quad = \iiint u(x, y, z)\, F(dx)F(dy)G(dz) - \iiint u(x, y, z)\, F(dy)G(dz)F(dx) \\
&\qquad - \iiint u(x, y, z)\, F(dx)G(dz)F(dy) - \iiint u(x, y, z)\, F(dx)F(dy)G(dz) \\
&\qquad + 2 \cdot \iiint u(x, y, z)\, F(dx)F(dy)G(dz) \\
&\quad = 0.
\end{aligned}$$

Each of the expectations in (A.20) has the form

$$\mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, Y_{i_3}) \cdot u^*(X_{i_4}, X_{i_5}, Y_{i_6})\big]$$

such that the $\{X_{i_1}, X_{i_2}, X_{i_4}, X_{i_5}\}$ are all unique, and $Y_{i_3} = Y_{i_6}$. In this case we may write

$$\mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, Y_{i_3}) \cdot u^*(X_{i_4}, X_{i_5}, Y_{i_6})\big] = \mathbb{E}\big[\mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, Y_{i_3}) \mid Y_{i_3}\big] \cdot u^*(X_{i_4}, X_{i_5}, Y_{i_6})\big] = 0,$$

since for the inner expectation

$$\begin{aligned}
&\mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, Y_{i_3}) \mid Y_{i_3}\big] \\
&\quad = \mathbb{E}\Big[u(X_{i_1}, X_{i_2}, Y_{i_3}) - \iint u(X_{i_1}, y, z)\, F(dy)G(dz) - \iint u(x, X_{i_2}, z)\, F(dx)G(dz) \\
&\qquad\quad - \iint u(x, y, Y_{i_3})\, F(dx)F(dy) + 2 \cdot \iiint u(x, y, z)\, F(dx)F(dy)G(dz) \mid Y_{i_3}\Big] \\
&\quad = \iint u(x, y, Y_{i_3})\, F(dx)F(dy) - \iiint u(x, y, z)\, F(dy)G(dz)F(dx) \\
&\qquad - \iiint u(x, y, z)\, F(dx)G(dz)F(dy) - \iint u(x, y, Y_{i_3})\, F(dx)F(dy) \\
&\qquad + 2 \cdot \iiint u(x, y, z)\, F(dx)F(dy)G(dz) \\
&\quad = 0.
\end{aligned}$$

Further, each of the expectations in (A.24) has the form

$$\mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, Y_{i_3}) \cdot u^*(X_{i_4}, X_{i_5}, Y_{i_6})\big]$$

such that exactly one of the $\{X_{i_1}, X_{i_2}\}$ is equal to exactly one of the $\{X_{i_4}, X_{i_5}\}$, and $Y_{i_3} \neq Y_{i_6}$. In the case that $X_{i_1} \in \{X_{i_4}, X_{i_5}\}$ the expectation vanishes since

$$\mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, Y_{i_3}) \cdot u^*(X_{i_4}, X_{i_5}, Y_{i_6})\big] = \mathbb{E}\big[\mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, Y_{i_3}) \mid X_{i_1}\big] \cdot u^*(X_{i_4}, X_{i_5}, Y_{i_6})\big],$$

and for the inner expectation

$$\begin{aligned}
&\mathbb{E}\big[u^*(X_{i_1}, X_{i_2}, Y_{i_3}) \mid X_{i_1}\big] \\
&\quad = \mathbb{E}\Big[u(X_{i_1}, X_{i_2}, Y_{i_3}) - \iint u(X_{i_1}, y, z)\, F(dy)G(dz) - \iint u(x, X_{i_2}, z)\, F(dx)G(dz) \\
&\qquad\quad - \iint u(x, y, Y_{i_3})\, F(dx)F(dy) + 2 \cdot \iiint u(x, y, z)\, F(dx)F(dy)G(dz) \mid X_{i_1}\Big] \\
&\quad = \iint u(X_{i_1}, y, z)\, F(dy)G(dz) - \iint u(X_{i_1}, y, z)\, F(dy)G(dz) \\
&\qquad - \iiint u(x, y, z)\, F(dx)G(dz)F(dy) - \iiint u(x, y, z)\, F(dx)F(dy)G(dz) \\
&\qquad + 2 \cdot \iiint u(x, y, z)\, F(dx)F(dy)G(dz) \\
&\quad = 0.
\end{aligned}$$

In the case that $X_{i_2} \in \{X_{i_4}, X_{i_5}\}$ completely analogous arguments show that the expectations are equal to zero as well, so that we have shown that the sums (A.20), (A.23) and (A.24) all vanish completely.

There are $\frac{4 \cdot m! \cdot n}{(m-3)!}$ summands in (A.21) and $\frac{2m! \cdot n}{(m-2)!}$ summands in (A.22) and $\frac{2 \cdot m! \cdot n!}{(m-2)!(n-2)!}$ summands in (A.25) all of which are bounded by $\mathbb{E}\big[u^*(X_1, X_2, Y_1)\big]^2$ due to the Cauchy-Schwarz inequality, so that altogether

$$\begin{aligned}
&\mathbb{E}\Big[U_{m,n} - \hat{U}_{m,n}\Big]^2 \\
&\quad \leq m^{-2}(m-1)^{-2}n^{-2} \cdot \left[\frac{4m! \cdot n}{(m-3)!} + \frac{2m! \cdot n}{(m-2)!} + \frac{2m! \cdot n!}{(m-2)!(n-2)!}\right] \cdot \mathbb{E}\big[u^*(X_1, X_2, Y_1)\big]^2 \\
&\quad = m^{-2}(m-1)^{-2}n^{-2} \cdot \Big[4m(m-1)(m-2)n + 2m(m-1)n + 2m(m-1)n(n-1)\Big] \\
&\qquad \times \mathbb{E}\big[u^*(X_1, X_2, Y_1)\big]^2 \\
&\quad = \left[\frac{4(m-2) + 2 + 2(n-1)}{m^{-1}(m-1)^{-1}n^{-1}}\right] \cdot \mathbb{E}\big[u^*(X_1, X_2, Y_1)\big]^2 \\
&\quad = \left[\frac{4(m-2) + 2n}{m^{-1}(m-1)^{-1}n^{-1}}\right] \cdot \mathbb{E}\big[u^*(X_1, X_2, Y_1)\big]^2 \\
&\quad = \Big[O(m^{-1}n^{-1}) + O(m^{-2})\Big] \cdot \mathbb{E}\big[u^*(X_1, X_2, Y_1)\big]^2.
\end{aligned}$$

□

Bibliography

Behnen, K. (1972). A characterization of certain rank-order tests with bounds for the asymptotic relative efficiency. *The Annals of Mathematical Statistics*, 43(6):1839–51.

Behnen, K. and Hušková, M. (1984). A simple algorithm for the adaptation of scores and power behavior of the corresponding rank test. *Communications in Statistics - Theory and Methods*, 13(3):305–25.

Behnen, K. and Neuhaus, G. (1983). Galton's test as a linear rank test with estimated scores and its local asymptotic efficiency. *The Annals of Statistics*, 11(2):588–99.

Behnen, K. and Neuhaus, G. (1989). *Rank Tests with Estimated Scores and Their Application*. Teubner, Stuttgart.

Behnen, K., Neuhaus, G., and Ruymgaart, F. (1983). Two sample rank estimators of optimal nonparametric score-functions and corresponding adaptive rank statistics. *The Annals of Statistics*, 11(4):1175–1189.

Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests*. Academic Press, New York, 1st edition.

Königsberger, K. (2004). *Analysis 1*. Springer Verlag, Berlin/Heidelberg, 6th edition.

Neuhaus, G. (1987). Local asymptotics for linear rank statistics with estimated score functions. *The Annals of Statistics*, 15(2):491–512.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83.

Eidesstattliche Erklärung

Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Ich stimme einer evtl. Überprüfung meiner Dissertation durch eine Antiplagiat-Software zu.

Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der *Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis* niedergelegt sind, eingehalten.

Marburg, den 06. Januar 2017

Brandon Greene