Test for the Model Selection from Two Competing Distribution Classes

Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

Fakultät Naturwissenschaften

Universität Hohenheim

Institut für Angewandte Mathematik und Statistik

Erscheinungsjahr: 2016

Autor: Hong Chen

Dekan: Prof. Dr. rer. nat. Heinz Breer

1.berichtende Person: Prof. Dr. Uwe Jensen
2.berichtende Person: Prof. Dr. Harro Walk

Mündliche Prüfung am: 2016/08/18

Firstly, I would like to express my sincere gratitude to my advisor Prof. Dr. Uwe Jensen for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.
My sincere thanks also goes to Dr. Maik Döring, Dr. André Erhardt, and Dr. Torsten Stefan for their insightful comments and encouragement.
Last but not the least, I would like to thank the DFG (German Research Foundation) for the support and funding of the research project (Ge: Je 162 / 10-1, Schi 457 / 12-1).

Contents

Chapter 1

Introduction

1.1 The Problem

One of the main tasks in statistics is to allocate an appropriate parametric distribution function to a given set of data. Statistical methods to do it include goodness-of-fit testing, selection procedures and model selection testing.

In all these methods a divergence measure has to be defined to describe the goodness-of-fit of the model to the data. A great number of divergence measures have been proposed in the literature such as f-divergences, Bregman divergences, α-divergences, Kullback-Leibler discrepancy, Kolmogorov-Smirnov discrepancy, Anderson-Darling discrepancy, Cramér-von Mises discrepancy and so on, see Basseville (2010) for a summary. Clearly, all of them may be used to construct goodness-of-fit tests, selection procedures or model selection tests.

In the context of goodness-of-fit test for the distribution function, the Kolmogorov-Smirnov discrepancy is used, for instance, in Durbin (1973, 1975, 1985), Khmaladze (1981) and Wooldridge (1990). These tests were extended to the case with right censored data by Sun (1997), Nikabadze and Stute (1997). For the case with covariates in random design settings, Andrews (1997) used the Kolmogorov-Smirnov distance, while Li and Tkacz (2011), Ducharme and Ferrigno (2012), Rothe and Wied (2013) applied the Cramér-von Mises distance to construct tests. Tests for conditional density functions based on the Kullback-Leibler information criterion and for conditional distributions have been considered in J.X. Zheng (2000) and X. Zheng (2012).

In terms of model selection procedures for density functions, the "Akaike information criterion" (Akaike (1973)) and "Bayesian information criterion" (Schwartz (1978)) are used mostly, that are both based on the Kullback-Leibler information criterion, see Claeskens and Hjort (2008) for a summary.

Generally, selection procedures are simple to apply, but it does not give the degree of confidence in the choice, while the model selection testing can control the extent of the certainty of the decision by adjusting the confidence level. In terms of the model selection testing between two parametric density models, the Kullback-Leibler information criterion is the most investigated divergence measure in the literature. The resulting test is the so-called likelihood ratio test, which was considered for instance by Nishii (1988), Vuong (1989), Sin and White (1996), Inoue and Kilian (2006), Shi (2015a). These tests have also been generalized to moment-based models by Kitamura (2001), Chen et al. (2007) and Shi (2015b). As Chen et al. (2007) pointed out alternative discrepancy measures that measure goodness-of-fit might be preferred in some applications.

In practice, which criterion to choose should depend on the aim of the estimation. In this thesis, we are interested in the estimation of the distribution function of the data. Thus, it is reasonable to use some criterion like Kolmogorov-Smirnov discrepancy, Anderson-Darling discrepancy, Cramér-von Mises discrepancy between the distribution functions. However, model selection tests based on these discrepancies were seldom used in the literature with some exceptions. For instance, Liebscher (2014) proposed a model selection test based on the Anderson-Darling distance in the case of i.i.d. data. For the case with covariates, Ng and Joe (2016) extend Vuong's (1989) tests with a generalized measure of distance, however, lots of measures are not included among others Kolmogorov-Smirnov discrepancy, Anderson-Darling discrepancy and Cramér-von Mises discrepancy. Recently, Chen et al. (2015) proposed a test based on the Cramér-von Mises distance.

In this thesis, we will extend the model selection tests in Chen et al. (2015) to the case with multi-dimensional covariates and right random censoring in a fixed design setting. Both censoring and fixed design in the context of the model selection from two competing distribution function models were rarely considered before, thus, this thesis can fill this gap.

Let $z \in \mathbb{R}^d$ represent the d dimensional vector of covariate with $d \in \mathbb{N}$ and $X_z \in \mathbb{R}$ the random variable at the covariate value z. Without loss of generality, we assume $z \in [0,1]^d$. The distribution function of X_z is denoted as $H(\cdot|z)$. Let $z_1, ..., z_{n_0}$ be the predetermined covariate values (fixed design). In particular, we assume $n_0 = \bar{n}_0^d$ with $\bar{n}_0 \in \mathbb{N}$ and the covariate values are equidistant grid points on $[0,1]^d$, i.e.

$$\{z_1, z_2, ..., z_{n_0}\} := \Big\{\Big(\frac{i_1}{\bar{n}_0}, \frac{i_2}{\bar{n}_0}, ..., \frac{i_d}{\bar{n}_0}\Big) : 1 \leq i_1, ..., i_d \leq \bar{n}_0, i_1, ..., i_d \in \mathbb{N}\Big\}.$$

Denote $X_1, ..., X_{n_0}$ as the corresponding independent random variables instead of $X_{z_1}, ..., X_{z_{n_0}}$. Further, for each $j \in \{1, ..., m-1\}$ with $m \in \mathbb{N}$, let $(X_{j\cdot n_0+1}, z_{j\cdot n_0+1}), ..., (X_{j\cdot n_0+n_0}, z_{j\cdot n_0+n_0})$ be i.i.d. copy of

$$(X_1, z_1), ..., (X_{n_0}, z_{n_0}),$$

i.e. for any $i \in \{1, ..., n_0\}$,

$$z_i = z_{n_0+i} = ... = z_{(m-1)\cdot n_0+i}$$

and $X_i, X_{n_0+i}, ..., X_{(m-1)\cdot n_0+i}$ are i.i.d. random variables. Let $n := n_0 \cdot m$ be the sample size, then we have the data set:

$$(X_1, z_1), ..., (X_n, z_n)$$

with n_0 different covariate values and m observations at each covariate value, i.e.

$$\begin{array}{llll} (X_1, z_1), & (X_2, z_2), & \ldots, & (X_{n_0}, z_{n_0}), \\ (X_{n_0+1}, z_1), & (X_{n_0+2}, z_2), & \ldots, & (X_{2n_0}, z_{n_0}), \\ & \cdots & & \\ (X_{(m-1)\cdot n_0+1}, z_1), & (X_{(m-1)\cdot n_0+2}, z_2), & \ldots, & (X_{m\cdot n_0}, z_{n_0}). \end{array}$$

The data structure in this thesis is inspired by a case study in which endurance tests on DC-motors under different load levels were conducted at the Institute of Design and Production in Precision Engineering of the University of Stuttgart, see Bobrowski et al. (2011, 2015) and the case study in Chapter 4 of this thesis. For each predetermined load level, the lifetimes of 16 DC-motors have been observed.

We consider two potential parametric model classes of distributions denoted by

$$
\begin{aligned}
\mathcal{F} =& \{F(\cdot|\theta, z) : \theta \in \Theta \subset \mathbb{R}^p, z \in [0,1]^d\}, \\
\mathcal{G} =& \{G(\cdot|\gamma, z) : \gamma \in \Gamma \subset \mathbb{R}^q, z \in [0,1]^d\},
\end{aligned}
$$

where Θ and Γ are compact intervals and $p, q \in \mathbb{N}$. For instance, the two distribution function model classes can be Weibull and log-normal distribution classes. The aim of this thesis is to propose model selection tests to answer the question which of the two model classes approximates the underlying family of distributions H better in different settings based on the Cramér-von Mises distance. The proposed tests in this thesis are consistent in the sense that with increasing number of data the tests lead to the model with closer distance to the underlying distribution function with probability approaching one.

In the remaining sections of Chapter 1 some basic concepts in statistics are introduced like the maximum likelihood estimation theory, kernel estimator of distribution function and right random censoring. These concepts will be used and extended in the main part of this thesis.

In Chapter 2 the Cramér-von Mises distance between the underlying distribution H and the competing parametric model classes will be introduced based on the maximum likelihood theory. Then the hypotheses are given for the model selection test. Further the test statistics will be defined and their asymptotic behavior will be derived for the cases with $m \to \infty$ and n_0 fixed or with m fixed $n_0 \to \infty$. In the end the decision rules will be formulated.

The results in Chapter 2 will be extended to the case with right random censoring in Chapter 3. Among other tools, the Kaplan-Meier estimator and Beran estimator are used.

Chapter 4 contains a case study using the data form endurance tests on DC-motors at the Institute of Design and Production in Precision Engineering of the University of Stuttgart.

In Chapter 5 simulation studies are carried out to show the performance of the test procedure with moderate sample size.

At the end of this thesis, the extension possibilities of the proposed tests will be discussed in a conclusion in Chapter 6. Some auxiliary lemmas are postponed to the Appendix A.

1.2 Notations

In this section, we introduce some notations which will be used through out this thesis.

For $a, b \in \mathbb{R}$, $x = (x_1, ..., x_d)^T$, $y = (y_1, ..., y_d)^T \in \mathbb{R}^d$ with $d \in \mathbb{N}$, denote

$$\lceil a \rceil := \max\{k : k \leq a, k \in \mathbb{Z}\}, \quad a \wedge b := \max(a, b),$$

$$a \cdot \mathbb{N} := \{a \cdot k : k \in \mathbb{N}\}, \quad |x| := (|x_1|, ..., |x_d|)^T,$$

$$x + y := (x_1 + y_1, ..., x_d + y_d)^T, \quad a \cdot x := (a \cdot x_1, ..., a \cdot x_d)^T,$$

further we write $x \leq y$ if $x_i \leq y_i$ holds for all $i \in \{1, ..., d\}$ and the indicator

$$I(x \leq y) := \begin{cases} 1, & x \leq y, \\ 0, & \text{otherwise.} \end{cases}$$

For real valued vectors and matrices $\|\cdot\|$ denotes the maximum norm. For any $i \in \{1, ..., n\}$, define the indicator function $\delta_i : [0,1]^d \to \{0,1\}$ with

$$\delta_i(z) := I(z_i = z).$$

For any function $\psi : \Theta \to \mathbb{R}$, let

$$\dot{\psi} := \left(\partial\psi/\partial\theta_1, ..., \partial\psi/\partial\theta_p\right)^T$$

be the column vector of the first partial derivatives of ψ with respect to θ. Further let $\ddot{\psi}$ denote the matrix of the second partial derivatives of ψ with respect to θ.

For a sequence of real valued random variables $(X_n)_{n\in\mathbb{N}}$ defined on a probability space $(\Omega, \mathcal{F}, P)$, we write

$$X_n \xrightarrow{d} \mathcal{N}(\mu, \sigma^2),$$

if X_n converges to some normally distributed random variable with expectation μ and variance σ^2 in distribution, as $n \to \infty$. Further, we write

$$X_n \xrightarrow{a.s.} X,$$

if X_n converges to the random variable X almost surely, as $n \to \infty$.

For the sequences of constants $(a_n)_{n\in\mathbb{N}}$, nonzero constants $(b_n)_{n\in\mathbb{N}}$ and real valued random variables $(X_n)_{n\in\mathbb{N}}$, the notation $a_n = o(b_n)$ means that $a_n/b_n \to$

0, the notation $a_n = O(b_n)$ means that the sequence $(a_n/b_n)_{n\in\mathbb{N}}$ is bounded. Further, $X_n = o_p(b_n)$ means that the sequence of values X_n/b_n converges to zero in probability as $n \to \infty$. The notation $X_n = O_p(b_n)$ means that the sequence of values $(X_n/b_n)_{n\in\mathbb{N}}$ is stochastically bounded, i.e. for any $\varepsilon > 0$, there exists a finite $M > 0$ such that for eventual all $n \in \mathbb{N}$

$$P(|X_n/b_n| > M) < \varepsilon.$$

The right endpoint of a distribution function F is defined as

$$\tau_F := \inf\{x : F(x) = 1\} \in (-\infty, +\infty].$$

To simplify the notation, we will use a generic constant $C > 0$ in the proofs, i.e., the value of C might be different in each term containing C. Further, we assume that the notations defined in the proof of a lemma or theorem is only valid within that particular proof.

1.3 Maximum Likelihood Theory

Let $X_1, ..., X_n$ be real valued i.i.d. random variables with $n \in \mathbb{N}$, their distribution function H can be estimated by the empirical distribution function

$$H_n(x) := \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq x), \tag{1.3.1}$$

for $x \in \mathbb{R}$. The properties of the function H_n are well investigated, see for example Van der Vaart (1998). We list here some of them. First, for each $x \in \mathbb{R}$

$$H_n(x) \xrightarrow{a.s.} H(x).$$

A stronger result, called the Glivenko-Cantelli theorem, states that the convergence holds uniformly over x, i.e.

$$\sup_{x\in\mathbb{R}} \big|H_n(x) - H(x)\big| \xrightarrow{a.s.} 0.$$

Further, the central limit theorem states the pointwise asymptotic normality:

$$\sqrt{n} \cdot \big(H_n(x) - H(x)\big) \overset{d}{\to} \mathcal{N}\big(0, H(x)\big(1 - H(x)\big)\big),$$

for $x \in \mathbb{R}$. These convergence properties of H_n can be extended to the so-called empirical integrals:

$$\int \psi(x) dH_n(x) = \frac{1}{n} \sum_{i=1}^{n} \psi(X_i),$$

where $\psi : \mathbb{R} \to \mathbb{R}$ is a given function. Notice that for any $x \in \mathbb{R}$,

$$H_n(x) = \int I(u \leq x) dH_n(u).$$

Hence, the empirical integral is a generalization of the empirical distribution function. If $\int |\psi(x)| dH(x) < \infty$, the strong law of large numbers yields

$$\int \psi(x) dH_n(x) \xrightarrow{a.s.} \int \psi(x) dH(x).$$

While under $\int \psi^2(x) dH(x) < \infty$, the central limit theorem gives

$$\sqrt{n} \cdot \Big(\int \psi(x) dH_n(x) - \int \psi(x) dH(x) \Big) \xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

where

$$\sigma^2 = \int \psi^2(x) dH(x) - \Big(\int \psi(x) dH(x) \Big)^2.$$

Another possibility to approximate the distribution function H is to use some parametric distribution model classes. For instance, if the random variables $X_1, ..., X_n$ represent some kind of lifetimes, the model class is often assumed to be exponential, Weibull or log-normal.

Denote the distribution model class by

$$\{F(\cdot|\theta) : \theta \in \Theta \subset \mathbb{R}^p, p \in \mathbb{N}\}.$$

Suppose the function $F(\cdot|\theta)$ has the density function $f(\cdot|\theta)$ for all $\theta \in \Theta$, an estimate for the parameter is the maximum likelihood estimator, which is defined as

$$\hat{\theta}_n := \underset{\theta \in \Theta}{\operatorname{argmax}} \frac{1}{n} \sum_{i=1}^{n} \log f(X_i|\theta).$$

For any function $\psi : \mathbb{R} \times \Theta \to \mathbb{R}$ with $(x, \theta) \mapsto \psi(x, \theta)$, we refer to it as dominated by an H integrable function, if there exists a function $M : \mathbb{R} \to \mathbb{R}$, such that $|\psi(x, \theta)| \leq M(x)$ for all $(x, \theta) \in \mathbb{R} \times \Theta$ and

$$\int M(x) dH(x) < \infty.$$

For the consistency and asymptotic normality of the maximum likelihood estimator, we make the following assumptions.

A1 The density $f(\cdot|\theta)$ is strictly positive H-a.s. for all $\theta \in \Theta$.

A2 The set Θ is compact and the function $\log f$ is twice continuously differentiable on Θ.

A3 The functions $\log f$ and $\|\partial^2 \log f/\partial\theta^2\|$ are dominated by H integrable functions.

A4 The function $\int \log f(x|\cdot)dH(x)$ has a unique maximum on Θ at θ_*, where θ_* is an interior point of Θ.

A5 The function $\|\partial \log f/\partial\theta\|^2$ is dominated by an H integrable function and the Hessian matrix

$$\int \partial^2 \log f(x|\theta_*)/\partial\theta^2 dH(x)$$

is invertible.

It follows from White (1982) that if A1–A5 hold, then

$$\hat{\theta}_n \xrightarrow{a.s.} \theta_*$$

and $\sqrt{n} \cdot (\hat{\theta}_n - \theta_*)$ convergences to a multi-dimensional normal distribution. The vector θ_* is called pseudo-true value for the parametric model class. In the case that there exists a θ_0, such that $H(\cdot) = F(\cdot|\theta_0)$, it holds $\theta_* = \theta_0$.

1.4 Censored Data

It is well-known that in practical studies the observation of the survival of a patient is subject to right censoring. Classical example of this type of censoring is that the patient died from other causes than those under study or the patient is still alive by the time of the end of the study.

Let $X_1, ..., X_n$ be positive real valued i.i.d. random variables with distribution function H representing the lifetime time of an individual. Let $C_1, ..., C_n$ be real valued i.i.d. random variables with distribution function J representing the random censoring times. The observable random variables are

$$Y_i := \min(X_i, C_i) \quad \text{and} \quad \Delta_i := I(X_i \le C_i),$$

for $i \in \{1, ..., n\}$. The $0-1$ valued variable Δ_i indicates whether Y_i is a censored time ($\Delta_i = 0$) or not ($\Delta_i = 1$). Denote the distribution function of Y_i by B. We assume that $X_1, ..., X_n$ and $C_1, ..., C_n$ are independent, thus for $x \in \mathbb{R}$

$$B(x) = 1 - \big(1 - H(x)\big)\big(1 - J(x)\big).$$

The distribution function H can be estimated by Kaplan-Meier product-limit estimate (1958):

$$H_n^{KM}(x) := 1 - \prod_{Y_{(i)} \leq x} \big(1 - \frac{\Delta_{(i)}}{n - i + 1}\big),$$

where $Y_{(1)} \leq ... \leq Y_{(n)}$ are the ordered $Y_1, ..., Y_n$, and $\Delta_{(1)}, ..., \Delta_{(n)}$ are the corresponding indicators to $Y_{(1)}, ..., Y_{(n)}$. Note that the Kaplan-Meier estimator is a step function and has jumps only at the uncensored observations. Further, in the case of no censoring, it reduces to the empirical distribution function.

The convergence properties of the Kaplan-Meier estimator have been investigated in many papers. Földes and Rejto (1981) showed the strong uniform consistency of the Kaplan-Meier estimator. Lo and Singh (1986) obtained an asymptotic representation, which decomposes $H_n^{KM}(x) - H(x)$ in an average of i.i.d. terms and a remainder term converging to zero in probability. Based on that representation the asymptotic normality was derived.

In terms of the so-called Kaplan-Meier integrals: $\int \psi(x) dH_n^{KM}(x)$ for a given function $\psi : \mathbb{R} \to \mathbb{R}$, Gill (1983) proved its convergence in distribution under the condition that ψ is a non-negative, continuous and nonincreasing function. Under the same conditions for ψ, Schick et al.(1988) obtained a representation of $\int \psi(x) dH_n^{KM}(x)$ as a sum of i.i.d. random variables plus a remainder. Both of their methods are based on integration by parts. Under some regularity conditions on H, Yang (1994) and Akritas (2000) extended the convergence of $\int \psi(x) dH_n^{KM}(x)$, to those functions ψ satisfying

$$\int_0^{\tau_B} \frac{\psi^2(x)}{1 - J(x)} dH(x) < \infty.$$

In a more general setting, Stute and Wang (1993) pointed out that we can write

$$H_n^{KM}(x) = \sum_{i=1}^{n} W_{in} \cdot I(X_i \leq x),$$

where

$$W_{in} := \frac{\Delta_{(i)}}{n-i+1} \prod_{j=1}^{n-1} \Big(1 - \frac{\Delta_{(j)}}{n-j+1}\Big).$$

Based on this expression, they showed that under the condition $\int_0^{\tau_B} |\psi(x)| dH(x) < \infty$, it holds

$$\int \psi(x) dH_n^{KM}(x) \xrightarrow{a.s.} \int_0^{\tau_B} \psi(x) dH(x).$$

The asymptotic normality was shown by Stute (1995): if

$$\int_0^{\tau_B} \frac{\psi^2(x)}{1-J(x)} dH(x) < \infty \quad \text{and} \quad \int_0^{\tau_B} \psi(x,z) C^{1/2}(x) dH(x) < \infty,$$

where

$$C(x) := \int_0^{x-} \frac{1}{\big(1-B(u)\big)\big(1-J(u)\big)} dJ(u) := \int \frac{I(u<x)}{\big(1-B(u)\big)\big(1-J(u)\big)} dJ(u),$$

then

$$\sqrt{n} \cdot \Big(\int \psi(x) dH_n^{KM}(x) - \int_0^{\tau_B} \psi(x) dH(x) \Big) \xrightarrow{d} \mathcal{N}(0, \sigma_1^2),$$

where

$$\sigma_1^2 = \int_0^{\tau_B} \frac{\psi^2(x)}{1-J(x)} dH(x) - \Big(\int_0^{\tau_B} \psi(x) dH(x) \Big)^2 - \int_0^{\tau_B} \Big(\int_x^{\tau_B} \psi(u) dH(u) \Big)^2 \cdot \frac{1-H(x)}{\big(1-B(x)\big)^2} dJ(x).$$

Note that in the case without censoring ($J(x) = 0$, for all x), σ_1^2 reduces to σ^2 as defined in Section 1.3. These results will be applied and extended to the case with covariates in this thesis.

1.5 Kernel Estimate for Conditional Distribution Function

Given the data $(X_1, z_1), ..., (X_n, z_n) \in \mathbb{R} \times [0,1]^d$ with $n \in \mathbb{N}$ as defined in Section 1.1. Suppose the distribution function $H(\cdot|z)$ does not change too fast with respect to z, then the distribution of X_i and X_j should be close if z_i and z_j are close. This motivates the construction of estimation of $H(\cdot|z)$ for a z value with help of those X_i, for which z_i are close to z.

For each $i \in \{1, ..., n\}$, denote the weight function $w_{ni} : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^+$ by

$$w_{ni}(z, h) := \frac{K(\frac{z_i - z}{h})}{\sum_{k=1}^n K(\frac{z_k - z}{h})},$$

where the function $K : \mathbb{R}^d \to \mathbb{R}^+$ is called kernel function and $h > 0$ bandwidth. The Nadaraya-Watson kernel estimate (Nadaraya (1964), Watson (1964)) is defined by

$$\hat{H}_n(x|z) := \sum_{i=1}^n w_{ni}(z, h) \cdot I(X_i \leq x). \tag{1.5.1}$$

The bandwidth h controls the smoothness of the estimate. Common choices for K are, for instance, uniform and Epanechnikov kernel, which are defined by

$$K(x) := \frac{1}{2} \cdot I(|x| \leq 1) \quad \text{and} \quad K(x) := \frac{3}{4}(1 - x^2) \cdot I(|x| \leq 1),$$

respectively.

Another possibility to choose weights is the so called Gasser-Müller weights (Gasser and Müller (1984)). In our setting, if z is one dimensional ($d = 1$) and $m = 1$, the weights of the Gasser-Müller weights are defined by

$$w_{ni}(z, h) := \int_{z_{i-1}}^{z_i} \frac{1}{h} K\Big(\frac{z - u}{h}\Big) du,$$

with $z_0 = 0$.

The convergence properties of the kernel estimator for distribution function were shown in for example Aerts et al.(1994), Györfi et al.(2002) and Li and Racine (2007). The main conditions for the asymptotic properties are: first, the distribution function H is differentiable with respect to z, so $H(\cdot|z)$ does not change too fast in z and can be estimated with data at z_i close to z. Secondly, $n_0 h \to \infty$ and $h \to 0$ as $n_0 \to \infty$, i.e. the number of data in any fixed small interval tends to infinity.

In the case with censoring, for each $i \in \{1, ..., n\}$ we denote the right censoring random variable at z_i as C_i with distribution function $J(\cdot|z_i)$. Further denote

$$Y_i := \min(X_i, C_i) \quad \text{and} \quad \Delta_i := I(X_i \leq C_i).$$

Beran (1981) introduced a kernel estimator for the conditional distribution. His estimator is a generalization of the Kaplan-Meier estimator and is some-times called conditional Kaplan-Meier estimator:

$$\hat{H}_n^{KM}(x|z) := 1 - \prod_{Y_{(i)} \le x} \Big(1 - \frac{w_{n(i)}(z,h) \cdot \Delta_{(i)}}{1 - \sum_{j=1}^{i-1} w_{n(j)}(z,h)}\Big).$$

Here $Y_{(1)} \le ... \le Y_{(n)}$ denote the ordered $Y_1, ..., Y_n$, while $\Delta_{(1)}, ..., \Delta_{(n)}$ and $w_{n(1)}(z,h), ..., w_{n(n)}(z,h)$ represent the corresponding indicators and weights of $Y_{(1)}, ..., Y_{(n)}$, respectively. Note that in the case without censoring $\hat{H}_n^{KM}(x|z)$ reduces to the kernel estimator $\hat{H}_n(x|z)$ defined in (1.5.1).

The uniform consistency and asymptotic normality of Beran's estimator was studied by Dabrowska (1987, 1989), Mckeague and Utikal (1990), Akritas (1994) and González and Cadarso (1994) in the random design case, where the covariate z is also assumed to be a random variable. It was extend to the case with discrete covariates by Du and Akritas (2002). The fixed design case was investigated by Van Keilegom and Veraverbeke (1996, 1997a, 1997b) using Gasser-Müller weights.

Chapter 2

Model Selection Testing

In this chapter, we assume that the observations (X_1, z_1),..., (X_n, z_n) and the distribution function $H(\cdot|z)$ are defined as in Section 1.1. Given two parametric distribution model classes:

$$\mathcal{F} := \{F(\cdot|\theta, z) : \theta \in \Theta \subset \mathbb{R}^p, z \in [0,1]^d\}$$

and

$$\mathcal{G} := \{G(\cdot|\gamma, z) : \gamma \in \Gamma \subset \mathbb{R}^q, z \in [0,1]^d\},$$

where Θ and Γ are compact intervals and the constant $d, p, q \in \mathbb{N}$, we will construct model selection tests to answer the question which of the two model classes approximates the underlying family of distributions better. The distances of the model classes and the underlying distributions will be defined based on the Cramér von-Mises distances, which is often used in the goodness-of-fit test. The test statistics are defined as the difference of the estimated distances. Asymptotic normality of the test statistics will be proven. Based on this asymptotic behavior decision rules for the tests will be formulated.

This chapter is organized as follows: in Section 2.1, some notations and the hypotheses of the model selection tests for this chapter will be introduced. Section 2.2 deals with the case that n_0 is fixed and $m \to \infty$, i.e. the number of covariates values is fixed and the number of observations at each covariates values tends to infinity. The underlying distribution function H will be estimated by the empirical distribution at each covariates value. In Section 2.3, it is assumed that m is fixed and $n_0 \to \infty$. The empirical distribution function is replaced by the kernel Nadaraya-Watson estimator as defined in Section 1.5.

For simplicity of notation, we assume that the notations defined in Section 2.1 are valid through out this chapter and the notations defined in Section 2.2 and Section 2.3 are only valid in that particular section.

2.1 Notations and Hypotheses

In this section, we will introduce a distance measure between the underlying distribution and the given model classes based on the likelihood theory. First, we define the joint distribution function by $Q : \mathbb{R} \times [0,1]^d \to [0,1]$ with

$$Q(x,z) := \int \int I(u \leq x) \cdot I(v \leq z) dH(u|v) dv,$$

where the inner integration is with respect to the variable u. The empirical distribution function at covariate z and the joint empirical distribution are then given by $H_n, Q_n : \mathbb{R} \times [0,1]^d \to [0,1]$ with

$$H_n(x|z) := \frac{1}{m} \sum_{i=1}^{n} \delta_i(z) \cdot I(X_i \leq x),$$

where $\delta_i(z) := I(z_i = z)$ and

$$Q_n(x,z) := \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x) \cdot I(z_i \leq z) = \frac{1}{n_0} \sum_{i=1}^{n_0} H_n(x|z_i) \cdot I(z_i \leq z).$$

For any function ψ: $\mathbb{R} \times \Theta \times [0,1]^d \to \mathbb{R}$, we get then

$$\int \psi(x,\theta,z) dQ(x,z) = \int \int \psi(x,\theta,z) dH(x|z) dz,$$
$$\int \psi(x,\theta,z) dQ_n(x,z) = \frac{1}{n} \sum_{i=1}^{n} \psi(X_i,\theta,z_i) = \frac{1}{n_0} \sum_{i=1}^{n_0} \int \psi(x,\theta,z_i) dH_n(x|z_i).$$

Denote the logarithmic likelihood function for the model class $\mathcal{F}$ as $\hat{L}_{f,n} : \Theta \to \mathbb{R}$ with

$$\hat{L}_{f,n}(\theta) := \int \log f(x|\theta,z) dQ_n(x,z),$$

where for each $(\theta, z) \in \Theta \times [0,1]^d$, the function $f(\cdot|\theta,z)$ denotes the density function of $F(\cdot|\theta,z)$. The maximum likelihood estimator $\hat{\theta}_n$ for the model class $\mathcal{F}$ is defined as a measurable selection:

$$\hat{\theta}_n := \operatorname*{argmax}_{\theta \in \Theta} \hat{L}_{f,n}(\theta).$$

By the compactness of the set Θ, the estimator $\hat{\theta}_n$ exists, if for any $(x,z) \in \mathbb{R} \times [0,1]^d$ the function $f(x|\cdot, z)$ is continuous in θ. Further, we define the functions $L_{f,n_0}, L_{f,\infty} : \Theta \to \mathbb{R}$ and the vectors $\theta_{n_0}, \theta_* \in \Theta$ by

$$\begin{aligned} L_{f,n_0}(\theta) &:= \frac{1}{n_0} \sum_{i=1}^{n_0} \int \log f(x|\theta, z_i)\, dH(x|z_i), \\ \theta_{n_0} &:= \operatorname*{argmax}_{\theta \in \Theta} L_{f,n_0}(\theta), \\ L_{f,\infty}(\theta) &:= \int \log f(x|\theta, z)\, dQ(x,z), \\ \theta_* &:= \operatorname*{argmax}_{\theta \in \Theta} L_{f,\infty}(\theta). \end{aligned}$$

We will show in the next two sections that under some regularity conditions the following relations hold:

$$\begin{array}{ccc} \hat{L}_{f,n} & \xrightarrow{m \to \infty} & L_{f,n_0} \\ {}_{n_0 \to \infty}\searrow & & \swarrow_{n_0 \to \infty} \\ & L_{f,\infty}. & \end{array}$$

And we have then

$$\begin{array}{ccc} \hat{\theta}_n & \xrightarrow{m \to \infty} & \theta_{n_0} \\ {}_{n_0 \to \infty}\searrow & & \swarrow_{n_0 \to \infty} \\ & \theta_*. & \end{array}$$

We define the distance $d_H(\mathcal{F})$ between the underlying family of distribution functions H and the model class $\mathcal{F}$ as

$$\frac{1}{n_0} \sum_{i=1}^{n_0} \int \big(H(x|z_i) - F(x|\theta_{n_0}, z_i)\big)^2\, dH(x|z_i), \tag{2.1.1}$$

for the case with n_0 fixed and $m \to \infty$ and

$$\int \big(H(x|z) - F(x|\theta_*, z)\big)^2\, dQ(x,z) \tag{2.1.2}$$

for the case with m fixed and $n_0 \to \infty$, respectively.

Let g, $\hat{\gamma}_n$, γ_{n_0}, γ_* and $d_H(\mathcal{G})$ denote the counterparts for the model class $\mathcal{G}$. We will propose model selection tests of the null hypothesis

$$\mathcal{H}^0 : \quad d_H(\mathcal{F}) = d_H(\mathcal{G})$$

meaning that the two models are equally close to H, against

$$\mathcal{H}^{\mathcal{F}}: \quad d_H(\mathcal{F}) < d_H(\mathcal{G})$$

meaning H is closer to $\mathcal{F}$ than to $\mathcal{G}$ or

$$\mathcal{H}^{\mathcal{G}}: \quad d_H(\mathcal{F}) > d_H(\mathcal{G})$$

meaning H is closer to $\mathcal{G}$ than to $\mathcal{F}$.

We call a function $\psi : \mathbb{R} \times \Theta \times [0,1]^d \to \mathbb{R}$ with $(x, \theta, z) \mapsto \psi(x, \theta, z)$ as H integrable or H square integrable, if for each $(\theta, z) \in \Theta \times [0,1]^d$,

$$\int \psi(x, \theta, z) dH(x|z) < \infty \quad \text{or} \quad \int \psi^2(x, \theta, z) dH(x|z) < \infty$$

holds.

We refer the function ψ as dominated by an H integrable function, if there exists an H integrable function $M : \mathbb{R} \times [0,1]^d \to \mathbb{R}$, such that $|\psi(x, \theta, z)| \leq M(x, z)$ for all $(x, \theta, z) \in \mathbb{R} \times \Theta \times [0,1]^d$.

If there exists a function $M : \mathbb{R} \to \mathbb{R}$ such that for all $(x, \theta, z) \in \mathbb{R} \times \Theta \times [0,1]^d$,

$$\left|\psi(x, \theta, z)\right| < M(x) \quad \text{and} \quad \int M(x) dx < \infty.$$

we call ψ dominated by a Lebesgue integrable function independent of z.

If there exists a function $M : \mathbb{R} \to \mathbb{R}$ such that for all $(x, \theta, z) \in \mathbb{R} \times \Theta \times [0,1]^d$,

$$\left|\psi(x, \theta, z)\right| < M(x) \quad \text{and} \quad \int M(x) dH(x|z) < \infty,$$

we refer ψ as dominated by an H integrable function independent of z.

Further the domination by an H square integrable function (independent of z) is defined analogously.

In Section 2.2, we assume all the convergences are taken by letting $m \to \infty$. In Section 2.3, we assume all the convergences are taken by letting $n_0 \to \infty$. Note that since $n = m \cdot n_0$, in Section 2.2, $n \in n_0 \cdot N := \{n_0 \cdot a : a \in N\}$, in Section 2.3, $n \in m \cdot N := \{m \cdot a : a \in N\}$, in both cases $n \to \infty$.

2.2 The Case with Number of Observations at Each Covariate Tending to Infinity

In this section, the distance $d_H(\mathcal{F})$ defined in (2.1.1) is estimated by

$$\hat{d}_{H,n}(\mathcal{F}) := \int \left(H_n(x|z) - F(x|\hat{\theta}_n, z)\right)^2 dQ_n(x,z).$$

For the class $\mathcal{G}$, the estimator $\hat{d}_{H,n}(\mathcal{G})$ is defined in an analogous way. As test statistic we take the difference of the estimated distances

$$T_n := \hat{d}_{H,n}(\mathcal{F}) - \hat{d}_{H,n}(\mathcal{G}).$$

The main results of this section is the asymptotic normality of $\sqrt{n}\cdot T_n$ and the determination of a consistent estimator for the asymptotic variance of $\sqrt{n} \cdot T_n$. Based on these results, decision rules for the model selection test will be formulated. In this section we make the following assumptions. They are stated in terms of the model class $\mathcal{F}$, it is understood that corresponding assumptions are also made for the model class $\mathcal{G}$.

B1 For each $(\theta, z) \in \Theta \times [0,1]^d$, the density function $f(\cdot|\theta, z) : \mathbb{R} \to \mathbb{R}$ is strictly positive $H(\cdot|z)$-a.s.

B2 The function $\log f$ is three times continuously differentiable in θ on Θ.

B3 The function $\log f$ is dominated by an H integrable function.

B4 The function L_{f,n_0} has a unique maximum on Θ at θ_{n_0}, which is an interior point of Θ.

B5 The functions $\|\partial \log f/\partial\theta\|$ and $\|\partial^2 \log f/\partial\theta^2\|$ are dominated by H square integrable functions. The Hessian matrix $\ddot{L}_{f,n_0}(\theta_{n_0})$ is invertible with inverse $\ddot{L}^{-1}_{f,n_0}(\theta_{n_0})$.

B6 For any $i,\ j,\ k \in \{1, 2, ..., p\}$, the function $\partial^3 \log f/\partial\theta_i\partial\theta_j\partial\theta_k$ is dominated by an H integrable function.

B7 The functions $\dot{F}$ and $\ddot{F}$ exist and they are bounded.

These assumptions are regular assumptions in the framework of the maximum likelihood theory. The asymptotic properties of $\hat{\theta}_n$, which we will show in this section, can be reached also under weaker conditions. However, the focus of this thesis is model selection test, therefore we use the more restrictive conditions to avoid technical difficulties.

Lemma 2.2.1. *Define the function $\psi : \mathbb{R} \times [0,1]^d \to \mathbb{R}$. If ψ is an H integrable function, then*

$$\int \psi(x,z)dQ_n(x,z) - \frac{1}{n_0}\sum_{i=1}^{n_0} \int \psi(x,z_i)dH(x|z_i) \xrightarrow{a.s.} 0. \tag{2.2.1}$$

If ψ is an H square integrable function, then

$$\sqrt{n} \cdot \Big(\int \psi(x,z)dQ_n(x,z) - \frac{1}{n_0}\sum_{i=1}^{n_0} \int \psi(x,z_i)dH(x|z_i)\Big) \xrightarrow{d} \mathcal{N}(0,\sigma^2), \tag{2.2.2}$$

where

$$\sigma^2 = \frac{1}{n_0}\sum_{i=1}^{n_0} \Big(\int \psi^2(x,z_i)dH(x|z_i) - \Big(\int \psi(x,z_i)dH(x|z_i)\Big)^2\Big).$$

Proof. We denote first for each $i \in \{1, ..., m\}$,

$$U_i := \frac{1}{n_0}\sum_{j=1}^{n_0} \psi(X_{(i-1)\cdot n_0+j}, z_{(i-1)\cdot n_0+j}).$$

Note that $U_1, ..., U_m$ are i.i.d. and we can write

$$\int \psi(x,z)dQ_n(x,z) = \frac{1}{m}\sum_{i=1}^{m} U_i.$$

Further, the expectation

$$E\Big[\int \psi(x,z)dQ_n(x,z)\Big] = \frac{1}{n_0}\sum_{i=1}^{n_0} \int \psi(x,z_i)dH(x|z_i)$$

and by independence of $X_1, ..., X_n$ the variance

$$Var\Big[\sqrt{n} \cdot \int \psi(x,z)dQ_n(x,z)\Big] = \frac{1}{n}\sum_{i=1}^{n} Var\big[\psi(X_i,z_i)\big] = \sigma^2.$$

Therefore, the assertions follow from the strong law of large numbers and central limit theorem for i.i.d data. □

Based on Lemma 2.2.1, we will show in the next two lemmas the consistency of the maximum likelihood estimator $\hat{\theta}_n$ to θ_{n_0} and the asymptotic properties of $\sqrt{n} \cdot (\hat{\theta}_n - \theta_{n_0})$, respectively.

Lemma 2.2.2. *If* B1–B5 *hold, then* $\|\hat{\theta}_n - \theta_{n_0}\| \to 0$ *a.s.*

Proof. Under B1–B3 the functions $\hat{L}_{f,n}$ and L_{f,n_0} are continuous on Θ. Under B4 the pseudo true value θ_{n_0} is unique and is a well separated maximizer of the function L_{f,n_0}. If we can show

$$\sup_{\theta \in \Theta} \left| \hat{L}_{f,n}(\theta) - L_{f,n_0}(\theta) \right| \to 0 \quad \text{a.s.},$$

then the assertion follows from an argmax theorem, see for example Theorem 2.12 in Kosorok (2008). Under B5, for each $z \in [0,1]^d$, there exists a function $M(\cdot, z) : \mathbb{R} \to \mathbb{R}$ and a constant $C > 0$, such that

$$\sup_{\theta \in \Theta} \|\partial \log f(x|\theta, z)/\partial \theta\| \leq M(x, z)$$

and

$$\frac{1}{n_0} \sum_{i=1}^{n_0} \int M(x, z_i)\, dH(x|z_i) < C.$$

By Lemma 2.2.1 with $\psi(x, z) = M(x, z)$,

$$\int M(x, z)\, dQ_n(x, z) \xrightarrow{a.s.} \frac{1}{n_0} \sum_{i=1}^{n_0} \int M(x, z_i)\, dH(x|z_i).$$

Hence, for eventually all $n \in n_0 \cdot \mathbb{N}$

$$\sup_{\theta \in \Theta} \int \left\| \frac{\partial \log f(x|\theta, z)}{\partial \theta} \right\| dQ_n(x, z) \leq \int M(x, z)\, dQ_n(x, z) < C \quad \text{a.s.} \qquad (2.2.3)$$

Since Θ is compact, for any constant $\varepsilon > 0$ and the constant C above, there exist compact non-empty subsets $S_{l,1}, ..., S_{l,l} \subseteq \Theta$ with $l \in \mathbb{N}$ such that

$$\Theta \subseteq \bigcup_{k=1}^{l} S_{l,k} \quad \text{and} \quad \sup_{\theta, \tilde{\theta} \in S_{l,k}} \|\theta - \tilde{\theta}\| \leq \frac{\varepsilon}{3Cp}. \qquad (2.2.4)$$

By the compactness of the sets $S_{l,k}$ there exist vectors $\theta_{nl,k}$, $\theta_{l,k} \in S_{l,k}$, such that

$$\sup_{\theta \in S_{l,k}} \hat{L}_{f,n}(\theta) = \hat{L}_{f,n}(\theta_{nl,k}) \quad \text{and} \quad \inf_{\theta \in S_{l,k}} L_{f,n_0}(\theta) = L_{f,n_0}(\theta_{l,k}).$$

For a fixed point $\dot\theta_{l,k} \in S_{l,k}$, by the triangle inequality, we get

$$\big|\sup_{\theta\in S_{l,k}} \hat L_{f,n}(\theta) - \inf_{\theta\in S_{l,k}} L_{f,n_0}(\theta)\big| \le \big|\hat L_{f,n}(\theta_{nl,k}) - \hat L_{f,n}(\dot\theta_{l,k})\big| + \big|\hat L_{f,n}(\dot\theta_{l,k}) - L_{f,n_0}(\dot\theta_{l,k})\big| + \big|L_{f,n_0}(\dot\theta_{l,k}) - L_{f,n_0}(\theta_{l,k})\big|. \quad (2.2.5)$$

By a Taylor expansion with an intermediate point $\tilde\theta_{nl,k}$ between $\theta_{nl,k}$ and $\dot\theta_{l,k}$, the first term on the right-hand side in (2.2.5) can be written as

$$\big|\hat L_{f,n}(\theta_{nl,k}) - \hat L_{f,n}(\dot\theta_{l,k})\big| = \Big|\Big(\int \frac{\partial \log f(x|\tilde\theta_{nl,k}, z)}{\partial\theta}\, dQ_n(x,z)\Big)^T \cdot (\theta_{nl,k} - \dot\theta_{l,k})\Big|.$$

By (2.2.3) and (2.2.4), for eventually all $n \in n_0 \cdot \mathbb{N}$, the right term of the last equation is bounded almost surely by

$$p \cdot \sup_{\theta,\tilde\theta\in S_{l,k}} \|\theta - \tilde\theta\| \cdot \int \Big\|\frac{\partial \log f(x|\tilde\theta_{nl,k}, z)}{\partial\theta}\Big\|\, dQ_n(x,z) \le p \cdot \frac{\varepsilon}{3Cp} \cdot C = \frac{\varepsilon}{3}.$$

Analogously, for the third term on the right-hand side in (2.2.5), we have the same result. For the second term on the right-hand side in (2.2.5) by Lemma 2.2.1 with $\psi(x,z) = \log f(x|\dot\theta_{l,k}, z)$ under B3, for n large enough we have

$$\big|\hat L_{f,n}(\dot\theta_{l,k}) - L_{f,n_0}(\dot\theta_{l,k})\big| < \frac{\varepsilon}{3} \quad \text{a.s.}$$

Hence, there exists an $N_{l,k}$, such that for all $n > N_{l,k}$,

$$\big|\sup_{\theta\in S_{l,k}} \hat L_{f,n}(\theta) - \inf_{\theta\in S_{l,k}} L_{f,n_0}(\theta)\big| < \varepsilon \quad \text{a.s.}$$

Analogously, there exists an $N'_{l,k} \in \mathbb{N}$, such that for all $n > N'_{l,k}$,

$$\big|\inf_{\theta\in S_{l,k}} \hat L_{f,n}(\theta) - \sup_{\theta\in S_{l,k}} L_{f,n_0}(\theta)\big| < \varepsilon \quad \text{a.s.}$$

Hence, for all $n > \max_{1\le k\le l}\{N_{l,k}, N'_{l,k}\}$,

$$\sup_{\theta\in\Theta}\big|\hat L_{f,n}(\theta) - L_{f,n_0}(\theta)\big| \le \max_k \sup_{\theta\in S_{l,k}} \big|\hat L_{f,n}(\theta) - L_{f,n_0}(\theta)\big|$$
$$\le \max_k \Big\{\big|\sup_{\theta\in S_{l,k}} \hat L_{f,n}(\theta) - \inf_{\theta\in S_{l,k}} L_{f,n_0}(\theta)\big|, \big|\inf_{\theta\in S_{l,k}} \hat L_{f,n}(\theta) - \sup_{\theta\in S_{l,k}} L_{f,n_0}(\theta)\big|\Big\} < \varepsilon \quad \text{a.s.}$$

□

Lemma 2.2.3. *If* B1–B6 *hold, then*

$$\sqrt{n} \cdot \|\hat\theta_n - \theta_{n_0}\| = O_p(1),$$
$$\sqrt{n} \cdot \big\|\hat\theta_n - \theta_{n_0} + \ddot L^{-1}_{f,n_0}(\theta_{n_0}) \cdot \dot{\hat L}_{f,n}(\theta_{n_0})\big\| = o_p(1).$$

Proof. By a Taylor expansion, there exists a $\tilde{\theta}_n$ lying between θ_{n_0} and $\hat{\theta}_n$ such that

$$\dot{\hat{L}}_{f,n}(\hat{\theta}_n) = \dot{\hat{L}}_{f,n}(\theta_{n_0}) + \ddot{\hat{L}}_{f,n}(\tilde{\theta}_n) \cdot (\hat{\theta}_n - \theta_{n_0}).$$

By the definition of $\hat{\theta}_n$ it follows that

$$\left\| \dot{\hat{L}}_{f,n}(\theta_{n_0}) + \ddot{\hat{L}}_{f,n}(\tilde{\theta}_n) \cdot (\hat{\theta}_n - \theta_{n_0}) \right\| = 0. \tag{2.2.6}$$

Under B5 and by Lemma 2.2.1 with $\psi(x,z) = \partial \log f(x|\theta_{n_0}, z)/\partial\theta_j$, for $j \in \{1, ..., p\}$, we have

$$\sqrt{n} \cdot \Big(\int \frac{\partial \log f(x|\theta_{n_0}, z)}{\partial \theta_j} \, dQ_n(x,z) - \frac{1}{n_0} \sum_{i=1}^{n_0} \int \frac{\partial \log f(x|\theta_{n_0}, z_i)}{\partial \theta_j} \, dH(x|z_i) \Big)$$

converges in distribution to a normal distribution. Further by the definition of θ_{n_0}, for each $j \in \{1, ..., p\}$, under B4 and B5,

$$\frac{1}{n_0} \sum_{i=1}^{n_0} \int \frac{\partial \log f(x|\theta_{n_0}, z_i)}{\partial \theta_j} \, dH(x|z_i) = 0. \tag{2.2.7}$$

Hence,

$$\sqrt{n} \cdot \left\| \dot{\hat{L}}_{f,n}(\theta_{n_0}) \right\| = \sup_{j \in \{1,\ldots,p\}} \sqrt{n} \cdot \left\| \int \frac{\partial \log f(x|\theta_{n_0}, z)}{\partial \theta_j} \, dQ_n(x,z) \right\| = O_p(1). \tag{2.2.8}$$

Under B3 and B5 we can switch the order of integration and differentiation in the Hessian matrix of $L_{f,n_0}(\theta_{n_0})$, i.e.

$$\ddot{L}_{f,n_0}(\theta_{n_0}) = \frac{1}{n_0} \sum_{i=1}^{n_0} \int \frac{\partial^2 \log f(x|\theta_{n_0}, z_i)}{\partial \theta^2} dH(x|z_i).$$

By a Taylor expansion, there exists a $\bar{\theta}_n$ lying between $\tilde{\theta}_n$ and θ_{n_0} such that

$$\begin{aligned}
&\left\| \ddot{\hat{L}}_{f,n}(\tilde{\theta}_n) - \ddot{L}_{f,n_0}(\theta_{n_0}) \right\| \leq \left\| \ddot{\hat{L}}_{f,n}(\tilde{\theta}_n) - \ddot{\hat{L}}_{f,n}(\theta_{n_0}) \right\| + \left\| \ddot{\hat{L}}_{f,n}(\theta_{n_0}) - \ddot{L}_{f,n_0}(\theta_{n_0}) \right\| \\
&\leq p \cdot \max_{i,j,k} \left| \int \frac{\partial^3 \log f(x|\bar{\theta}_n, z)}{\partial \theta_i \partial \theta_j \partial \theta_k} \, dQ_n(x,z) \right| \|\tilde{\theta}_n - \theta_{n_0}\| \\
&\quad + \max_{i,j} \left| \int \frac{\partial^2 \log f(x|\theta_{n_0}, z)}{\partial \theta_i \partial \theta_j} dQ_n(x,z) - \frac{1}{n_0} \sum_{i=1}^{n_0} \int \frac{\partial^2 \log f(x|\theta_{n_0}, z_i)}{\partial \theta_i \partial \theta_j} dH(x|z_i) \right|.
\end{aligned} \tag{2.2.9}$$

Under B1–B6, by Lemma 2.2.2, the first term on the right side of (2.2.9) tends to zero. Further, for each i, $j \in \{1, ..., p\}$ by Lemma 2.2.1 with $\psi(x,z) =$

$\partial^2 \log f(x|\theta_{n_0}, z)/\partial\theta_i\partial\theta_j$ under B5 the second term tends to zero as well. Therefore,

$$\|\hat{\ddot{L}}_{f,n}(\tilde{\theta}_n) - \ddot{L}_{f,n_0}(\theta_{n_0})\| \to 0 \quad \text{a.s.}$$

Since the matrix $\ddot{L}_{f,n_0}(\theta_{n_0})$ is invertible under B5, by the continuity of the determinant function the matrix $\hat{\ddot{L}}_{f,n}(\tilde{\theta}_n)$ is also invertible for eventually all $n \in n_0 \cdot \mathbb{N}$. Further, by the continuity of the inversion operator we obtain

$$\|\hat{\ddot{L}}^{-1}_{f,n}(\tilde{\theta}_n) - \ddot{L}^{-1}_{f,n_0}(\theta_{n_0})\| \to 0 \quad \text{a.s.} \tag{2.2.10}$$

as well. It follows then from (2.2.6), (2.2.8) and (2.2.10) that

$$\begin{aligned}
&\sqrt{n} \cdot \left\|\hat{\theta}_n - \theta_{n_0} + \ddot{L}^{-1}_{f,n_0}(\theta_{n_0}) \cdot \hat{\dot{L}}_{f,n}(\theta_{n_0})\right\| \\
=&\sqrt{n} \cdot \left\|\hat{\theta}_n - \theta_{n_0} + \hat{\ddot{L}}^{-1}_{f,n}(\tilde{\theta}_n) \cdot \hat{\dot{L}}_{f,n}(\theta_{n_0}) - \left(\hat{\ddot{L}}^{-1}_{f,n}(\tilde{\theta}_n) - \ddot{L}^{-1}_{f,n_0}(\theta_{n_0})\right) \cdot \hat{\dot{L}}_{f,n}(\theta_{n_0})\right\| \\
\leq& p \cdot \left\|\hat{\ddot{L}}^{-1}_{f,n}(\tilde{\theta}_n) - \ddot{L}^{-1}_{f,n_0}(\theta_{n_0})\right\| \cdot \left\|\sqrt{n} \cdot \hat{\dot{L}}_{f,n}(\theta_{n_0})\right\| = o_p(1).
\end{aligned}$$

By the boundedness of $\ddot{L}^{-1}_{f,n_0}(\theta_{n_0})$ and (2.2.8) we get

$$\begin{aligned}
\sqrt{n} \cdot \|\hat{\theta}_n - \theta_{n_0}\| \leq \sqrt{n} \cdot \|\hat{\theta}_n - \theta_{n_0} + \ddot{L}^{-1}_{f,n_0}(\theta_{n_0}) \cdot \hat{\dot{L}}_{f,n}(\theta_{n_0})\| \\
+ p \cdot \|\ddot{L}^{-1}_{f,n_0}(\theta_{n_0})\| \cdot \sqrt{n} \cdot \|\hat{\dot{L}}_{f,n}(\theta_{n_0})\| = O_p(1).
\end{aligned}$$

□

In order to state the main theorems we introduce the functions $C_{\mathcal{F}} : \Theta \to \mathbb{R}^p$, $N_{\mathcal{F}} : \mathbb{R} \times \Theta \times [0,1]^d \to \mathbb{R}$ and a constant $\sigma^2_{\mathcal{F}}$ with

$$\begin{aligned}
C_{\mathcal{F}}(\theta) :=& \frac{1}{n_0} \sum_{i=1}^{n_0} \int \big(H(x|z_i) - F(x|\theta, z_i)\big) \cdot \dot{F}(x|\theta, z_i)\, dH(x|z_i), \\
N_{\mathcal{F}}(x, \theta, z) :=& \big(H(x|z) - F(x|\theta, z)\big)^2 + 2\int_x^\infty \big(H(u|z) - F(u|\theta, z)\big)\, dH(u|z) \\
&+ 2C^T_{\mathcal{F}}(\theta_{n_0}) \cdot \ddot{L}^{-1}_{f,n_0}(\theta_{n_0}) \cdot \frac{\partial \log f(x|\theta, z)}{\partial\theta}, \\
\sigma^2_{\mathcal{F}} :=& \frac{1}{n_0} \sum_{i=1}^{n_0} \Big(\int N^2_{\mathcal{F}}(x, \theta_{n_0}, z_i)\, dH(x|z_i) - \Big(\int N_{\mathcal{F}}(x, \theta_{n_0}, z_i)\, dH(x|z_i)\Big)^2\Big).
\end{aligned}$$

Under B5 and B7, the functions $C_{\mathcal{F}}$ and $N_{\mathcal{F}}$ exist, the constant $\sigma^2_{\mathcal{F}} < \infty$. In the next theorem we show the asymptotic normality of the estimated distance.

Theorem 2.2.4. *Let* B1–B7 *be satisfied, then*

$$\sqrt{n} \cdot \left(\hat{d}_{H,n}(\mathcal{F}) - d_H(\mathcal{F})\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\mathcal{F}}).$$

Proof. Note that for $(x,z)\in\mathbb{R}\times[0,1]^d$, $H_n(x|z)-F(x|\hat{\theta}_n,z)$ can be represented as

$$\big(H_n(x|z)-H(x|z)\big)+\big(H(x|z)-F(x|\theta_{n_0},z)\big)-\big(F(x|\hat{\theta}_n,z)-F(x|\theta_{n_0},z)\big).$$

Hence, we can write

$$\begin{aligned}\sqrt{n}\cdot\hat{d}_{H,n}(\mathcal{F})=&\sqrt{n}\cdot\int\big(H_n(x|z)-H(x|z)\big)^2\,dQ_n(x,z)\\&+2\sqrt{n}\cdot\int\big(H_n(x|z)-H(x|z)\big)\cdot\big(H(x|z)-F(x|\theta_{n_0},z)\big)\,dQ_n(x,z)\\&-2\sqrt{n}\cdot\int\big(H(x|z)-F(x|\theta_{n_0},z)\big)\cdot\big(F(x|\hat{\theta}_n,z)-F(x|\theta_{n_0},z)\big)\,dQ_n(x,z)\\&+\sqrt{n}\cdot\int\big(H(x|z)-F(x|\theta_{n_0},z)\big)^2dQ_n(x,z)\\&-2\sqrt{n}\cdot\int\big(H_n(x|z)-H(x|z)\big)\cdot\big(F(x|\hat{\theta}_n,z)-F(x|\theta_{n_0},z)\big)\,dQ_n(x,z)\\&+\sqrt{n}\cdot\int\big(F(x|\hat{\theta}_n,z)-F(x|\theta_{n_0},z)\big)^2\,dQ_n(x,z)=:\sum_{k=1}^{6}T_{kn}.\end{aligned}$$

For T_{1n} we have

$$\begin{aligned}T_{1n}=&\sqrt{n}\cdot\int\big(H_n^2(x|z)-2H(x|z)H_n(x|z)+H^2(x|z)\big)\,dQ_n(x,z)\\=&\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int H_n^2(x|z_i)dH_n(x|z_i)-2\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int H(x|z_i)H_n(x|z_i)dH_n(x|z_i)\\&+\sqrt{n}\cdot\int H^2(x|z)dQ_n(x,z)\\=&\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int\int\int I(u\le x)I(t\le x)dH_n(u|z_i)dH_n(t|z_i)dH_n(x|z_i)\\&-2\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int\int I(u\le x)H(x|z_i)dH_n(u|z_i)dH_n(x|z_i)\\&+\sqrt{n}\cdot\int H^2(x|z)dQ_n(x,z).\end{aligned}\tag{2.2.11}$$

By Lemma A.1.1 with $k=3$, $X_{ij}=X_j$ for $i\in\{1,2,3\}$, $j\in\{1,...,n\}$ and

$$\psi(u,t,x,z)=I(u\le x)I(t\le x),$$

since

$$E[\psi^2(X_{i_1},X_{i_2},X_{i_3},z)]=E[I(X_{i_1}\le X_{i_3})I(X_{i_2}\le X_{i_3})]\le 1$$

for any $i_1, i_2, i_3 \in \{1, ..., n\}$, thus the first term on the right-hand side of (2.2.11) can be written as

$$\begin{aligned}
&\sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int\int\int I(u \leq x) I(t \leq x) dH_n(u|z_i) dH(t|z_i) dH(x|z_i) \\
&+ \sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int\int\int I(u \leq x) I(t \leq x) dH(u|z_i) dH_n(t|z_i) dH(x|z_i) \\
&+ \sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int\int\int I(u \leq x) I(t \leq x) dH(u|z_i) dH(t|z_i) dH_n(x|z_i) \\
&- 2\sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int\int\int I(u \leq x) I(t \leq x) dH(u|z_i) dH(t|z_i) dH(x|z_i) + o_p(1) \\
=&2\sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int H(x|z_i) H_n(x|z_i) dH(x|z_i) + \sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int H^2(x|z_i) dH_n(x|z_i) \\
&- 2\sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int H^2(x|z_i) dH(x|z_i) + o_p(1).
\end{aligned}$$

Analogously, by Lemma A.1.1 with $k = 2$, $X_{ij} = X_j$ for $i \in \{1, 2\}$, $j \in \{1, ..., n\}$ and

$$\psi(u, x, z) = I(u \leq x) H(x|z),$$

since

$$E[\psi^2(X_{i_1}, X_{i_2}, z)] = E[I(X_{i_1} \leq X_{i_2}) H^2(X_{i_2}|z)] \leq 1$$

for any $i_1, i_2 \in \{1, ..., n\}$, hence the second term on the right-hand side of (2.2.11) can be written as

$$\begin{aligned}
&- 2\sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int\int I(u \leq x) H(x|z_i) dH_n(u|z_i) dH(x|z_i) \\
&- 2\sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int\int I(u \leq x) H(x|z_i) dH(u|z_i) dH_n(x|z_i) \\
&+ 2\sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int\int I(u \leq x) H(x|z_i) dH(u|z_i) dH(x|z_i) + o_p(1) \\
=& - 2\sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int H(x|z_i) H_n(x|z_i) dH(x|z_i) - 2\sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int H^2(x|z_i) dH_n(x|z_i) \\
&+ 2\sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int H^2(x|z_i) dH(x|z_i) + o_p(1).
\end{aligned}$$

Hence, we get

$$
\begin{aligned}
T_{1n} =& 2\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int H(x|z_i)H_n(x|z_i)dH(x|z_i)\\
&+\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int H^2(x|z_i)dH_n(x|z_i)-2\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int H^2(x|z_i)dH(x|z_i)\\
&-2\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int H(x|z_i)H_n(x|z_i)dH(x|z_i)-2\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int H^2(x|z_i)dH_n(x|z_i)\\
&+2\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int H^2(x|z_i)dH(x|z_i)\\
&+\sqrt{n}\cdot\int H^2(x|z)dQ_n(x,z)+o_p(1)=o_p(1).
\end{aligned}
\tag{2.2.12}
$$

With the same arguments, it can be shown that

$$
\begin{aligned}
T_{2n} =& 2\sqrt{n}\cdot\int\int_x^{\infty}\big(H(u|z)-F(u|\theta_{n_0},z)\big)\,dH(u|z)\,dQ_n(x,z)\\
&-2\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int\int_x^{\infty}\big(H(u|z_i)-F(u|\theta_{n_0},z_i)\big)\,dH(u|z_i)\,dH(x|z_i)+o_p(1).
\end{aligned}
$$

By a Taylor expansion, there exists a $\tilde{\theta}_n$ between $\hat{\theta}_n$ and θ_{n_0}, such that

$$
\begin{aligned}
T_{3n} =& -2\sqrt{n}\cdot\Big(\int\big(H(x|z)-F(x|\theta_{n_0},z)\big)\cdot\dot{F}(x|\theta_{n_0},z)\,dQ_n(x,z)\Big)^T\cdot(\hat{\theta}_n-\theta_{n_0})\\
&-2\sqrt{n}\cdot(\hat{\theta}_n-\theta_{n_0})^T\cdot\int\big(H(x|z)-F(x|\theta_{n_0},z)\big)\cdot\ddot{F}(x|\tilde{\theta}_n,z)\,dQ_n(x,z)\cdot(\hat{\theta}_n-\theta_{n_0}).
\end{aligned}
$$

For the first term on the right-hand side of the last equation, under B7, each component of the vector $\big(H(x|z)-F(x|\theta_{n_0},z)\big)\cdot\dot{F}(x|\theta_{n_0},z)$ is H integrable, hence, by Lemma 2.2.1 we get

$$\Big\|\int\big(H(x|z)-F(x|\theta_{n_0},z)\big)\cdot\dot{F}(x|\theta_{n_0},z)\,dQ_n(x,z)-C_{\mathcal{F}}(\theta_{n_0})\Big\|=o_p(1).$$

For the second term note that under B7,

$$\Big\|\int\big(H(x|z)-F(x|\theta_{n_0},z)\big)\cdot\ddot{F}(x|\tilde{\theta}_n,z)\,dQ_n(x,z)\Big\|$$

is stochastically bounded. Hence, by Lemma 2.2.3 under B1–B6, the second term on the right-hand side is equal to $o_p(1)$ and

$$T_{3n}=-2\sqrt{n}\cdot C_{\mathcal{F}}^T(\theta_{n_0})\cdot(\hat{\theta}_n-\theta_{n_0})+o_p(1)=2\sqrt{n}\cdot C_{\mathcal{F}}^T(\theta_{n_0})\cdot\ddot{L}_{f,n_0}^{-1}(\theta_{n_0})\cdot\dot{\hat{L}}_{f,n}(\theta_{n_0})+o_p(1).$$

Analogously, we can show that

$$T_{5n} = 2\sqrt{n} \cdot \Big(\int \big(H_n(x|z) - H(x|z)\big) \cdot \dot{F}(x|\theta_{n_0}, z)\, dQ_n(x,z)\Big)^T \cdot (\hat{\theta}_n - \theta_{n_0}) + o_p(1).$$

By Cauchy-Schwarz's inequality, we get

$$\begin{aligned}
&\Big\| \int \big(H_n(x|z) - H(x|z)\big) \cdot \dot{F}(x|\theta_{n_0}, z)\, dQ_n(x,z)\Big\| \\
&\le \int \big|H_n(x|z) - H(x|z)\big| \cdot \|\dot{F}(x|\theta_{n_0}, z)\|\, dQ_n(x,z) \\
&\le \Big(\int \big(H_n(x|z) - H(x|z)\big)^2 dQ_n(x,z)\Big)^{1/2} \cdot \Big(\int \|\dot{F}(x|\theta_{n_0}, z)\|^2\, dQ_n(x,z)\Big)^{1/2} \\
&\le \Big(T_{1n} \cdot n^{-1/2}\Big)^{1/2} \cdot \Big(\int \|\dot{F}(x|\theta_{n_0}, z)\|^2\, dQ_n(x,z)\Big)^{1/2} = o_p(1)
\end{aligned}$$

where the last step follows from (2.2.12) and B7. Hence, it follows from Lemma 2.2.3 that $T_{5n} = o_p(1)$.

By the same arguments, there exists a $\bar{\theta}_n$ between $\hat{\theta}_n$ and θ_{n_0}, such that

$$|T_{6n}| \le \sqrt{n} \cdot p^2 \int \|\dot{F}(x|\bar{\theta}_n, z)\|^2 dQ_n(x,z) \cdot \|\hat{\theta}_n - \theta_{n_0}\|^2 = o_p(1).$$

Therefore, by the definition of the function $N_{\mathcal{F}}$,

$$\begin{aligned}
\sqrt{n} \cdot \hat{d}_{H,n}(\mathcal{F}) = &\sqrt{n} \cdot \int N_{\mathcal{F}}(x, \theta_{n_0}, z)\, dQ_n(x,z) \\
&- 2\sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int \int_x^\infty \big(H(u|z_i) - F(u|\theta_{n_0}, z_i)\big)\, dH(u|z_i)\, dH(x|z_i) + o_p(1).
\end{aligned}$$

Note that by the definition of θ_{n_0}, under B3 and B5,

$$\begin{aligned}
&\frac{1}{n_0} \sum_{i=1}^{n_0} \int C_{\mathcal{F}}^T(\theta_{n_0}) \cdot \ddot{L}_{f,n_0}^{-1}(\theta_{n_0}) \cdot \frac{\partial \log f(x|\theta_{n_0}, z_i)}{\partial \theta} dH(x|z_i) \\
=&C_{\mathcal{F}}^T(\theta_{n_0}) \cdot \ddot{L}_{f,n_0}^{-1}(\theta_{n_0}) \cdot \dot{L}_{f,n_0}(\theta_{n_0}) = 0.
\end{aligned}$$

Thus, we can write

$$\begin{aligned}
d_H(\mathcal{F}) =&\frac{1}{n_0} \sum_{i=1}^{n_0} \int N_{\mathcal{F}}(x, \theta_{n_0}, z_i)\, dH(x|z_i) \\
&- 2 \cdot \frac{1}{n_0} \sum_{i=1}^{n_0} \int \int_x^\infty \big(H(u|z_i) - F(u|\theta_{n_0}, z_i)\big)\, dH(u|z_i)\, dH(x|z_i).
\end{aligned}$$

Therefore, we have

$$\sqrt{n} \cdot \big(\hat{d}_{H,n}(\mathcal{F}) - d_H(\mathcal{F})\big)$$

$$=\sqrt{n}\cdot\Big(\int N_{\mathcal{F}}(x,\theta_{n_0},z)\,dQ_n(x,z)-\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F}}(x,\theta_{n_0},z_i)\,dH(x|z_i)\Big)+o_p(1). \tag{2.2.13}$$

Under B5 and B7, the function $N_{\mathcal{F}}$ is H square integrable. Hence, the assertion follows from Lemma 2.2.1 with $\psi(x,z)=N_{\mathcal{F}}(x,\theta_{n_0},z)$. □

For the estimation of the asymptotic variance $\sigma^2_{\mathcal{F}}$ we define the functions $C_{\mathcal{F},n}:\Theta\to\mathbb{R}^p$, and $N_{\mathcal{F},n}:\mathbb{R}\times\Theta\times[0,1]^d\to\mathbb{R}$ with

$$\begin{aligned}C_{\mathcal{F},n}(\theta):=&\int\big(H_n(x|z)-F(x|\theta,z)\big)\dot F(x|\theta,z)dQ_n(x,z),\\ N_{\mathcal{F},n}(x,\theta,z):=&\big(H_n(x|z)-F(x|\theta,z)\big)^2+2\int_x^\infty\big(H_n(u|z)-F(u|\theta,z)\big)dH_n(u|z)\\ &+2C^T_{\mathcal{F},n}(\hat\theta_n)\cdot\hat{\ddot L}^{-1}_{f,n}(\hat\theta_n)\cdot\frac{\partial\log f(x|\theta,z)}{\partial\theta}.\end{aligned}$$

In the next lemma, we show that $\sigma^2_{\mathcal{F}}$ can be estimated consistently by

$$\hat\sigma^2_{\mathcal{F},n}:=\int N^2_{\mathcal{F},n}(x,\hat\theta_n,z)dQ_n(x,z)-\frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int N_{\mathcal{F},n}(x,\hat\theta_n,z_i)dH_n(x|z_i)\Big)^2.$$

Lemma 2.2.5. *If* B1–B7 *hold, then we have*

$$\hat\sigma^2_{\mathcal{F},n}=\sigma^2_{\mathcal{F}}+o_p(1).$$

Proof. First, we show that for each $\theta\in\Theta$,

$$\int N^2_{\mathcal{F},n}(x,\theta,z)dQ_n(x,z)-\frac{1}{n_0}\sum_{i=1}^{n_0}\int N^2_{\mathcal{F}}(x,\theta,z_i)dH(x|z_i)=o_p(1). \tag{2.2.14}$$

Note that $\int\big(H_n(x|z)-F(x|\theta,z)\big)^4dQ_n(x,z)$ is a part of

$$\int N^2_{\mathcal{F},n}(x,\theta,z)dQ_n(x,z)$$

and its counterpart in

$$\frac{1}{n_0}\sum_{i=1}^{n_0}\int N^2_{\mathcal{F}}(x,\theta,z_i)dH(x|z_i)$$

is

$$\frac{1}{n_0}\sum_{i=1}^{n_0}\int\big(H(x|z_i)-F(x|\theta,z_i)\big)^4dH(x|z_i).$$

In the sequel, we show that

$$\begin{aligned}&\int \big(H_n(x|z)-F(x|\theta,z)\big)^4 dQ_n(x,z)\\ =&\frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(H(x|z_i)-F(x|\theta,z_i)\big)^4 dH(x|z_i)+o_p(1).\end{aligned} \tag{2.2.15}$$

Note that

$$\begin{aligned}&\int \big(H_n(x|z)-F(x|\theta,z)\big)^4 dQ_n(x,z)\\ =&\frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(H_n^4(x|z_i)-4H_n^3(x|z_i)F(x|\theta,z_i)+6H_n^2(x|z_i)F^2(x|\theta,z_i)\\ &-4H_n(x|z_i)F^3(x|\theta,z_i)+F^4(x|\theta,z_i)\big)dH_n(x|z_i).\end{aligned} \tag{2.2.16}$$

By Corollary A.1.2, with $k=5$ and $X_{ij}=X_j$ for $i\in\{1,...,5\}, j\in\{1,...,n\}$ and

$$\psi(x_1,...,x_5,z)=\prod_{j=1}^{4} I(x_j\le x_5),$$

since

$$E[\psi^2(X_{i_1},...,X_{i_5},z)]=E\Big[\prod_{j=1}^{4} I(X_{i_j}\le X_{i_5})\Big]\le 1,$$

we get

$$\begin{aligned}&\frac{1}{n_0}\sum_{i=1}^{n_0}\int H_n^4(x|z_i)dH_n(x|z_i)\\ =&\frac{1}{n_0}\sum_{i=1}^{n_0}\int\cdots\int\prod_{j=1}^{4} I(x_j\le x_5)dH_n(x_1|z_i)\cdots dH_n(x_5|z_i)\\ =&\frac{1}{n_0}\sum_{i=1}^{n_0}\int\cdots\int\prod_{j=1}^{4} I(x_j\le x_5)dH(x_1|z_i)\cdots dH(x_5|z_i)+o_p(1)\\ =&\frac{1}{n_0}\sum_{i=1}^{n_0}\int H^4(x|z_i)dH(x|z_i)+o_p(1).\end{aligned}$$

With similar arguments, we can show that similar results hold for the other terms on the right-hand side of (2.2.16). Therefore, (2.2.15) holds.

Analogously, it can be shown that

$$\int\int_x^\infty \big(H_n(x|z)-F(x|\theta,z)\big)^2\big(H_n(u|z)-F(u|\theta,z)\big)dH_n(u|z)dQ_n(x,z)$$

$$=\frac{1}{n_0}\sum_{i=1}^{n_0}\int\int_x^\infty \big(H(x|z_i)-F(x|\theta,z_i)\big)^2\big(H(u|z_i)-F(u|\theta,z_i)\big)dH(u|z_i)dH(x|z_i)+o_p(1) \tag{2.2.17}$$

and

$$\begin{aligned}&\int\Big(\int_x^\infty \big(H_n(u|z)-F(u|\theta,z)\big)dH_n(u|z)\Big)^2 dQ_n(x,z)\\=&\frac{1}{n_0}\sum_{i=1}^{n_0}\int\Big(\int_x^\infty \big(H(u|z_i)-F(u|\theta,z_i)\big)dH(u|z_i)\Big)^2 dH(x|z_i)+o_p(1).\end{aligned} \tag{2.2.18}$$

For the rest terms of

$$\int N^2_{\mathcal{F},n}(x,\theta,z)dQ_n(x,z) \quad\text{and}\quad \frac{1}{n_0}\sum_{i=1}^{n_0}\int N^2_{\mathcal{F}}(x,\theta,z_i)dH(x|z_i),$$

note that

$$\|C^T_{\mathcal{F},n}(\hat\theta_n)-C^T_{\mathcal{F}}(\theta_{n_0})\|\le\|C^T_{\mathcal{F},n}(\hat\theta_n)-C^T_{\mathcal{F},n}(\theta_{n_0})\|+\|C^T_{\mathcal{F},n}(\theta_{n_0})-C^T_{\mathcal{F}}(\theta_{n_0})\|.$$

Under B7, the derivative of $C^T_{\mathcal{F},n}$ is stochastically bounded on Θ. Hence, by Lemma 2.2.2 under B1-B5

$$\|C^T_{\mathcal{F},n}(\hat\theta_n)-C^T_{\mathcal{F},n}(\theta_{n_0})\|=o_p(1).$$

Further, it follows from Corollary A.1.2 and Lemma 2.2.1 that under B7

$$\|C^T_{\mathcal{F},n}(\theta_{n_0})-C^T_{\mathcal{F}}(\theta_{n_0})\|=o_p(1).$$

Therefore,

$$\|C^T_{\mathcal{F},n}(\hat\theta_n)-C^T_{\mathcal{F}}(\theta_{n_0})\|=o_p(1).$$

Further, analogously to (2.2.10) under B6, it can be shown that

$$\|\ddot{\hat{L}}^{-1}_{f,n}(\hat\theta_n)-\ddot{L}^{-1}_{f,n_0}(\theta_{n_0})\|=o_p(1).$$

Thus, under B5 Corollary A.1.2 implies

$$\begin{aligned}&\int\big(H_n(x|z)-F(x|\theta,z)\big)^2\cdot C^T_{\mathcal{F},n}(\hat\theta_n)\cdot\ddot{\hat{L}}^{-1}_{f,n}(\hat\theta_n)\cdot\frac{\partial\log f(x|\theta,z)}{\partial\theta}dQ_n(x,z)\\=&C^T_{\mathcal{F},n}(\hat\theta_n)\cdot\ddot{L}^{-1}_{f,n}(\hat\theta_n)\cdot\int\big(H_n(x|z)-F(x|\theta,z)\big)^2\cdot\frac{\partial\log f(x|\theta,z)}{\partial\theta}dQ_n(x,z)\end{aligned}$$

$$=\frac{1}{n_0}\sum_{i=1}^{n_0} C_{\mathcal{F}}^T(\theta_{n_0})\cdot \ddot{L}_{f,n_0}^{-1}(\theta_{n_0})\cdot \int \big(H(x|z_i)-F(x|\theta,z_i)\big)^2\cdot \frac{\partial \log f(x|\theta,z_i)}{\partial\theta}dH(x|z_i)+o_p(1)$$

(2.2.19)

With the same arguments, it follows further

$$\int\int_x^\infty \big(H_n(u|z)-F(u|\theta,z)\big)dH_n(u|z)\cdot C_{\mathcal{F},n}^T(\hat\theta_n)\cdot \ddot{\hat{L}}_{f,n}^{-1}(\hat\theta_n)\cdot \frac{\partial \log f(x|\theta,z)}{\partial\theta}dQ_n(x,z)$$
$$=\frac{1}{n_0}\sum_{i=1}^{n_0} C_{\mathcal{F}}^T(\theta_{n_0})\cdot \ddot{L}_{f,n_0}^{-1}(\theta_{n_0})$$
$$\times\int\int_x^\infty \big(H(u|z_i)-F(u|\theta,z_i)\big)dH(u|z_i)\cdot \frac{\partial \log f(x|\theta,z_i)}{\partial\theta}dH(x|z_i)+o_p(1)$$

(2.2.20)

and

$$\int \Big(C_{\mathcal{F},n}^T(\hat\theta_n)\cdot \ddot{\hat{L}}_{f,n}^{-1}(\hat\theta_n)\cdot \frac{\partial \log f(x|\theta,z)}{\partial\theta}\Big)^2 dQ_n(x,z)$$
$$=\frac{1}{n_0}\sum_{i=1}^{n_0}\int \Big(C_{\mathcal{F}}^T(\theta_{n_0})\cdot \ddot{L}_{f,n_0}^{-1}(\theta_{n_0})\cdot \frac{\partial \log f(x|\theta,z_i)}{\partial\theta}\Big)^2 dH(x|z_i)+o_p(1). \quad (2.2.21)$$

By (2.2.15)–(2.2.21), for any $\theta\in\Theta$, (2.2.14) holds. Hence,

$$\int N_{\mathcal{F},n}^2(x,\hat\theta_n,z)dQ_n(x,z)-\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F}}^2(x,\theta_{n_0},z_i)dH(x|z_i)$$
$$=\Big(\int N_{\mathcal{F},n}^2(x,\hat\theta_n,z)dQ_n(x,z)-\int N_{\mathcal{F},n}^2(x,\theta_{n_0},z)dQ_n(x,z)\Big)$$
$$+\Big(\int N_{\mathcal{F},n}^2(x,\theta_{n_0},z)dQ_n(x,z)-\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F}}^2(x,\theta_{n_0},z_i)dH(x|z_i)\Big)$$
$$=\int N_{\mathcal{F},n}^2(x,\hat\theta_n,z)dQ_n(x,z)-\int N_{\mathcal{F},n}^2(x,\theta_{n_0},z)dQ_n(x,z)+o_p(1).$$

By definition of the function $N_{\mathcal{F},n}$, under B7 there exists a constant $C>0$ such that

$$\big\|N_{\mathcal{F},n}(x,\theta,z)\cdot \dot N_{\mathcal{F},n}(x,\theta,z)\big\|\le C\Big(1+\Big\|\frac{\partial \log f(x|\theta,z)}{\partial\theta}\Big\|\Big)\Big(1+\Big\|\frac{\partial^2 \log f(x|\theta,z)}{\partial\theta^2}\Big\|\Big).$$

Thus, under B5

$$\int N_{\mathcal{F},n}^2(x,\cdot,z)dQ_n(x,z)$$

has a stochastically bounded derivative on Θ. Therefore, by Lemma 2.2.2,

$$\int N_{\mathcal{F},n}^2(x,\hat\theta_n,z)dQ_n(x,z)-\int N_{\mathcal{F},n}^2(x,\theta_{n_0},z)dQ_n(x,z)=o_p(1).$$

Consequently,

$$\int N^2_{\mathcal{F},n}(x,\hat{\theta}_n,z)dQ_n(x,z)-\frac{1}{n_0}\sum_{i=1}^{n_0}\int N^2_{\mathcal{F}}(x,\theta_{n_0},z_i)dH(x|z_i)=o_p(1). \quad (2.2.22)$$

Analogously, we can show

$$\begin{aligned}
&\frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int N_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)dH_n(x|z_i)\Big)^2\\
=&\frac{1}{n_0}\sum_{i=1}^{n_0}\int\int N_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)N_{\mathcal{F},n}(u,\hat{\theta}_n,z_i)dH_n(u|z_i)dH_n(x|z_i)\\
=&\frac{1}{n_0}\sum_{i=1}^{n_0}\int\int N_{\mathcal{F}}(x,\theta_{n_0},z_i)N_{\mathcal{F}}(u,\theta_{n_0},z_i)dH(u|z_i)dH(x|z_i)+o_p(1)\\
=&\frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int N_{\mathcal{F}}(x,\theta_{n_0},z_i)dH(x|z_i)\Big)^2+o_p(1). \qquad (2.2.23)
\end{aligned}$$

The assertion follows then from (2.2.22) and (2.2.23). □

For the Model $\mathcal{G}$, let $C_{\mathcal{G}}$, $N_{\mathcal{G}}$, $\sigma^2_{\mathcal{G}}$ and their estimates be defined accordingly. Further we denote the constants σ^2 and $\hat{\sigma}^2_n$ by

$$\begin{aligned}
\sigma^2 :=& \frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(N_{\mathcal{F}}(x,\theta_{n_0},z_i)-N_{\mathcal{G}}(x,\gamma_{n_0},z_i)\big)^2 dH(x|z_i)\\
&-\frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int \big(N_{\mathcal{F}}(x,\theta_{n_0},z_i)-N_{\mathcal{G}}(x,\gamma_{n_0},z_i)\big)dH(x|z_i)\Big)^2,\\
\hat{\sigma}^2_n :=& \int \big(N_{\mathcal{F},n}(x,\hat{\theta}_n,z)-N_{\mathcal{G},n}(x,\hat{\gamma}_n,z)\big)^2 dQ_n(x,z)\\
&-\frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int \big(N_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)-N_{\mathcal{G},n}(x,\hat{\gamma}_n,z_i)\big)dH_n(x|z_i)\Big)^2.
\end{aligned}$$

Next we show the asymptotic normality of test statistic T_n and the consistency of $\hat{\sigma}^2_n$ to σ^2.

Theorem 2.2.6. *If* B1–B7 *hold then*

$$\sqrt{n}\cdot\Big(T_n-\big(d_H(\mathcal{F})-d_H(\mathcal{G})\big)\Big)\xrightarrow{d}\mathcal{N}(0,\sigma^2) \quad \text{and} \quad \hat{\sigma}^2_n\to\sigma^2.$$

Proof. Analogously to (2.2.13), we can show

$$\sqrt{n}\cdot\Big(T_n-\big(d_H(\mathcal{F})-d_H(\mathcal{G})\big)\Big)$$

$$
\begin{aligned}
=&\sqrt{n}\int \left(N_{\mathcal{F}}(x,\theta_{n_0},z)-N_{\mathcal{G}}(x,\gamma_{n_0},z)\right)dQ_n(x,z)\\
&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int \left(N_{\mathcal{F}}(x,\theta_{n_0},z_i)-N_{\mathcal{G}}(x,\gamma_{n_0},z_i)\right)dH(x|z_i)+o_p(1).
\end{aligned}
$$

Thus, the first part of the assertion follows from Lemma 2.2.1 with $\psi(x,z) = N_{\mathcal{F}}(x,\theta_{n_0},z) - N_{\mathcal{G}}(x,\gamma_{n_0},z)$. For the second part of the assertion, note that

$$
\begin{aligned}
\hat{\sigma}_n^2 =&\hat{\sigma}_{\mathcal{F},n}^2+\hat{\sigma}_{\mathcal{G},n}^2-2\int N_{\mathcal{F},n}(x,\hat{\theta}_n,z)\cdot N_{\mathcal{G},n}(x,\hat{\gamma}_n,z)dQ_n(x,z)\\
&+\frac{2}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)dH_n(x|z_i)\cdot\int N_{\mathcal{G},n}(x,\hat{\gamma}_n,z_i)dH_n(x|z_i)
\end{aligned}
$$

and

$$
\begin{aligned}
\sigma^2 =&\sigma_{\mathcal{F}}^2+\sigma_{\mathcal{G}}^2-\frac{2}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F}}(x,\theta_{n_0},z_i)\cdot N_{\mathcal{G}}(x,\gamma_{n_0},z_i)dH(x|z_i)\\
&+\frac{2}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F}}(x,\theta_{n_0},z_i)dH(x|z_i)\cdot\int N_{\mathcal{G}}(x,\gamma_{n_0},z_i)dH(x|z_i).
\end{aligned}
$$

Analogously to (2.2.22), we can show

$$
\begin{aligned}
&\int N_{\mathcal{F},n}(x,\hat{\theta}_n,z)\cdot N_{\mathcal{G},n}(x,\hat{\gamma}_n,z)dQ_n(x,z)\\
=&\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F}}(x,\theta_{n_0},z_i)\cdot N_{\mathcal{G}}(x,\gamma_{n_0},z_i)dH(x|z_i)=o_p(1)
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)dH_n(x|z_i)\cdot\int N_{\mathcal{G},n}(x,\hat{\gamma}_n,z_i)dH_n(x|z_i)\\
=&\frac{1}{n_0}\sum_{i=1}^{n_0}\int\int N_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)\cdot N_{\mathcal{G},n}(u,\hat{\gamma}_n,z_i)dH_n(u|z_i)dH_n(x|z_i)\\
=&\frac{1}{n_0}\sum_{i=1}^{n_0}\int\int N_{\mathcal{F}}(x,\theta_{n_0},z_i)\cdot N_{\mathcal{G}}(u,\gamma_{n_0},z_i)dH(u|z_i)dH(x|z_i)+o_p(1)\\
=&\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F}}(x,\theta_{n_0},z_i)dH(x|z_i)\cdot\int N_{\mathcal{G}}(x,\gamma_{n_0},z_i)dH(x|z_i)+o_p(1).
\end{aligned}
$$

Hence, the second part of the assertion follows from Lemma 2.2.5. □

Now we can formulate the asymptotic behavior of the test statistic under the hypotheses as in the following theorem.

Theorem 2.2.7. *Let* B1-B7 *be satisfied.*

(1) If $\mathcal{H}^{\mathcal{F}}$ holds, then $\sqrt{n} \cdot T_n$ tends to $-\infty$ in probability.

(2) If $\mathcal{H}^{\mathcal{G}}$ holds, then $\sqrt{n} \cdot T_n$ tends to $+\infty$ in probability.

(3) If $\mathcal{H}^0$ holds, then $\sqrt{n} \cdot T_n \xrightarrow{d} \mathcal{N}(0, \sigma^2)$.

Proof. The assertions follow directly from Theorem 2.2.6. □

By Theorem 2.2.6 and Theorem 2.2.7, if $\sigma^2 > 0$ and $\mathcal{H}^0$ hold true, then

$$\frac{\sqrt{n} \cdot T_n}{\hat{\sigma}_n} \xrightarrow{d} \mathcal{N}(0, 1).$$

The decision rules of our test are given as follows: for a given significance level α we will decide for the hypothesis $\mathcal{H}^0$, if $|\sqrt{n} \cdot T_n/\hat{\sigma}_n| \leq z_{1-\alpha/2}$, where z_α denotes the α-quantile of a standard normal distribution. In the case of $\sqrt{n} \cdot T_n/\hat{\sigma}_n < -z_{1-\alpha/2}$ we reject $\mathcal{H}^0$ in favor of $\mathcal{H}^{\mathcal{F}}$. If $\sqrt{n} \cdot T_n/\hat{\sigma}_n > z_{1-\alpha/2}$, we reject $\mathcal{H}^0$ in favour of $\mathcal{H}^{\mathcal{G}}$. However, we propose to use the model with less parameters, ever if $\mathcal{H}^0$ is not rejected.

A non-degenerate test, which works in the case $\sigma^2 = 0$ as well, can also be constructed based on our theorems by using similar arguments as in Shi (2015b). However, it would go beyond the scope of this thesis.

In the following, we show that $\sigma^2 > 0$ for a concrete example. Without loss of generality, we assume $d = 1$. Let $\mathcal{F}$ be Weibull and $\mathcal{G}$ Log Normal distribution class with parameters depending linearly on z, i.e.

$$F(x|\theta, z) = 1 - \exp\big[-(\frac{x}{a(z)})^{b(z)}\big],$$
$$G(x|\gamma, z) = \frac{1}{\sqrt{2\pi}\sigma(z)} \int_0^x \frac{1}{t} \exp\big[-\frac{1}{2}(\frac{\ln t - \mu(z)}{\sigma(z)})^2\big] dt,$$

where

$$\theta = (a_0, a_1, b_0, b_1) \in \mathbb{R}^4, \gamma = (c_0, c_1, d_0, d_1) \in \mathbb{R}^4,$$

$$a(z) = a_0 + a_1 z, b(z) = b_0 + b_1 z, \mu(z) = c_0 + c_1 z, \sigma(z) = d_0 + d_1 z,$$

for $(x, z) \in \mathbb{R}^+ \times [0, 1]$. Further, we assume that the function $H(\cdot|z)$ has a density function for each $z \in [0, 1]^d$.

Note that by Jensen's inequality, for each $z \in \{z_1, ..., z_{n_0}\}$,

$$\int \big(N_{\mathcal{F}}(x, \theta_{n_0}, z) - N_{\mathcal{G}}(x, \gamma_{n_0}, z)\big)^2 dH(x|z)$$

$$- \Big(\int \big(N_{\mathcal{F}}(x,\theta_{n_0},z) - N_{\mathcal{G}}(x,\gamma_{n_0},z)\big)\, dH(x|z) \Big)^2 \geq 0,$$

thus, suppose $\sigma^2 = 0$, it muss hold that

$$\int \big(N_{\mathcal{F}}(x,\theta_{n_0},z) - N_{\mathcal{G}}(x,\gamma_{n_0},z)\big)^2 dH(x|z) - \Big(\int \big(N_{\mathcal{F}}(x,\theta_{n_0},z) - N_{\mathcal{G}}(x,\gamma_{n_0},z)\big)\, dH(x|z) \Big)^2 = 0.$$

Consequently, there exists a constant $k \in \mathbb{R}$, such that for all $(x,z) \in \mathbb{R}^+ \times \{z_1,...,z_{n_0}\}$,

$$N_{\mathcal{F}}(x,\theta_{n_0},z) - N_{\mathcal{G}}(x,\gamma_{n_0},z) = k. \quad H\text{-a.s.} \tag{2.2.24}$$

Denote the vectors $v = (v_1,v_2,v_3,v_4)^T, v' = (v'_1,v'_2,v'_3,v'_4)^T \in \mathbb{R}^4$ with

$$v^T := 2C^T_{\mathcal{F},n}(\hat{\theta}_n) \cdot \ddot{\hat{L}}^{-1}_{f,n}(\hat{\theta}_n), \quad v'^T := 2C^T_{\mathcal{G},n}(\hat{\gamma}_n) \cdot \ddot{L}^{-1}_{g,n}(\hat{\gamma}_n)$$

and the function $\omega : (x,z) \in \mathbb{R}^+ \times [0,1] \longrightarrow \mathbb{R}$ with

$$\omega(x,z) := \big(H(x|z) - F(x|\theta_{n_0},z)\big)^2 - \big(H(x|z) - G(x|\gamma_{n_0},z)\big)^2 + 2\int_x^\infty \big(G(u|\gamma_{n_0},z) - F(u|\theta_{n_0},z)\big)\, dH(u|z).$$

Then Equality (2.2.24) implies

$$\frac{\partial \log f(x|\theta_{n_0},z)}{\partial a_0} \cdot (v_1 + v_2 \cdot z) + \frac{\partial \log f(x|\theta_{n_0},z)}{\partial b_0} \cdot (v_3 + v_4 \cdot z) - \frac{\partial \log g(x|\gamma_{n_0},z)}{\partial c_0} \cdot (v'_1 + v'_2 \cdot z) - \frac{\partial \log g(x|\gamma_{n_0},z)}{\partial d_0} \cdot (v'_3 + v'_4 \cdot z) = -\omega(x,z) + k, \tag{2.2.25}$$

where

$$\frac{\partial \log f(x|\theta,z)}{\partial a_0} = -\frac{a(z)}{b(z)} + \frac{a(z)}{b(z)}\Big(\frac{x}{b(z)}\Big)^{a(z)},$$

$$\frac{\partial \log f(x|\theta,z)}{\partial b_0} = \frac{1}{a(z)} + \log x - \log b(z) - \Big(\frac{x}{b(z)}\Big)^{a(z)} \log\Big(\frac{x}{b(z)}\Big),$$

$$\frac{\partial \log g(x|\gamma,z)}{\partial c_0} = \frac{\log x - \mu(z)}{\sigma(z)^2},$$

$$\frac{\partial \log g(x|\gamma,z)}{\partial d_0} = -\frac{1}{\sigma(z)} + \frac{\big(\log x - \mu(z)\big)^2}{\sigma(z)^3}.$$

Note for any $z \in \{z_1, ..., z_{n_0}\}$, as $x \to \infty$,

$$\Big|\frac{\partial \log f(x|\theta_{n_0}, z)}{\partial a_0}\Big|, \ \Big|\frac{\partial \log f(x|\theta_{n_0}, z)}{\partial b_0}\Big|, \ \Big|\frac{\partial \log g(x|\gamma_{n_0}, z)}{\partial c_0}\Big|, \ \Big|\frac{\partial \log g(x|\gamma_{n_0}, z)}{\partial d_0}\Big|$$

all tend to infinity, however, the converge rate are all different. Hence, if

$$(v_1 + v_2 \cdot z)^2 + (v_3 + v_4 \cdot z)^2 + (v_1' + v_2' \cdot z)^2 + (v_3' + v_4' \cdot z)^2 \neq 0,$$

then as $x \to \infty$,

$$\Big|\frac{\partial \log f(x|\theta_{n_0}, z)}{\partial a_0} \cdot (v_1 + v_2 \cdot z) + \frac{\partial \log f(x|\theta_{n_0}, z)}{\partial b_0} \cdot (v_3 + v_4 \cdot z) - \frac{\partial \log g(x|\gamma_{n_0}, z)}{\partial c_0} \cdot (v_1' + v_2' \cdot z) - \frac{\partial \log g(x|\gamma_{n_0}, z)}{\partial d_0} \cdot (v_3' + v_4' \cdot z)\Big| \to \infty.$$

But by definition for all (x, z), $|\omega(x, z)| \leq 4$, which is a contradiction to Equality (2.2.25). Hence, the assumption $\sigma^2 = 0$ does not hold and we get $\sigma^2 > 0$. If

$$(v_1 + v_2 \cdot z)^2 + (v_3 + v_4 \cdot z)^2 + (v_1' + v_2' \cdot z)^2 + (v_3' + v_4' \cdot z)^2 = 0,$$

by (2.2.25), we have then for all $(x, z) \in \mathbb{R}^+ \times \{z_1, ..., z_{n_0}\}$,

$$\omega(x, z) = k \quad H\text{-a.s.} \tag{2.2.26}$$

By the definition of the two competing model classes, they are disjoint. Further, the function H has a density function. Hence, for any $z \in \{z_1, ..., z_{n_0}\}$ there exists an $x_z > 0$ and $\delta > 0$ such that $F(x_z|\theta_{n_0}, z) \neq G(x_z|\gamma_{n_0}, z)$ and the density function of $H(\cdot|z)$ is bounded away from zero on $(x_z - \delta, x_z + \delta)$.

Without loss of generality, we can assume that $F(x_z|\theta_{n_0}, z) > G(x_z|\gamma_{n_0}, z)$. Define $m_1, m_2 \in \mathbb{R} \cup \{\infty\}$ with

$$m_{z1} := \sup\{x : F(x|\theta_{n_0}, z) = G(x|\gamma_{n_0}, z) \ \text{ and } \ x < x_z\},$$
$$m_{z2} := \inf\{x : F(x|\theta_{n_0}, z) = G(x|\gamma_{n_0}, z) \ \text{ and } \ x > x_z\},$$

where we let $\inf\{\emptyset\} = \infty$. Since F and G are continuous functions in x and

$$F(0|\theta_{n_0}, z) = G(0|\gamma_{n_0}, z) = 0$$
$$\lim_{x\to\infty} F(x|\theta_{n_0}, z) = \lim_{x\to\infty} G(x|\gamma_{n_0}, z) = 0,$$

thus, m_{z1} and m_{z2} exist and $F(\cdot|\theta_{n_0}, z) - G(\cdot|\gamma_{n_0}, z) > 0$ on (m_{z1}, m_{z2}). Consequently,

$$\omega(m_{z1}, z) - \omega(m_{z2}, z) = 2\int_{m_{z1}}^{m_{z2}} \big(G(u|\gamma_{n_0}, z) - F(u|\theta_{n_0}, z)\big)\, dH(u|z) < 0.$$

whereby $\omega(m_{z2}, z) = \lim_{x\to\infty} \omega(x, z)$ if $m_{z2} = \infty$. But it contradicts Equality (2.2.26). Hence, it muss hold that $\sigma^2 > 0$.

2.3 The Case with Number of Covariates Tending to Infinity

With m fixed, the underlying distribution function $H(\cdot|z)$ can not be estimated by the empirical distribution at covariate value z consistently any more. Instead it will be estimated by the kernel Nadaraya-Watson estimator:

$$\hat{H}_n(x|z) := \sum_{i=1}^{n} w_{ni}(z,h) \cdot I(X_i \leq x)$$

where the function $w_{ni} : [0,1]^d \times (0,\infty) \to \mathbb{R}^+$ is given by

$$w_{ni}(z,h) := \frac{K(\frac{z_i - z}{h})}{\sum_{k=1}^{n} K(\frac{z_k - z}{h})}$$

with kernel function $K : \mathbb{R}^d \to \mathbb{R}^+$ and bandwidth $h > 0$. Further we denote the kernel estimator for the joint distribution function as

$$\hat{Q}_n(x,z) := \frac{1}{n_0} \sum_{i=1}^{n_0} \hat{H}_n(x|z_i) \cdot I(z_i \leq z).$$

The distance $d_H(\mathcal{F})$ defined in (2.1.2) can then be estimated by

$$\hat{d}_{H,n}(\mathcal{F}) := \int \left(\hat{H}_n(x|z) - F(x|\hat{\theta}_n, z)\right)^2 d\hat{Q}_n(x,z).$$

For the class $\mathcal{G}$, the estimator $\hat{d}_{H,n}(\mathcal{G})$ is defined in an analogous way. As test statistic we take the difference of the estimated distances again

$$T_n := \hat{d}_{H,n}(\mathcal{F}) - \hat{d}_{H,n}(\mathcal{G}).$$

In this section, we will show similar results as in Section 2.2. For the asymptotic properties of the kernel estimator, we assume the following conditions hold true throughout this section.

(i) The function H has bounded derivative and Hessian matrix with respect to z. The function $\|\partial^2 H/\partial z \partial x\|$ is dominated by a Lebesgue integrable function independent of z.

(ii) As $n_0 \to \infty$, $h \to 0$ and $n_0 h^{2d} \to \infty$.

(iii) Let K be a bounded positive integrable function on $[-1,1]^d$, zero otherwise. Further, for all $x \in \mathbb{R}^d$, $K(x) = K(|x|)$.

Here the symmetry assumption on the kernel is only for simplicity of proofs.

Again, the assumptions on the distribution model classes are formulated in terms of $\mathcal{F}$ and it is understood that corresponding conditions are also made on $\mathcal{G}$.

C1 The distribution $F(\cdot|\theta, z)$ has a density function $f(\cdot|\theta, z)$, which is strictly positive $H(\cdot|z)$-a.s. for each $(\theta, z) \in \Theta \times [0, 1]^d$.

C2 The function $\log f$ is three times continuously differentiable in θ on Θ.

C3 The function $\log f$ is dominated by an H square integrable function independent of z.

C4 For each $n_0 \in \mathbb{N}$ the function L_{f,n_0} reaches maximum at θ_{n_0} on Θ, which are interior points of Θ.

C5 The functions $\|\partial \log f/\partial\theta\|^4$ and $\|\partial^2 \log f/\partial\theta^2\|^4$ are dominated by H integrable functions independent of z.

C6 For any $i,\ j,\ k \in \{1, 2, ..., p\}$, the function $\partial^3 \log f/\partial\theta_i\partial\theta_j\partial\theta_k$ is dominated by an H square integrable function independent of z.

C7 The functions $\dot{F}$ and $\ddot{F}$ exist and they are bounded.

C8 The function $L_{f,\infty}$ has a unique maximizer on Θ at θ_*, which is an interior point of Θ. The Hessian matrix $\ddot{L}_{f,\infty}(\theta_*)$ is invertible with inverse $\ddot{L}^{-1}_{f,\infty}(\theta_*)$.

C9 The function $\|\partial\dot{F}/\partial z\|$ is dominated by an H integrable function independent of z and the functions $\|\partial F/\partial z\|$ and $\|\partial^2 \log f/\partial z\partial\theta\|$ are dominated by H square integrable functions independent of z.

Note that the grade of the integrability are doubled in C3, C5 and C6 in comparison to the assumptions in Section 2.2 because the data can not be seen as i.i.d. as $n_0 \to \infty$.

In the following lemmas we show first the relations among $\hat{\theta}_n$, θ_* and θ_{n_0}.

Lemma 2.3.1. *If* C1–C3, C5 *and* C8 *hold, then* $\|\hat{\theta}_n - \theta_*\| = o_p(1)$.

Proof. The assertion can be shown analogously to Lemma 2.2.2 based on Lemma A.2.7. □

Lemma 2.3.2. *If* C1–C5 *and* C8 *hold, then* $\|\theta_{n_0} - \theta_*\| = o(1)$.

Proof. For each $\theta \in \Theta$, by the definition of Riemann integral, under C3

$$\frac{1}{n_0}\sum_{i=1}^{n_0}\int \log f(x|\theta, z_i)dH(x|z_i) - \int \log f(x|\theta, z)dQ(x, z) = o(1).$$

Thus, under C1–C5 and C8, the assertion follows by the similar arguments used in Lemma 2.2.2. □

Corollary 2.3.3. *If* C1–C5 *and* C8 *hold, then* $\|\hat{\theta}_n - \theta_{n_0}\| = o_p(1)$.

Proof. The assertion follows directly from Lemma 2.3.1 and Lemma 2.3.2. □

Lemma 2.3.4. *If* C1–C6 *and* C8 *hold, then*

$$\sqrt{n}\cdot\|\hat{\theta}_n - \theta_{n_0}\| = O_p(1),$$
$$\sqrt{n}\cdot\left\|\hat{\theta}_n - \theta_{n_0} + \ddot{L}^{-1}_{f,\infty}(\theta_*)\cdot\dot{\hat{L}}_{f,n}(\theta_{n_0})\right\| = o_p(1).$$

Proof. For any $a \in \{1, ..., p\}$, by Lemma 2.3.2 and Lemma A.2.8 with

$$\psi_{1n}(x, z) = 0 \quad\text{and}\quad \psi_{2n}(x, z) = \frac{\partial \log f(x|\theta_{n_0}, z)}{\partial\theta_a}$$

for $n \in m\cdot\mathbb{N}$ and $z \in [0, 1]^d$, under C5,

$$\sqrt{n}\cdot\int\frac{\partial \log f(x|\theta_{n_0}, z)}{\partial\theta_a}dQ_n(x, z) - \sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int\frac{\partial \log f(x|\theta_{n_0}, z_i)}{\partial\theta_a}dH(x|z_i)$$

convergences to a normal distribution. Further, by the definition of θ_{n_0},

$$\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int\frac{\partial \log f(x|\theta_{n_0}, z_i)}{\partial\theta_a}dH(x|z_i) = 0.$$

Thus, for any $a \in \{1, ..., p\}$,

$$\sqrt{n}\cdot\int\frac{\partial \log f(x|\theta_{n_0}, z)}{\partial\theta_a}dQ_n(x, z)$$

convergences to a normal distribution. Therefore,

$$\sqrt{n}\cdot\left\|\dot{\hat{L}}_{f,n}(\theta_{n_0})\right\| = \sqrt{n}\cdot\left\|\int\frac{\partial \log f(x|\theta_{n_0}, z)}{\partial\theta}dQ_n(x, z)\right\| = O_p(1). \quad (2.3.1)$$

Under C1-C6 and C8, the rest of the proof can be stated similarly as in the proof of Lemma 2.2.3. □

The reason we still work with θ_{n_0} in this section is that we do not have

$$\sqrt{n} \cdot \left\| \hat{\theta}_n - \theta_* + \ddot{L}_{f,\infty}^{-1}(\theta_*) \cdot \dot{\hat{L}}_{f,n}(\theta_*) \right\| = o_p(1)$$

in general. Because $\sqrt{n} \cdot \left\| \dot{\hat{L}}_{f,n}(\theta_*) \right\| = O_p(1)$ does not hold.

In order to state the main theorems we introduce the functions $C_{\mathcal{F}} : \Theta \to \mathbb{R}^p$, and $N_{\mathcal{F},n}, N_{\mathcal{F}}, N_{\mathcal{F},n}^1, N_{\mathcal{F}}^1 : \mathbb{R} \times \Theta \times [0,1]^d \to \mathbb{R}$ with

$$\begin{aligned}
C_{\mathcal{F}}(\theta) &:= \int \big(H(x|z) - F(x|\theta, z)\big) \cdot \dot{F}(x|\theta, z)\, dQ(x,z), \\
N_{\mathcal{F},n}(x,\theta,z) &:= \big(E[\hat{H}_n(x|z)] - F(x|\theta,z)\big)^2 + 2\int_x^\infty \big(E[\hat{H}_n(u|z)] - F(u|\theta,z)\big)\, dE[\hat{H}_n(u|z)], \\
N_{\mathcal{F}}(x,\theta,z) &:= \big(H(x|z) - F(x|\theta,z)\big)^2 + 2\int_x^\infty \big(H(u|z) - F(u|\theta,z)\big)\, dH(u|z), \\
N_{\mathcal{F},n}^1(x,\theta,z) &:= N_{\mathcal{F},n}(x,\theta,z) + 2C_{\mathcal{F}}^T(\theta_*) \cdot \ddot{L}_{f,\infty}^{-1}(\theta_*) \cdot \frac{\partial \log f(x|\theta,z)}{\partial \theta}, \\
N_{\mathcal{F}}^1(x,\theta,z) &:= N_{\mathcal{F}}(x,\theta,z) + 2C_{\mathcal{F}}^T(\theta_*) \cdot \ddot{L}_{f,\infty}^{-1}(\theta_*) \cdot \frac{\partial \log f(x|\theta,z)}{\partial \theta}.
\end{aligned}$$

Further for each $n \in m \cdot \mathbb{N}$, let $d_{H,n}(\mathcal{F}) \in \mathbb{R}$ be defined as

$$d_{H,n}(\mathcal{F}) := \frac{1}{n_0} \sum_{i=1}^{n_0} \int \big(E[\hat{H}_n(x|z_i)] - F(x|\theta_{n_0}, z_i)\big)^2\, dE[\hat{H}_n(x|z_i)].$$

Under C7 the function $C_{\mathcal{F}}$ is bounded on Θ. Thus, by the boundedness of F and H, under C5 the function $N_{\mathcal{F}}^1$ is H square integrable function. Therefore, we can define the constant

$$\sigma_{\mathcal{F}}^2 := \int \left(\int \big(N_{\mathcal{F}}^1(x,\theta_*,z)\big)^2\, dH(x|z) - \left(\int N_{\mathcal{F}}^1(x,\theta_*,z)\, dH(x|z) \right)^2 \right) dz.$$

Theorem 2.3.5. *Let* C1–C9 be satisfied, then we have

$$\sqrt{n} \cdot \big(\hat{d}_{H,n}(\mathcal{F}) - d_{H,n}(\mathcal{F})\big) \xrightarrow{d} \mathcal{N}(0, \sigma_{\mathcal{F}}^2),$$

and

$$d_{H,n}(\mathcal{F}) \to d_H(\mathcal{F}).$$

Proof. Note that $\hat{H}_n(x|z) - F(x|\hat{\theta}_n, z)$ can be written as

$$\big(\hat{H}_n(x|z) - E[\hat{H}_n(x|z)]\big) + \big(E[\hat{H}_n(x|z)] - F(x|\theta_{n_0}, z)\big) - \big(F(x|\hat{\theta}_n, z) - F(x|\theta_{n_0}, z)\big).$$

Hence,

$$\sqrt{n} \cdot \hat{d}_{H,n}(\mathcal{F}) = \sqrt{n} \cdot \int \big(\hat{H}_n(x|z) - E[\hat{H}_n(x|z)]\big)^2\, d\hat{Q}_n(x,z)$$

$$
\begin{aligned}
&+ 2\sqrt{n} \cdot \int \big(\hat{H}_n(x|z) - E[\hat{H}_n(x|z)]\big) \cdot \big(E[\hat{H}_n(x|z)] - F(x|\theta_{n_0}, z)\big)\, d\hat{Q}_n(x,z) \\
&- 2\sqrt{n} \cdot \int \big(E[\hat{H}_n(x|z)] - F(x|\theta_{n_0}, z)\big) \cdot \big(F(x|\hat{\theta}_n, z) - F(x|\theta_{n_0}, z)\big)\, d\hat{Q}_n(x,z) \\
&+ \sqrt{n} \cdot \int \big(E[\hat{H}_n(x|z)] - F(x|\theta_{n_0}, z)\big)^2 d\hat{Q}_n(x,z) \\
&- 2\sqrt{n} \cdot \int \big(\hat{H}_n(x|z) - E[\hat{H}_n(x|z)]\big) \cdot \big(F(x|\hat{\theta}_n, z) - F(x|\theta_{n_0}, z)\big)\, d\hat{Q}_n(x,z) \\
&+ \sqrt{n} \cdot \int \big(F(x|\hat{\theta}_n, z) - F(x|\theta_{n_0}, z)\big)^2\, d\hat{Q}_n(x,z) =: \sum_{k=1}^{6} T_{kn}.
\end{aligned}
$$

As in the proof of Theorem 2.2.4, by Lemma A.2.9 it can be shown that

$$T_{1n} = o_p(1),$$

and

$$
\begin{aligned}
T_{2n} =& 2\sqrt{n} \cdot \int \int_x^\infty \big(E[\hat{H}_n(u|z)] - F(u|\theta_{n_0}, z)\big)\, dE[\hat{H}_n(u|z)]\, d\hat{Q}_n(x,z) \\
&- 2\sqrt{n} \cdot \int \int_x^\infty \big(E[\hat{H}_n(u|z)] - F(u|\theta_{n_0}, z)\big)\, dE[\hat{H}_n(u|z)]\, dE[\hat{Q}_n(x,z)] + o_p(1).
\end{aligned}
$$

By Lemma 2.3.4 under C7,

$$T_{3n} = -2\sqrt{n} \cdot \Big(\int \big(E[\hat{H}_n(x|z)] - F(x|\theta_{n_0}, z)\big)\dot{F}(x|\theta_{n_0}, z)\, d\hat{Q}_n(x,z)\Big)^T \cdot (\hat{\theta}_n - \theta_{n_0}) + o_p(1).$$

In the following, we show that

$$\Big\| \int \big(E[\hat{H}_n(x|z)] - F(x|\theta_{n_0}, z)\big) \cdot \dot{F}(x|\theta_{n_0}, z)\, d\hat{Q}_n(x,z) - C_{\mathcal{F}}^T(\theta_*)\Big\| = o_p(1). \tag{2.3.2}$$

Note first

$$
\begin{aligned}
&\Big\| \int \big(E[\hat{H}_n(x|z)] - F(x|\theta_{n_0}, z)\big) \cdot \dot{F}(x|\theta_{n_0}, z)\, d\hat{Q}_n(x,z) - C_{\mathcal{F}}^T(\theta_*)\Big\| \\
\le& \Big\| \int \big(E[\hat{H}_n(x|z)] - F(x|\theta_{n_0}, z)\big) \cdot \dot{F}(x|\theta_{n_0}, z)\, d\big(\hat{Q}_n(x,z) - E[\hat{Q}_n(x,z)]\big)\Big\| \\
&+ \Big\| \int \big(E[\hat{H}_n(x|z)] - F(x|\theta_{n_0}, z)\big) \cdot \dot{F}(x|\theta_{n_0}, z)\, dE[\hat{Q}_n(x,z)] \\
&\qquad - \int \big(E[\hat{H}_n(x|z)] - F(x|\theta_*, z)\big) \cdot \dot{F}(x|\theta_*, z)\, dE[\hat{Q}_n(x,z)]\Big\| \\
&+ \Big\| \int \big(E[\hat{H}_n(x|z)] - H(x|z)\big) \cdot \dot{F}(x|\theta_*, z)\, dE[\hat{Q}_n(x,z)]\Big\| \\
&+ \Big\| \int \big(H(x|z) - F(x|\theta_*, z)\big) \cdot \dot{F}(x|\theta_*, z)\, dE[\hat{Q}_n(x,z)] - C_{\mathcal{F}}^T(\theta_*)\Big\|.
\end{aligned} \tag{2.3.3}
$$

For $a \in \{1, ..., p\}$, by Lemma A.2.11 with

$$\psi_n(x,z) = \big(E[\hat{H}_n(x|z)] - F(x|\theta_{n_0}, z)\big) \cdot \frac{\partial F(x|\theta_{n_0}, z)}{\partial \theta_a},$$

under C7,

$$\int \big(E[\hat{H}_n(x|z)] - F(x|\theta_{n_0}, z)\big) \cdot \frac{\partial F(x|\theta_{n_0}, z)}{\partial \theta_a} \, d\big(\hat{Q}_n(x,z) - E[\hat{Q}_n(x,z)]\big) = o_p(1).$$

Hence, the first term on the right-hand side of (2.3.3) equals $o_p(1)$. Note that under C7 the function

$$\int \big(E[\hat{H}_n(x|z)] - F(x|\cdot, z)\big) \cdot \dot{F}(x|\cdot, z) \, dE[\hat{Q}_n(x,z)]$$

has a bounded derivative on Θ. Thus, by Lemma 2.3.2, the second term equals $o(1)$ as well. By Lemma A.2.6, under C7 there exists a constant $C > 0$ such that the third term term on the right-hand side of (2.3.3) is bounded by

$$C \cdot \max_{(x,z) \in \mathbb{R} \times [0,1]^d} \big| E[\hat{H}_n(x|z)] - H(x|z) \big| = o(1).$$

For each $a \in \{1, ..., p\}$, by Lemma A.2.12 with

$$\psi(x,z) = \big(H(x|z) - F(x|\theta_*, z)\big) \cdot \frac{\partial F(x|\theta_*, z)}{\partial \theta_a},$$

under C7, C9 and Assumption (i) the last term on the right-hand side of (2.3.3) is equal to $o(1)$ as well. Consequently, (2.3.2) follows. Thus, by Lemma 2.3.4, under C1–C6 and C8,

$$T_{3n} = -2\sqrt{n} \cdot C_{\mathcal{F}}^T(\theta_*) \cdot (\hat{\theta}_n - \theta_{n_0}) + o_p(1) = 2\sqrt{n} \cdot C_{\mathcal{F}}^T(\theta_*) \cdot \ddot{L}_{f,\infty}^{-1}(\theta_*) \cdot \dot{\hat{L}}_{f,n}(\theta_{n_0}) + o_p(1).$$

As in the Theorem 2.2.4, by Lemma 2.3.4, it can be shown that T_{5n} and T_{6n} converge to zero in probability. Therefore,

$$\begin{aligned}
\sqrt{n} \cdot \hat{d}_{H,n}(\mathcal{F}) &= \sqrt{n} \cdot \int N_{\mathcal{F},n}(x, \theta_{n_0}, z) \, d\hat{Q}_n(x,z) \\
&\quad - 2\sqrt{n} \int \int_x^\infty \big(E[\hat{H}_n(u|z)] - F(u|\theta_{n_0}, z)\big) \, dE[\hat{H}_n(u|z)] \, dE[\hat{Q}_n(x,z)] + o_p(1) \\
&\quad + 2\sqrt{n} \cdot C_{\mathcal{F}}^T(\theta_*) \cdot \ddot{L}_{f,\infty}^{-1}(\theta_*) \cdot \int \frac{\partial \log f(x|\theta_{n_0}, z)}{\partial \theta} dQ_n(x,z). \qquad (2.3.4)
\end{aligned}$$

Note that by the definition of θ_{n_0}, under C3 and C5,

$$C_{\mathcal{F}}^T(\theta_*) \cdot \ddot{L}_{f,\infty}^{-1}(\theta_*) \cdot \int \frac{\partial \log f(x|\theta_{n_0}, z)}{\partial \theta} dE[Q_n(x,z)]$$

$$=C_{\mathcal{F}}^T(\theta_*)\cdot\ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\dot{L}_{f,n_0}(\theta_{n_0})=0.$$

Hence, we can write

$$\begin{aligned}
d_{H,n}(\mathcal{F}) &= \int N_{\mathcal{F},n}(x,\theta_{n_0},z)\,dE[\hat{Q}_n(x,z)]\\
&-2\int\int_x^\infty \big(E[\hat{H}_n(u|z)]-F(u|\theta_{n_0},z)\big)\,dE[\hat{H}_n(u|z)]\,dE[\hat{Q}_n(x,z)]\\
&+2C_{\mathcal{F}}^T(\theta_*)\cdot\ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\int\frac{\partial\log f(x|\theta_{n_0},z)}{\partial\theta}dE[Q_n(x,z)].
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\sqrt{n}\cdot\big(\hat{d}_{H,n}(\mathcal{F})-d_{H,n}(\mathcal{F})\big)\\
=&\sqrt{n}\cdot\int N_{\mathcal{F},n}(x,\theta_{n_0},z)\,d\big(\hat{Q}_n(x,z)-E[\hat{Q}_n(x,z)]\big)\\
&+2\sqrt{n}\cdot C_{\mathcal{F}}^T(\theta_*)\cdot\ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\int\frac{\partial\log f(x|\theta_{n_0},z)}{\partial\theta}d\big(Q_n(x,z)-E[Q_n(x,z)]\big)+o_p(1).
\end{aligned}$$

Note that for all $(x,\theta,z)\in\mathbb{R}\times\Theta\times[0,1]^d$, by partial integration,

$$\begin{aligned}
&\big|N_{\mathcal{F},n}(x,\theta,z)-N_{\mathcal{F}}(x,\theta,z)\big|\\
\le&\Big|\big(E[\hat{H}_n(x|z)]-H(x|z)\big)\big(E[\hat{H}_n(x|z)]+H(x|z)-2\cdot F(x|\theta,z)\big)\Big|\\
&+2\cdot\Big|\int_x^\infty \big(E[\hat{H}_n(u|z)]-F(u|\theta,z)\big)-\big(H(u|z)-F(u|\theta,z)\big)\,dE[\hat{H}_n(u|z)]\Big|\\
&+2\cdot\Big|\int_x^\infty \big(H(u|z)-F(u|\theta,z)\big)\,d\big(E[\hat{H}_n(u|z)]-H(u|z)\big)\Big|\\
\le&4\cdot\max_x\big|E[\hat{H}_n(x|z)]-H(x|z)\big|+2\cdot\max_x\big|E[\hat{H}_n(x|z)]-H(x|z)\big|\\
&+2\cdot\Big|\big(E[\hat{H}_n(x|z)]-H(x|z)\big)\big(H(x|z)-F(x|\theta,z)\big)\Big|\\
&+2\cdot\Big|\int_x^\infty \big(E[\hat{H}_n(u|z)]-H(u|z)\big)\,d\big(H(u|z)-F(u|\theta,z)\big)\Big|\\
\le&12\cdot\max_x\big|E[\hat{H}_n(x|z)]-H(x|z)\big|.
\end{aligned}\tag{2.3.5}$$

Thus, by (A.2.10) and Lemma A.2.6, we can show the variance

$$Var\Big[\sqrt{n}\cdot\int\big(N_{\mathcal{F},n}(x,\theta_{n_0},z)-N_{\mathcal{F}}(x,\theta_{n_0},z)\big)d\hat{Q}_n(x,z)\Big]=o(1).$$

Further, the expectation

$$E\Big[\sqrt{n}\cdot\int\big(N_{\mathcal{F},n}(x,\theta_{n_0},z)-N_{\mathcal{F}}(x,\theta_{n_0},z)\big)\,d\big(\hat{Q}_n(x,z)-E[\hat{Q}_n(x,z)]\big)\Big]=0.$$

Hence, by Chebyshev's inequality,

$$\begin{aligned}&\sqrt{n}\cdot\int N_{\mathcal{F},n}(x,\theta_{n_0},z)\,d\big(\hat{Q}_n(x,z)-E[\hat{Q}_n(x,z)]\big)\\=&\sqrt{n}\cdot\int N_{\mathcal{F}}(x,\theta_{n_0},z)\,d\big(\hat{Q}_n(x,z)-E[\hat{Q}_n(x,z)]\big)+o_p(1).\end{aligned}\tag{2.3.6}$$

Therefore, we can write

$$\begin{aligned}&\sqrt{n}\cdot\big(\hat{d}_{H,n}(\mathcal{F})-d_{H,n}(\mathcal{F})\big)\\=&\sqrt{n}\cdot\int N_{\mathcal{F}}(x,\theta_{n_0},z)\,d\big(\hat{Q}_n(x,z)-E[\hat{Q}_n(x,z)]\big)\\&+2\sqrt{n}\cdot C_{\mathcal{F}}^T(\theta_*)\cdot\ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\int\frac{\partial\log f(x|\theta_{n_0},z)}{\partial\theta}d\big(Q_n(x,z)-E[Q_n(x,z)]\big)+o_p(1).\end{aligned}$$

We define

$$\sigma_{\mathcal{F},n}^2:=\frac{1}{n}\sum_{i=1}^n\Big(\int\big(N_{\mathcal{F}}^1(x,\theta_{n_0},z_i)\big)^2\,dH(x|z_i)-\Big(\int N_{\mathcal{F}}^1(x,\theta_{n_0},z_i)\,dH(x|z_i)\Big)^2\Big).$$

In the sequel, we will show that the conditions of Lemma A.2.8 with

$$\psi_{1n}(x,z)=N_{\mathcal{F}}(x,\theta_{n_0},z),$$

and

$$\psi_{2n}(x,z)=2\cdot C_{\mathcal{F}}^T(\theta_*)\cdot\ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\frac{\partial\log f(x|\theta_{n_0},z)}{\partial\theta}$$

are fulfilled. Note first that under C5, for each n, $\|\psi_{1n}(x,z)\|^4$ and $\|\psi_{2n}(x,z)\|^4$ are dominated by the same H integrable function. Further under C5 and C7, the functions $N_{\mathcal{F}}^1$ and $\dot{N}_{\mathcal{F}}^1$ are both dominated by H square integrable functions independent of z. Therefore

$$\frac{1}{n}\sum_{i=1}^n\int\big(N_{\mathcal{F}}^1(x,\cdot,z_i)\big)^2dH(x|z_i)$$

has a bounded derivative on Θ. By Lemma 2.3.2, we get then

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n\int\big(N_{\mathcal{F}}^1(x,\theta_{n_0},z_i)\big)^2dH(x|z_i)=&\frac{1}{n}\sum_{i=1}^n\int\big(N_{\mathcal{F}}^1(x,\theta_*,z_i)\big)^2dH(x|z_i)+o(1)\\=&\int\big(N_{\mathcal{F}}^1(x,\theta_*,z)\big)^2dQ(x,z)+o(1).\end{aligned}$$

With the same arguments, it can be shown that

$$\frac{1}{n}\sum_{i=1}^n\Big(\int N_{\mathcal{F}}^1(x,\theta_{n_0},z_i)dH(x|z_i)\Big)^2=\int\Big(\int N_{\mathcal{F}}^1(x,\theta_*,z)dH(x|z)\Big)^2dz+o(1).$$

Hence, $\sigma^2_{\mathcal{F},n} = \sigma^2_{\mathcal{F}} + o(1)$. Since H and F are bounded, under C9 and Assumption (i), there exists a constant $C > 0$ such that

$$\begin{aligned}\Big\|\frac{\partial N_{\mathcal{F}}(x,\theta,z)}{\partial z}\Big\| \leq & 2\cdot\Big\|\big(H(x|z) - F(x|\theta,z)\big)\Big(\frac{\partial H(x|z)}{\partial z} - \frac{\partial F(x|\theta,z)}{\partial z}\Big)\Big\| \\ & + 2\cdot\Big\|\int_x^\infty \Big(\frac{\partial H(u|z)}{\partial z} - \frac{\partial F(u|\theta,z)}{\partial z}\Big)\, dH(u|z)\Big\| \\ & + 2\cdot\Big\|\int_x^\infty \big(H(u|z) - F(u|\theta,z)\big)\cdot\frac{\partial H(u|z)}{\partial z\partial u}\, du\Big\| \\ \leq & C + C\cdot\Big\|\frac{\partial F(x|\theta,z)}{\partial z}\Big\|. \end{aligned} \tag{2.3.7}$$

Thus, under C9 the function $\|\partial N_{\mathcal{F}}/\partial z\|^2$ is dominated by an H integrable function independent of z. Therefore, the first part of the assertion follows from Lemma A.2.8.

For the second part of the assertion, we write first

$$\begin{aligned} & \big|d_{H,n}(\mathcal{F}) - d_H(\mathcal{F})\big| \\ = & \Big|\frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(E[\hat{H}_n(x|z_i)] - F(x|\theta_{n_0}, z_i)\big)^2 - \big(H(x|z_i) - F(x|\theta_*, z_i)\big)^2\, dE[\hat{H}_n(x|z_i)]\Big| \\ & + \Big|\frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(H(x|z_i) - F(x|\theta_*, z_i)\big)^2\, d\big(E[\hat{H}_n(x|z_i)] - H(x|z_i)\big)\Big|. \end{aligned} \tag{2.3.8}$$

The first term on the right side of (2.3.8) is bounded by

$$\begin{aligned} & 4\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int \big|E[\hat{H}_n(x|z_i)] - H(x|z_i) + F(x|\theta_*, z_i) - F(x|\theta_{n_0}, z_i)\big|\, dE[\hat{H}_n(x|z_i)] \\ \leq & 4\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\max_{x\in\mathbb{R}}\big|E[\hat{H}_n(x|z_i)] - H(x|z_i)\big| + \max_{x\in\mathbb{R},\theta\in\Theta}\big\|\dot{F}(x|\theta,z_i)\big\|\cdot\big\|\theta_* - \theta_{n_0}\big\|\Big). \end{aligned}$$

By Lemma 2.3.2 and Lemma A.2.6, under C7 it can be shown to be $o(1)$. By partial integration, the second term on the right side of (2.3.8) is bounded by

$$\begin{aligned} & \Big|2\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(E[\hat{H}_n(x|z_i)] - H(x|z_i)\big)\big(H(x|z_i) - F(x|\theta_*, z_i)\big)\, d\big(H(x|z_i) - F(x|\theta_*, z_i)\big)\Big| \\ \leq & 4\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\max_{x\in\mathbb{R}}\big|E[\hat{H}_n(x|z_i)] - H(x|z_i)\big|. \end{aligned}$$

Hence, with the same arguments, it equals $o(1)$ as well. Therefore,

$$d_{H,n}(\mathcal{F}) - d_H(\mathcal{F}) = o(1).$$

□

Let the functions $\hat{C}_{\mathcal{F},n} : \Theta \to \mathbb{R}^p$ and $\hat{N}^1_{\mathcal{F},n} : \mathbb{R} \times \Theta \times [0,1]^d \to \mathbb{R}$ be defined with

$$\hat{C}_{\mathcal{F},n}(\theta) := \int \big(\hat{H}_n(x|z) - F(x|\theta,z)\big) \cdot \dot{F}(x|\theta,z)\, d\hat{Q}_n(x,z),$$

$$\hat{N}^1_{\mathcal{F},n}(x,\theta,z) := \big(\hat{H}_n(x|z) - F(x|\theta,z)\big)^2 + 2\int_x^\infty \big(\hat{H}_n(u|z) - F(u|\theta,z)\big)\, d\hat{H}_n(u|z) + 2\hat{C}^T_{\mathcal{F},n}(\hat{\theta}_n) \cdot \ddot{\hat{L}}^{-1}_{f,n}(\hat{\theta}_n) \cdot \frac{\partial \log f(x|\theta,z)}{\partial\theta}.$$

In the next lemma we show that the variance σ^2 can be estimated consistently by the plug-in estimator:

$$\hat{\sigma}^2_{\mathcal{F},n} := \frac{1}{n_0}\sum_{i=1}^{n_0}\bigg(\int \big(\hat{N}^1_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)\big)^2\, d\hat{H}_n(x|z_i) - \Big(\int \hat{N}^1_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)\, d\hat{H}_n(x|z_i)\Big)^2\bigg).$$

Lemma 2.3.6. *If* C1–C9 *hold, then we have*

$$\hat{\sigma}^2_{\mathcal{F},n} = \sigma^2_{\mathcal{F}} + o_p(1).$$

Proof. First, we show that

$$\frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(\hat{N}^1_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)\big)^2\, d\hat{H}_n(x|z_i) = \frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(N^1_{\mathcal{F}}(x,\theta_*,z_i)\big)^2\, dH(x|z_i) + o_p(1). \quad (2.3.9)$$

Note that we have

$$\begin{aligned}
&\Big|\frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(\hat{N}^1_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)\big)^2\, d\hat{H}_n(x|z_i) - \frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(N^1_{\mathcal{F}}(x,\theta_*,z_i)\big)^2\, dH(x|z_i)\Big|\\
\le&\Big|\frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(\hat{N}^1_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)\big)^2\, d\hat{H}_n(x|z_i) - \frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(N^1_{\mathcal{F},n}(x,\theta_*,z_i)\big)^2\, dE[\hat{H}_n(x|z_i)]\Big|\\
&+\Big|\frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(N^1_{\mathcal{F},n}(x,\theta_*,z_i)\big)^2\, dE[\hat{H}_n(x|z_i)] - \frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(N^1_{\mathcal{F}}(x,\theta_*,z_i)\big)^2\, dE[\hat{H}_n(x|z_i)]\Big|\\
&+\Big|\frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(N^1_{\mathcal{F}}(x,\theta_*,z_i)\big)^2\, dE[\hat{H}_n(x|z_i)] - \frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(N^1_{\mathcal{F}}(x,\theta_*,z_i)\big)^2\, dH(x|z_i)\Big|\\
=:&\sigma_{1n} + \sigma_{2n} + \sigma_{3n}.
\end{aligned}$$

In the sequel, we show that $\sigma_{in} = o_p(1)$, for all $i \in \{1,2,3\}$.

Analogously to Lemma 2.2.5, under C1–C8, by Lemma 2.3.1, Corollary A.2.10 and Lemma A.2.11, it can be shown that $\sigma_{1n} = o_p(1)$.

Further by Cauchy-Schwarz's inequality, we get

$$\sigma_{2n} \le \frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(N^1_{\mathcal{F},n}(x,\theta_*,z_i) - N^1_{\mathcal{F}}(x,\theta_*,z_i)\big)^2 dE[\hat{H}_n(x|z_i)]$$
$$\times \int \big(N^1_{\mathcal{F},n}(x,\theta_*,z_i) + N^1_{\mathcal{F}}(x,\theta_*,z_i)\big)^2 dE[\hat{H}_n(x|z_i)].$$

Under C3, C5 and C7, the functions $N^1_{\mathcal{F},n}$ are $N^1_{\mathcal{F}}$ are dominated by H square integrable functions independent of z, hence, the function

$$\int \big(N^1_{\mathcal{F},n}(x,\theta_*,\cdot) + N^1_{\mathcal{F}}(x,\theta_*,\cdot)\big)^2 dE[\hat{H}_n(x|\cdot)]$$

is bounded on $[0,1]^d$. It suffices to show that

$$\frac{1}{n_0}\sum_{i=1}^{n_0}\int \big(N^1_{\mathcal{F},n}(x,\theta_*,z_i) - N^1_{\mathcal{F}}(x,\theta_*,z_i)\big)^2 dE[\hat{H}_n(x|z_i)] = o(1),$$

which follows by (2.3.5) and Lemma A.2.6 because for any $(x,\theta,z) \in \mathbb{R}\times\Theta\times[0,1]^d$,

$$N^1_{\mathcal{F},n}(x,\theta,z) - N^1_{\mathcal{F}}(x,\theta,z) = N_{\mathcal{F},n}(x,\theta,z) - N_{\mathcal{F}}(x,\theta,z).$$

Thus, $\sigma_{2n} = o_p(1)$ as well.

Note that there exists a constant $C > 0$ such that

$$\|\partial\big(N^1_{\mathcal{F}}\big)^2/\partial z\| \le 2d\cdot\|N^1_{\mathcal{F}}\|\cdot\|\partial N^1_{\mathcal{F}}/\partial z\|$$
$$\le 2d\cdot\|N^1_{\mathcal{F}}\|\cdot\Big(\|\partial N_{\mathcal{F}}/\partial z\| + 2\Big\|C^T_{\mathcal{F}}(\theta_*)\cdot\ddot{L}^{-1}_{f,\infty}(\theta_*)\cdot\frac{\partial^2\log f}{\partial z\partial\theta}\Big\|\Big)$$
$$\le C\cdot\Big(1+\Big\|\frac{\partial\log f}{\partial\theta}\Big\|\Big)\cdot\Big(1+\|\partial F/\partial z\| + \Big\|\frac{\partial^2\log f}{\partial z\partial\theta}\Big\|\Big).$$

Thus, under C5 and C9, $\big\|\partial\big(N^1_{\mathcal{F}}\big)^2/\partial z\big\|$ is dominated by an H integrable function independent of z. Therefore, with $\psi(x,z) = \big(N^1_{\mathcal{F}}(x,\theta_*,z)\big)^2$, Lemma A.2.12 implies $\sigma_{3n} = o_p(1)$.

Consequently, Equality (2.3.9) holds. Analogously, we can show

$$\frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int \hat{N}^1_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)\,d\hat{H}_n(x|z_i)\Big)^2$$
$$= \frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int N^1_{\mathcal{F}}(x,\theta_*,z_i)dH(x|z_i)\Big)^2 + o_p(1).$$

Thus, the assertion follows from the convergence of the Riemann sum. □

For the Model $\mathcal{G}$, let $N^1_{\mathcal{G}}$, $N^1_{\mathcal{G},n}$ and $d_{H,n}(\mathcal{G})$ be defined accordingly. Further denote the variance and its estimator as

$$\sigma^2 := \int\int \left(N^1_{\mathcal{F}}(x,\theta_*,z) - N^1_{\mathcal{G}}(x,\gamma_*,z)\right)^2 dH(x|z)dz - \int\Big(\int N^1_{\mathcal{F}}(x,\theta_*,z) - N^1_{\mathcal{G}}(x,\gamma_*,z)\,dH(x|z)\Big)^2 dz$$

$$\hat{\sigma}^2_n := \frac{1}{n_0}\sum_{i=1}^{n_0}\int \left(\hat{N}^1_{\mathcal{F},n}(x,\hat{\theta}_n,z_i) - \hat{N}^1_{\mathcal{G},n}(x,\hat{\gamma}_n,z_i)\right)^2 d\hat{H}_n(x|z_i) - \frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int \hat{N}^1_{\mathcal{F},n}(x,\hat{\theta}_n,z_i) - \hat{N}^1_{\mathcal{G},n}(x,\hat{\gamma}_n,z_i)\,d\hat{H}_n(x|z_i)\Big)^2.$$

As in Section 2.2, based on Theorem 2.3.5 and Lemma 2.3.6, the asymptotic behavior of the test statistics can be stated as in the following two theorems. The proofs are omitted.

Theorem 2.3.7. *Let* C1–C9 *be satisfied, then we have*

$$\sqrt{n}\cdot T_n - \sqrt{n}\cdot\left(d_{H,n}(\mathcal{F}) - d_{H,n}(\mathcal{G})\right) \xrightarrow{d} \mathcal{N}(0,\sigma^2),$$

$$d_{H,n}(\mathcal{F}) - d_{H,n}(\mathcal{G}) \to d_H(\mathcal{F}) - d_H(\mathcal{G}), \qquad \hat{\sigma}^2_n \to \sigma^2.$$

Theorem 2.3.8. *Let* C1–C9 *be satisfied.*

(1) If $\mathcal{H}^{\mathcal{F}}$ holds, then $\sqrt{n}\cdot T_n$ tends to $-\infty$ in probability.

(2) If $\mathcal{H}^{\mathcal{G}}$ holds, then $\sqrt{n}\cdot T_n$ tends to $+\infty$ in probability.

(3) If $\mathcal{H}^0$ holds and $\sqrt{n}\cdot\left(d_{H,n}(\mathcal{F}) - d_{H,n}(\mathcal{G})\right) = o(1)$, then

$$\sqrt{n}\cdot T_n \xrightarrow{d} \mathcal{N}(0,\sigma^2).$$

It remains the question, whether

$$\sqrt{n}\cdot\left(d_{H,n}(\mathcal{F}) - d_{H,n}(\mathcal{G})\right) = o(1)$$

holds under $\mathcal{H}^0$ in general. Unfortunately, it does not hold except for the case $d = 1$. In the following Lemma, we assume that $d = 1$ ($n_0 = \bar{n}_0$), we will give conditions, under which

$$\sqrt{n}\cdot d_{H,n}(\mathcal{F}) = \sqrt{n}\cdot d_H(\mathcal{F}) + o(1)$$

i.e. $\sqrt{n}\cdot\left(d_{H,n}(\mathcal{F}) - d_{H,n}(\mathcal{G})\right) = o(1)$ holds under $\mathcal{H}^0$.

Lemma 2.3.9. *Assume that $d = 1$, $n_0 h^4 \to 0$ and the function*

$$\int \Big\|\frac{\partial \log f(x|\theta,\cdot)}{\partial\theta}\Big\| \cdot \Big\|\frac{\partial^2 H(x|\cdot)}{\partial z \partial x}\Big\| dx$$

is bounded on $[0,1]$, *then under* C1–C6, C8 and C9 we have

$$\sqrt{n}\cdot d_{H,n}(\mathcal{F}) = \sqrt{n}\cdot d_H(\mathcal{F}) + o(1).$$

Proof. First we have

$$\begin{aligned}
&\sqrt{n}\cdot\big|d_{H,n}(\mathcal{F}) - d_H(\mathcal{F})\big| \\
\leq&\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int\Big|\big(E[\hat{H}_n(x|z_i)] - F(x|\theta_{n_0}, z_i)\big)^2 - \big(H(x|z_i) - F(x|\theta_{n_0}, z_i)\big)^2\Big|\,dE[\hat{H}_n(x|z_i)] \\
&+\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int\Big|\big(H(x|z_i) - F(x|\theta_{n_0}, z_i)\big)^2 - \big(H(x|z_i) - F(x|\theta_*, z_i)\big)^2\Big|\,dE[\hat{H}_n(x|z_i)] \\
&+\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\Big|\int\big(H(x|z_i) - F(x|\theta_*, z_i)\big)^2\,d\big(E[\hat{H}_n(x|z_i)] - H(x|z_i)\big)\Big| \\
&+\sqrt{n}\cdot\Big|\frac{1}{n_0}\sum_{i=1}^{n_0}\int\big(H(x|z_i) - F(x|\theta_*, z_i)\big)^2\,dH(x|z_i) - d_H(\mathcal{F})\Big| \\
=:&d_{1n} + d_{2n} + d_{3n} + d_{4n}.
\end{aligned}$$

By Lemma A.2.6, there exists a $C > 0$ such that,

$$\begin{aligned}
d_{1n} \leq& 4\sqrt{n}\cdot\frac{1}{n}\sum_{i=1}^{n}\int\big|E[\hat{H}_n(x|z_i)] - H(x|z_i)\big|\,dE[\hat{H}_n(x|z_i)] \\
=&4\sqrt{n}\cdot\frac{1}{n}\sum_{z_i\in S_h}\int\big|E[\hat{H}_n(x|z_i)] - H(x|z_i)\big|\,dE[\hat{H}_n(x|z_i)] \\
&+4\sqrt{n}\cdot\frac{1}{n}\sum_{z_i\notin S_h}\int\big|E[\hat{H}_n(x|z_i)] - H(x|z_i)\big|\,dE[\hat{H}_n(x|z_i)] \\
\leq&C\sqrt{n}\cdot\frac{1}{n}\sum_{z_i\in S_h}h^2 + C\sqrt{n}\cdot\frac{1}{n}\sum_{z_i\notin S_h}h \\
\leq&C\sqrt{n}\cdot h^2 + C\sqrt{n}\cdot\frac{1}{n}\cdot(2\lceil n_0 h\rceil + 1)\cdot mh = o(1),
\end{aligned}$$

where the last step follows from the assumption $n_0 h^4 \to 0$. Thus, $d_{1n} = o(1)$.

Under C7, there exists a constant $C > 0$ such that

$$d_{2n} \leq 4\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int\big|F(x|\theta_*, z_i) - F(x|\theta_{n_0}, z_i)\big|\,dE[\hat{H}_n(x|z_i)] \leq C\sqrt{n}\cdot\|\theta_* - \theta_{n_0}\|.$$

In the following, we show that under C1–C6 and C8, $\sqrt{n} \cdot \|\theta_* - \theta_{n_0}\| = o(1)$. By a Taylor expansion, there exists a $\tilde{\theta}_n \in \Theta$ such that

$$\dot{L}_{f,n_0}(\theta_*) = \dot{L}_{f,n_0}(\theta_{n_0}) + \ddot{L}_{f,n_0}(\tilde{\theta}_n) \cdot (\theta_* - \theta_{n_0}).$$

With the same arguments used in Lemma 2.2.3, by Lemma 2.3.2, for eventually all $n \in m \cdot \mathbb{N}$, $\ddot{L}_{f,n_0}(\tilde{\theta}_n)$ is invertible. Further, by the definition of θ_* and θ_{n_0},

$$\dot{L}_{f,\infty}(\theta_*) = \dot{L}_{f,n_0}(\theta_{n_0}) = 0.$$

Thus, we obtain

$$\begin{aligned}\sqrt{n} \cdot (\theta_* - \theta_{n_0}) =& \sqrt{n} \cdot \ddot{L}^{-1}_{f,n_0}(\tilde{\theta}_n) \cdot \big(\dot{L}_{f,n_0}(\theta_*) - \dot{L}_{f,n_0}(\theta_{n_0})\big) \\ =& \sqrt{n} \cdot \ddot{L}^{-1}_{f,n_0}(\tilde{\theta}_n) \cdot \big(\dot{L}_{f,n_0}(\theta_*) - \dot{L}_{f,\infty}(\theta_*)\big).\end{aligned} \tag{2.3.10}$$

Note that for each $i \in \{1, ..., p\}$, the derivative of the function

$$\int \frac{\partial \log f(x|\theta_*, \cdot)}{\partial \theta_i} dH(x|\cdot)$$

with respect to z is

$$\int \frac{\partial \log f(x|\theta_*, \cdot)}{\partial z \partial \theta_i} dH(x|\cdot) + \int \frac{\partial \log f(x|\theta_*, \cdot)}{\partial \theta_i} \cdot \frac{\partial^2 H(x|\cdot)}{\partial z \partial x} dx,$$

which is bounded on $[0, 1]$ under the assumptions of this lemma and C9. Thus, for each $i \in \{1, ..., p\}$ the function

$$\int \frac{\partial \log f(x|\theta_*, \cdot)}{\partial \theta_i} dH(x|\cdot)$$

is Lipschitz continuous on $[0, 1]$. Hence, with

$$\psi(z) = \int \frac{\partial \log f(x|\theta_*, z)}{\partial \theta_i} dH(x|z)$$

Lemma A.2.13 implies

$$\sqrt{n} \cdot \big\|\dot{L}_{f,n_0}(\theta_*) - \dot{L}_{f,\infty}(\theta_*)\big\| = o(1).$$

Further, analogously to (2.2.10), by Lemma 2.3.2 we can show

$$\|\ddot{L}^{-1}_{f,n_0}(\tilde{\theta}_n) - \ddot{L}^{-1}_{f,\infty}(\theta_*)\| = o(1).$$

Therefore, (2.3.10) implies $\sqrt{n} \cdot \|\theta_* - \theta_{n_0}\| = o(1)$, thus, $d_{2n} = o(1)$.

Note further that by partial integration d_{3n} can be written as

$$2\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\Big|\int\big(E[\hat{H}_n(x|z_i)]-H(x|z_i)\big)\big(H(x|z_i)-F(x|\theta_*,z_i)\big)\,d\big(H(x|z_i)-F(x|\theta_*,z_i)\big)\Big|.$$

By similar arguments for d_{1n}, we get $d_{3n}=o(1)$ as well.

For the term d_{4n} note that for each $\theta\in\Theta$, the derivative of the function

$$\int(H-F)^2\,dH$$

with respect to z is

$$\int(H-F)(\frac{\partial H}{\partial z}-\frac{\partial F}{\partial z})\,dH+\int(H-F)^2\frac{\partial^2 H}{\partial z\partial x}\,dx,$$

which is bounded under Assumption (i) and C9. Thus, the function

$$\int\big(H(x|\cdot)-F(x|\theta_*,\cdot)\big)^2\,dH(x|\cdot)$$

is Lipschitz continuous on $[0,1]$. Hence, with

$$\psi(z)=\int\big(H(x|z)-F(x|\theta_*,z)\big)^2\,dH(x|z)$$

Lemma A.2.13 implies $d_{4n}=o(1)$. Therefore, the assertion follows. $\square$

The decision rules of the test for the case $d=1$ can then be formulated as in Section 2.2. For a given significance level α, we will decide for the hypothesis $\mathcal{H}^0$, if $|\sqrt{n}\cdot T_n/\hat{\sigma}_n|\le z_{1-\alpha/2}$, where z_α denotes the α-quantile of a standard normal distribution. In the case of $\sqrt{n}\cdot T_n/\hat{\sigma}_n<-z_{1-\alpha/2}$ we reject $\mathcal{H}^0$ in favor of $\mathcal{H}^{\mathcal{F}}$. If $\sqrt{n}\cdot T_n/\hat{\sigma}_n>z_{1-\alpha/2}$, we reject $\mathcal{H}^0$ in favour of $\mathcal{H}^{\mathcal{G}}$.

For the case $d>1$, the Equality $\sqrt{n}\cdot\|\theta_*-\theta_{n_0}\|=o(1)$ does not hold in general. A one-sided test can then be carried out with

$$\mathcal{H}_a^0:\quad d_H(\mathcal{F})-d_H(\mathcal{G})<a$$

against

$$\mathcal{H}_a^1:\quad d_H(\mathcal{F})-d_H(\mathcal{G})\ge a,$$

where a is a constant. Given a significance level α, we reject the hypothesis $\mathcal{H}_a^0$ in favour of $\mathcal{H}_a^1$, in the case of $\sqrt{n}\cdot(T_n-a)/\hat{\sigma}_n>z_{1-\alpha}$, otherwise $\mathcal{H}_a^0$ will be accepted.

Chapter 3

The Case with Right Censoring

In this chapter, we will extend the results of Chapter 2 to the case with right censoring, i.e. we assume that the random variable X_z at covariate value $z \in [0,1]^d$ is subject to right random censoring. The corresponding censoring random variable is denoted by C_z. The observable random vector at z is then (Y_z, Δ_z), where

$$Y_z := \min(X_z, C_z) \quad \text{and} \quad \Delta_z := I\{X_z \leq C_z\}.$$

Let $H(\cdot|z)$, $J(\cdot|z)$ and $B(\cdot|z)$ denote the distribution functions of X_z, C_z and Y_z, respectively. Assume that X_z and C_z are independent, thus

$$B(\cdot|z) = 1 - \big(1 - H(\cdot|z)\big)\big(1 - J(\cdot|z)\big).$$

Again, let $z_1, ..., z_n$ be the fixed covariate values defined in Section 1.1, at which the random variables are observed. For each covariate value z_i, we write X_i, C_i, Y_i, Δ_i instead of $X_{z_i}, C_{z_i}, Y_{z_i}, \Delta_{z_i}$. The random variables $X_1, ..., X_n, C_1, ..., C_n$ are assumed to be independent. For simplicity of notation let $\tau_{B(\cdot|z)} = \tau_B \in \mathbb{R}$ for any $z \in [0,1]^d$.

The two competing parametric model classes are still denoted as

$$\mathcal{F} := \{F(\cdot|\theta, z) : \theta \in \Theta \subset \mathbb{R}^p, z \in [0,1]^d\}$$

and

$$\mathcal{G} := \{G(\cdot|\gamma, z) : \gamma \in \Gamma \subset \mathbb{R}^q, z \in [0,1]^d\},$$

where Θ and Γ are compact, p and $q \in \mathbb{N}$.

This chapter is organized similarly as Chapter 2. In Section 3.1 we will introduce the basic notations and hypotheses. Amongst other things the discrepancy (distance) between the model class and the underlying distribution functions will be defined. The case with fixed n_0 and m tending to infinity will be shown in Section 3.2. The difference to Section 2.2 lies in the estimation of H by the Kaplan-Meier estimator. Section 3.3 deals with the case fixed m and n_0 tending to infinity, where the conditional Kaplan-Meier estimation is used. For simplicity of notation, we assume that the notations defined in Section 3.1 are valid through out this chapter and the notations defined in Section 3.2 and Section 3.3 are only valid in that particular section.

3.1 Notations and Hypotheses

For each $z \in [0,1]^d$, we denote the subdistribution functions $B^1, B^0 : \mathbb{R} \times [0,1]^d \to [0,1]$ with

$$B^1(x|z) := \mathrm{P}(Y_z \leq x, \Delta_z = 1) = \int_{-\infty}^{x} \big(1 - J(u|z)\big) dH(u|z),$$
$$B^0(x|z) := \mathrm{P}(Y_z \leq x, \Delta_z = 0) = \int_{-\infty}^{x} \big(1 - H(u|z)\big) dJ(u|z).$$

The corresponding empirical distribution functions are given by

$$B^1_n(x|z) := \frac{1}{m} \sum_{i=1}^{n} \Delta_i \cdot \delta_i(z) \cdot I(Y_i \leq x),$$
$$B^0_n(x|z) := \frac{1}{m} \sum_{i=1}^{n} (1 - \Delta_i) \cdot \delta_i(z) \cdot I(Y_i \leq x).$$

Further, the empirical distribution for B is denoted by

$$B_n(x|z) := \frac{1}{m} \sum_{i=1}^{n} \delta_i(z) I(Y_i \leq x).$$

Note that we have the relations

$$B(x|z) = B^1(x|z) + B^0(x|z),$$

$$B_n(x|z) = B^1_n(x|z) + B^0_n(x|z).$$

Denote again the joint distribution by $Q : \mathbb{R} \times [0,1]^d \to [0,1]$ with

$$Q(x,z) := \int \int I(u \leq x) \cdot I(v \leq z) dH(u|v) dv,$$

where the inner integration is with respect to the variable u. In addition we denote for $r \in \{0,1\}$ the joint distribution functions $Q_n^r, Q^r : \mathbb{R}\times[0,1]^d \to [0,1]$ with

$$Q_n^r(x,z) := \frac{1}{n_0}\sum_{i=1}^{n_0} B_n^r(x|z_i)\cdot I(z_i \le z)$$
$$Q^r(x,z) := \int\int I(u \le x)\cdot I(v \le z) dB^r(u|v)dv,$$

where the inner integration is with respect to the variable u. For any function $\psi : \mathbb{R}\times[0,1]^d \to \mathbb{R}$ and $r \in \{0,1\}$ we have as in Chapter 2 that

$$\int \psi(x,z)dQ_n^r(x,z) = \frac{1}{n_0}\sum_{i=1}^{n_0}\int \psi(x,z_i)dB_n^r(x|z_i).$$

$$\int \psi(x,z)dQ^r(x,z) = \int\int \psi(x,z)dB^r(x|z)dz.$$

Denote further the functions $\gamma, C : (-\infty, \tau_B]\times[0,1]^d \to \mathbb{R}$ with

$$\gamma(x|z) := \exp\Big(\int_{-\infty}^{x-} \frac{1}{1-B(u|z)} dB^0(u|z)\Big),$$
$$C(x|z) := \int_{-\infty}^{x-} \frac{1}{\big(1-B(u|z)\big)^2} dB^0(u|z).$$

Note that

$$\gamma(x|z) = \exp\Big(\int_{-\infty}^{x-} \frac{1-H(u|z)}{1-B(u|z)} dJ(u|z)\Big) = \exp\Big(\int_{-\infty}^{x-} \frac{1}{1-J(u|z)} dJ(u|z)\Big)$$
$$= \exp\Big(-\log\big(1-J(x-|z)\big)\Big) = \frac{1}{1-J(x-|z)}.$$

Let $\tau < \tau_B$ be a constant. In order to ignore the tail effect of the Kaplan-Meier estimator later, we define the logarithmic likelihood function for the model class $\mathcal{F}$ as function $\hat{L}_{f,n} : \Theta \to \mathbb{R}$ with

$$\hat{L}_{f,n}(\theta) := \frac{1}{n}\sum_{i=1}^{n}\Big(\Delta_i \log f(Y_i|\theta, z_i) + (1-\Delta_i)\log\big(1 - F(Y_i|\theta, z_i)\big)\Big)\cdot I(Y_i \le \tau)$$
$$= \int \log f(x|\theta,z)\cdot I(x\le\tau)dQ_n^1(x,z) + \int \log\big(1-F(x|\theta,z)\big)\cdot I(x\le\tau)dQ_n^0(x,z).$$

The maximum likelihood estimator of θ is then defined as

$$\hat{\theta}_n := \underset{\theta\in\Theta}{\operatorname{argmax}}\, \hat{L}_{f,n}(\theta).$$

Further we define functions $L_{f,n_0}, L_{f,\infty} : \Theta \to \mathbb{R}$ and the vectors $\theta_{n_0}, \theta_* \in \Theta$ with

$$\begin{aligned} L_{f,n_0}(\theta) &:= \frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int_{-\infty}^{\tau} \log f(x|\theta, z_i)dB^1(x|z_i) + \int_{-\infty}^{\tau} \log\big(1 - F(x|\theta, z_i)\big)dB^0(x|z_i)\Big), \\ \theta_{n_0} &:= \operatorname*{argmax}_{\theta\in\Theta} L_{f,n_0}(\theta), \\ L_{f,\infty}(\theta) &:= \int \log f(x|\theta, z)\cdot I(x \le \tau)dQ^1(x, z) + \int \log\big(1 - F(x|\theta, z)\big)\cdot I(x \le \tau)dQ^0(x, z), \\ \theta_* &:= \operatorname*{argmax}_{\theta\in\Theta} L_{f,\infty}(\theta). \end{aligned}$$

Analogously to the case without censoring, the following asymptotic relations hold:

$$\begin{array}{ccc} \hat{L}_{f,n} & \xrightarrow{m\to\infty} & L_{f,n_0} \\ {\scriptstyle n_0\to\infty}\searrow & & \swarrow{\scriptstyle n_0\to\infty} \\ & L_{f,\infty}. & \end{array}$$

Under some regularity conditions, we obtain

$$\begin{array}{ccc} \hat{\theta}_n & \xrightarrow{m\to\infty} & \theta_{n_0} \\ {\scriptstyle n_0\to\infty}\searrow & & \swarrow{\scriptstyle n_0\to\infty} \\ & \theta_*. & \end{array}$$

The distance $d_H(\mathcal{F})$ between the underlying family of distribution functions H and the model class $\mathcal{F}$ is defined as

$$\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\big(H(x|z_i) - F(x|\theta_{n_0}, z_i)\big)^2\, dH(x|z_i), \tag{3.1.1}$$

for the case with n_0 fixed and $m \to \infty$ and

$$\int \big(H(x|z) - F(x|\theta_*, z)\big)^2 \cdot I(x \le \tau)\, dQ(x, z) \tag{3.1.2}$$

for the case with m fixed and $n_0 \to \infty$, respectively. Again, let $\hat{\gamma}_n$, γ_{n_0}, γ_* and $d_H(\mathcal{G})$ denote the counterpart for the model class $\mathcal{G}$. We will propose model selection tests with the null hypothesis

$$\mathcal{H}^0: \quad d_H(\mathcal{F}) = d_H(\mathcal{G})$$

meaning that the two models are equally close to H, against

$$\mathcal{H}^{\mathcal{F}}: \quad d_H(\mathcal{F}) < d_H(\mathcal{G})$$

meaning H is closer to $\mathcal{F}$ than to $\mathcal{G}$ or

$$\mathcal{H}^{\mathcal{G}}: \quad d_H(\mathcal{F}) > d_H(\mathcal{G})$$

meaning H is closer to $\mathcal{G}$ than to $\mathcal{F}$.

In this chapter, it is assumed that the integrability of a function defined in Chapter 2 holds for x on $(-\infty, \tau]$. For instance, for any function $\psi : \mathbb{R} \times \Theta \times [0,1]^d \to \mathbb{R}$ with $(x, \theta, z) \mapsto \psi(x, \theta, z)$, we refer it as dominated by a B^1 integrable function, if for each $z \in [0,1]^d$, there exists a function $M(\cdot, z) : \mathbb{R} \to \mathbb{R}$, such that $|\psi(x, \theta, z)| \leq M(x, z)$ for all $(x, \theta) \in (-\infty, \tau] \times \Theta$ and $\int_{-\infty}^{\tau} M(x, z) dB^1(x|z) < \infty$.

It is assumed that all the convergences are taken by letting $m \to \infty$ in Section 3.2 and all the convergences are taken by letting $n_0 \to \infty$ in Section 3.3. Note that since $n = m \cdot n_0$, in both cases $n \to \infty$.

3.2 The Case with Number of Observations at Each Covariate Tending to Infinity

In this section, for each $z \in [0,1]^d$ the distribution function $H(\cdot|z)$ is estimated by Kaplan-Meier estimator, which in our setting is defined by

$$H_n^{KM}(x|z) := 1 - \prod_{Y_{(i)} \leq x} \left(1 - \frac{\delta_{(i)}(z) \cdot \Delta_{(i)}}{m - \sum_{j=1}^{i-1} \delta_{(j)}(z)}\right),$$

where $Y_{(1)} \leq Y_{(2)} \leq ... \leq Y_{(n)}$ are the ordered observations and $\Delta_{(i)}$ and $\delta_{(i)}(z)$ are the corresponding indicators for $Y_{(i)}$. Let the joint empirical distribution function be defined as

$$Q_n^{KM}(x, z) := \frac{1}{n_0} \sum_{i=1}^{n_0} H_n^{KM}(x|z_i) \cdot I(z_i \leq z).$$

The distance $d_H(\mathcal{F})$ defined in (3.1.1) can then be estimated by

$$\hat{d}_{H,n}(\mathcal{F}) := \int \left(H_n^{KM}(x|z) - F(x|\hat{\theta}_n, z)\right)^2 \cdot I(x \leq \tau)\, dQ_n^{KM}(x, z).$$

For the class $\mathcal{G}$, the estimator $\hat{d}_{H,n}(\mathcal{G})$ is defined in an analogous way. Again the test statistic is defined as the difference of the estimated distances

$$T_n := \hat{d}_{H,n}(\mathcal{F}) - \hat{d}_{H,n}(\mathcal{G}).$$

The assumptions are formulated as follows. They are extensions of the assumptions in Section 2.2 to the setting in this section.

D1 For each $(\theta, z) \in \Theta \times [0,1]^d$, the distribution $F(\cdot|\theta, z)$ has a density function $f(\cdot|\theta, z) : \mathbb{R} \to \mathbb{R}$. The functions $f(\cdot|\theta, z)$ and $1 - F(\cdot|\theta, z)$ are strictly positive $B^1(\cdot|z)$- and $B^0(\cdot|z)$-a.s. respectively.

D2 The functions $\log f$ and $\log(1 - F)$ are three times continuously differentiable in θ on Θ.

D3 The function $\log f$ is dominated by a B^1 integrable function. The function $\log(1 - F)$ is dominated by a B^0 integrable function.

D4 The function L_{f,n_0} has a unique maximum on Θ at θ_{n_0}, which is an interior point of Θ.

D5 The functions $\|\partial \log f/\partial\theta\|$ and $\|\partial^2 \log f/\partial\theta^2\|$ are dominated by B^1 square integrable functions. The functions $\|\partial \log(1-F)/\partial\theta\|$ and $\|\partial^2 \log(1-F)/\partial\theta^2\|$ are dominated by B^0 square integrable functions. The Hessian matrix $\ddot{L}_{f,n_0}(\theta_{n_0})$ is invertible with inverse $\ddot{L}^{-1}_{f,n_0}(\theta_{n_0})$.

D6 For any $i,\ j,\ k \in \{1, 2, ..., p\}$, the function $\partial^3 \log f/\partial\theta_i\partial\theta_j\partial\theta_k$ is dominated by a B^1 integrable function. The function $\partial^3 \log(1-F)/\partial\theta_i\partial\theta_j\partial\theta_k$ is dominated by a B^0 integrable function.

D7 The functions $\dot{F}$ and $\ddot{F}$ exists and they are bounded.

Analogously to Lemma 2.2.1, we can show the following lemma.

Lemma 3.2.1. *For $r \in \{0, 1\}$, define the function $\psi_r : \mathbb{R} \times [0,1]^d \to 0$. If ψ_r is B^r integrable, then*

$$\sum_{r=0}^{1} \Big(\int \psi_r(x,z) dQ_n^r(x,z) - \frac{1}{n_0} \sum_{i=1}^{n_0} \int \psi_r(x, z_i) dB^r(x|z_i) \Big) \xrightarrow{a.s.} 0.$$

If ψ_r is B^r square integrable, then

$$\sqrt{n}\cdot\sum_{r=0}^{1}\Big(\int\psi_r(x,z)dQ_n^r(x,z)-\frac{1}{n_0}\sum_{i=1}^{n_0}\int\psi_r(x,z_i)dB^r(x|z_i)\Big)\xrightarrow{d}\mathcal{N}(0,\sigma^2),$$

where

$$\sigma^2=\frac{1}{n_0}\sum_{i=1}^{n_0}\sum_{r=0}^{1}\Big(\int\psi_r^2(x,z_i)dB^r(x|z_i)-\Big(\int\psi_r(x,z_i)dB^r(x|z_i)\Big)^2\Big).$$

Proof. We denote first for each $i\in\{1,...,m\}$,

$$U_i:=\frac{1}{n_0}\sum_{j=1}^{n_0}\Big(\Delta_{(i-1)\cdot n_0+j}\cdot\psi_1(Y_{(i-1)\cdot n_0+j},z_j)+\big(1-\Delta_{(i-1)\cdot n_0+j}\big)\cdot\psi_0(Y_{(i-1)\cdot n_0+j},z_j)\Big).$$

Notice that U_1,..., U_m are i.i.d. and

$$\sum_{r=0}^{1}\int\psi_r(x,z)dQ_n^r(x,z)=\frac{1}{n}\sum_{i=1}^{n}\Big(\Delta_i\psi_1(Y_i,z_i)+\big(1-\Delta_i\big)\psi_0(Y_i,z_i)\Big)=\frac{1}{m}\sum_{i=1}^{m}U_i.$$

The expectation

$$\begin{aligned}E\Big[\sum_{r=0}^{1}\int\psi_r(x,z)dQ_n^r(x,z)\Big]=&\frac{1}{n}\sum_{i=1}^{n}E\Big[\Delta_i\psi_1(Y_i,z_i)+\big(1-\Delta_i\big)\psi_0(Y_i,z_i)\Big]\\=&\frac{1}{n}\sum_{i=1}^{n}\sum_{r=0}^{1}\int\psi_r(x,z_i)dB^r(x|z_i)\\=&\frac{1}{n_0}\sum_{i=1}^{n_0}\sum_{r=0}^{1}\int\psi_r(x,z_i)dB^r(x|z_i).\end{aligned}$$

Because $(Y_1,\Delta_1),...,(Y_n,\Delta_n)$ are independent, we get further

$$\begin{aligned}&Var\Big[\sqrt{n}\cdot\sum_{r=0}^{1}\int\psi_r(x,z)dQ_n^r(x,z)\Big]\\=&\frac{1}{n}\sum_{i=1}^{n}Var\Big[\Delta_i\psi_1(Y_i,z_i)+\big(1-\Delta_i\big)\psi_0(Y_i,z_i)\Big]=\sigma^2.\end{aligned}$$

Thus, the assertions follow from the strong law of large numbers and central limit theorem for i.i.d data.

□

Based on Lemma 3.2.1, the following two lemmas on the convergence of $\hat{\theta}_n$ can be shown as the case without censoring. The proofs are omitted.

Lemma 3.2.2. *If* D1–D5 *hold, we have* $\|\hat{\theta}_n - \theta_{n_0}\| \to 0$ *a.s.*

Lemma 3.2.3. *If* D1–D6 *holds, then*

$$\sqrt{n} \cdot \|\hat{\theta}_n - \theta_{n_0}\| = O_p(1),$$
$$\sqrt{n} \cdot \left\|\hat{\theta}_n - \theta_{n_0} + \ddot{L}^{-1}_{f,n_0}(\theta_{n_0}) \cdot \dot{\hat{L}}_{f,n}(\theta_{n_0})\right\| = o_p(1).$$

In order to state the main theorem we introduce the functions $C_{\mathcal{F}} : \Theta \to \mathbb{R}^p$, and $N_{\mathcal{F}} : (-\infty, \tau_B] \times \Theta \times [0,1]^d \to \mathbb{R}$ with

$$\begin{aligned} C_{\mathcal{F}}(\theta) :=& \frac{1}{n_0} \sum_{i=1}^{n_0} \int_{-\infty}^{\tau} \big(H(x|z_i) - F(x|\theta, z_i)\big) \cdot \dot{F}(x|\theta, z_i)\, dH(x|z_i), \\ N_{\mathcal{F}}(x,\theta,z) :=& \big(H(x|z) - F(x|\theta,z)\big)^2 \cdot I(x \le \tau) \\ &+ 2\int_x^{\tau} \big(H(u|z) - F(u|\theta,z)\big)\, dH(u|z) \cdot I(x \le \tau). \end{aligned}$$

Further we denote $N^1_{\mathcal{F}}, N^0_{\mathcal{F}} : (-\infty, \tau_B] \times \Theta \times [0,1]^d \to \mathbb{R}$ with

$$\begin{aligned} N^1_{\mathcal{F}}(x,\theta,z) :=& N_{\mathcal{F}}(x,\theta,z)\gamma(x|z) \\ &+ 2C^T_{\mathcal{F}}(\theta_{n_0}) \cdot \ddot{L}^{-1}_{f,n_0}(\theta_{n_0}) \cdot \frac{\partial \log f(x|\theta,z)}{\partial\theta} \cdot I(x \le \tau) \\ &- \int N_{\mathcal{F}}(u,\theta,z) C(x \wedge u|z) dH(u|z), \\ N^0_{\mathcal{F}}(x,\theta,z) :=& \frac{1}{1 - B(x|z)} \int N_{\mathcal{F}}(u,\theta,z) \cdot I(x < u) dH(u|z) \\ &+ 2C^T_{\mathcal{F}}(\theta_{n_0}) \cdot \ddot{L}^{-1}_{f,n_0}(\theta_{n_0}) \cdot \frac{\partial \log\big(1 - F(x|\theta,z)\big)}{\partial\theta} \cdot I(x \le \tau) \\ &- \int N_{\mathcal{F}}(u,\theta,z) C(x \wedge u|z) dH(u|z) \end{aligned}$$

and the constant $\sigma^2_{\mathcal{F}} \in \mathbb{R}$,

$$\begin{aligned} \sigma^2_{\mathcal{F}} :=& \frac{1}{n_0} \sum_{i=1}^{n_0} \Big(\int N^1_{\mathcal{F}}(x,\theta_{n_0},z_i)^2\, dB^1(x|z_i) - \big(\int N^1_{\mathcal{F}}(x,\theta_{n_0},z_i)\, dB^1(x|z_i)\big)^2 \Big) \\ &+ \frac{1}{n_0} \sum_{i=1}^{n_0} \Big(\int N^0_{\mathcal{F}}(x,\theta_{n_0},z_i)^2\, dB^0(x|z_i) - \big(\int N^0_{\mathcal{F}}(x,\theta_{n_0},z_i)\, dB^0(x|z_i)\big)^2 \Big). \end{aligned}$$

Theorem 3.2.4. *Let* D1–D7 *be satisfied, then*

$$\sqrt{n} \cdot \big(\hat{d}_{H,n}(\mathcal{F}) - d_H(\mathcal{F})\big) \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\mathcal{F}}).$$

Proof. Analogously to Theorem 2.2.4, under D1–D7, by Lemma 3.2.3, Lemma A.3.2 and A.3.3, we can write

$$\begin{aligned}\sqrt{n}\cdot\hat{d}_{H,n}(\mathcal{F}) =& \sqrt{n}\cdot\int N_{\mathcal{F}}(x,\theta_{n_0},z)\,dQ_n^{KM}(x,z)\\ &- 2\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\int_x^{\tau}\big(H(u|z_i)-F(u|\theta,z_i)\big)\,dH(u|z_i)\,dH(x|z_i)\\ &+ 2\sqrt{n}\cdot C_{\mathcal{F}}^T(\theta_{n_0})\cdot\ddot{L}_{f,n_0}^{-1}(\theta_{n_0})\cdot\dot{\hat{L}}_{f,n}(\theta_{n_0})+o_p(1).\end{aligned}$$

By Lemma A.3.1, we get then

$$\begin{aligned}\sqrt{n}\cdot\hat{d}_{H,n}(\mathcal{F}) =& \sqrt{n}\cdot\int N_{\mathcal{F}}^1(x,\theta_{n_0},z)dQ_n^1(x,z)+\int N_{\mathcal{F}}^0(x,\theta_{n_0},z)dQ_n^0(x,z)\\ &- 2\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\int_x^{\tau}\big(H(u|z_i)-F(u|\theta,z_i)\big)\,dH(u|z_i)\,dH(x|z_i)+o_p(1).\end{aligned}$$

Notice that by the definition of θ_{n_0},

$$\begin{aligned}&\frac{1}{n_0}\sum_{i=1}^{n_0}\int C_{\mathcal{F}}^T(\theta_{n_0})\cdot\ddot{L}_{f,n_0}^{-1}(\theta_{n_0})\cdot\frac{\partial\log f(x|\theta_{n_0},z_i)}{\partial\theta}\cdot I(x\le\tau)dB^1(x|z_i)\\ &+\frac{1}{n_0}\sum_{i=1}^{n_0}\int C_{\mathcal{F}}^T(\theta_{n_0})\cdot\ddot{L}_{f,n_0}^{-1}(\theta_{n_0})\cdot\frac{\partial\log\big(1-F(x|\theta_{n_0},z_i)\big)}{\partial\theta}\cdot I(x\le\tau)dB^0(x|z_i)\\ =&C_{\mathcal{F}}^T(\theta_{n_0})\cdot\ddot{L}_{f,n_0}^{-1}(\theta_{n_0})\cdot\dot{L}_{f,n_0}(\theta_{n_0})=0.\end{aligned}$$

Further for each $z\in[0,1]^d$,

$$\begin{aligned}&\int\int N_{\mathcal{F}}(u,\theta_{n_0},z)C(x\wedge u|z)dH(u|z)dB(x|z)\\ =&\int\int\int\frac{N_{\mathcal{F}}(u,\theta_{n_0},z)I(t<x)I(t<u)}{\big(1-B(t|z)\big)^2}dB^0(t|z)dH(u|z)dB(x|z)\\ =&\int\int\int\frac{N_{\mathcal{F}}(u,\theta_{n_0},z)I(t<x)I(t<u)}{\big(1-B(t|z)\big)^2}dB(x|z)dB^0(t|z)dH(u|z)\\ =&\int\int\frac{N_{\mathcal{F}}(u,\theta_{n_0},z)I(t<u)\big(1-B(t|z)\big)}{\big(1-B(t|z)\big)^2}dB^0(t|z)dH(u|z)\\ =&\int\int\frac{N_{\mathcal{F}}(u,\theta_{n_0},z)I(t<u)}{1-B(t|z)}dB^0(t|z)dH(u|z)\\ =&\int\int\frac{N_{\mathcal{F}}(u,\theta_{n_0},z)\cdot I(x<u)}{1-B(x|z)}dH(u|z)dB^0(x|z) \qquad (3.2.1)\end{aligned}$$

Hence,

$$\frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int N_{\mathcal{F}}^1(x,\theta_{n_0},z_i)dB^1(x|z_i)+\int N_{\mathcal{F}}^0(x,\theta_{n_0},z_i)dB^0(x|z_i)\Big)$$

$$
\begin{aligned}
=&\frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int N_{\mathcal{F}}(x,\theta_{n_0},z_i)\gamma(x|z_i)dB^1(x|z_i)\\
&+\int\int \frac{N_{\mathcal{F}}(u,\theta_{n_0},z_i)\cdot I(x<u)}{1-B(x|z_i)}dH(u|z_i)dB^0(x|z_i)\\
&\quad-\int\int N_{\mathcal{F}}(u,\theta_{n_0},z_i)C(x\wedge u|z_i)dH(u|z_i)dB(x|z_i)\Big)\\
=&\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F}}(x,\theta_{n_0},z_i)dH(x|z_i)\\
=&2\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\int_{x}^{\tau}\big(H(u|z_i)-F(u|\theta_{n_0},z_i)\big)\,dH(u|z_i)\,dH(x|z_i)+d_H(\mathcal{F}).
\end{aligned}
$$

Consequently, we get

$$
\begin{aligned}
d_H(\mathcal{F})=&\frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int N^1_{\mathcal{F}}(x,\theta_{n_0},z_i)dB^1(x|z_i)+\int N^0_{\mathcal{F}}(x,\theta_{n_0},z_i)dB^0(x|z_i)\Big)\\
&-2\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\int_{x}^{\tau}\big(H(u|z_i)-F(u|\theta_{n_0},z_i)\big)\,dH(u|z_i)\,dH(x|z_i).
\end{aligned}
$$

Then, we can write

$$
\begin{aligned}
&\sqrt{n}\cdot\big(\hat{d}_{H,n}(\mathcal{F})-d_H(\mathcal{F})\big)\\
=&\sqrt{n}\cdot\int N^1_{\mathcal{F}}(x,\theta_{n_0},z)d\big(Q^1_n(x,z)-E[Q^1_n(x,z)]\big)\\
&+\sqrt{n}\cdot\int N^0_{\mathcal{F}}(x,\theta_{n_0},z)d\big(Q^0_n(x,z)-E[Q^0_n(x,z)]\big)+o_p(1).
\end{aligned}
$$

Note that under D5 and D7 the functions $N^1_{\mathcal{F}}(\cdot,\theta_{n_0},\cdot)$ and $N^0_{\mathcal{F}}(\cdot,\theta_{n_0},\cdot)$ are B^1 and B^0 square integrable, receptively. Thus, the assertion follows from Lemma 3.2.1 with $\psi_1(x,z)=N^1_{\mathcal{F}}(x,\theta_{n_0},z)$ and $\psi_0(x,z)=N^0_{\mathcal{F}}(x,\theta_{n_0},z)$. □

Notice that in the case without censoring ($J=B^0=0$) the variance reduces to the variance defined in Theorem 2.2.4.

For the estimation of the variance $\sigma^2_{\mathcal{F}}$ we denote for each $n\in n_0\cdot\mathbb{N}$ the functions $\gamma_n, C_n:(-\infty,\tau_B]\times[0,1]^d\to\mathbb{R}$, $C_{\mathcal{F},n}:\Theta\to\mathbb{R}^p$, $N_{\mathcal{F},n}:\mathbb{R}\times\Theta\times[0,1]^d\to\mathbb{R}$, $N^1_{\mathcal{F},n},N^0_{\mathcal{F},n}:(-\infty,\tau_B]\times\Theta\times[0,1]^d\to\mathbb{R}$ with

$$
\begin{aligned}
\gamma_n(x|z):=&\exp\Big(\int_{-\infty}^{x-}\frac{dB^0_n(u|z)}{1-B_n(u|z)}\Big),\\
C_n(x|z):=&\int_{-\infty}^{x-}\frac{1}{\big(1-B_n(u|z)\big)^2}dB^0_n(u|z),
\end{aligned}
$$

$$
\begin{aligned}
C_{\mathcal{F},n}(\theta) :=& \int \Big(\int_{-\infty}^{x} \gamma_n(u|z) dB_n^1(u|z) - F(x|\theta, z) \Big) \dot{F}(x|\theta, z) \gamma_n(x|z) \cdot I(x \le \tau) dQ_n^1(x, z), \\
N_{\mathcal{F},n}(x, \theta, z) :=& \Big(\int_{-\infty}^{x} \gamma_n(u|z) dB_n^1(u|z) - F(x|\theta, z) \Big)^2 \cdot I(x \le \tau) \\
&+ 2 \int_x^{\tau} \Big(\int_{-\infty}^{u} \gamma_n(t|z) dB_n^1(t|z) - F(u|\theta, z) \Big) \gamma_n(u|z) dB_n^1(u|z) \cdot I(x \le \tau), \\
N_{\mathcal{F},n}^1(x, \theta, z) :=& N_{\mathcal{F},n}(x, \theta, z) \gamma_n(x|z) + 2 C_{\mathcal{F},n}^T(\hat{\theta}_n) \cdot \ddot{\hat{L}}_{f,n}^{-1}(\hat{\theta}_n) \cdot \frac{\partial \log f(x|\theta, z)}{\partial \theta} \cdot I(x \le \tau) \\
&- \int N_{\mathcal{F},n}(u, z) C_n(x \wedge u|z) \gamma_n(u|z) dB_n^1(u|z), \\
N_{\mathcal{F},n}^0(x, \theta, z) :=& \frac{1}{1 - B_n(x|z)} \int N_{\mathcal{F},n}(u, z) \cdot I(x < u) \gamma_n(u|z) dB_n^1(u|z) \\
&+ 2 C_{\mathcal{F},n}^T(\hat{\theta}_n) \cdot \ddot{\hat{L}}_{f,n}^{-1}(\hat{\theta}_n) \cdot \frac{\partial \log \big(1 - F(x|\theta, z)\big)}{\partial \theta} \cdot I(x \le \tau) \\
&- \int N_{\mathcal{F},n}(u, z) C_n(x \wedge u|z) \gamma_n(u|z) dB_n^1(u|z).
\end{aligned}
$$

In the next lemma, we show that $\sigma_{\mathcal{F}}^2$ can be estimated consistently by

$$
\begin{aligned}
\hat{\sigma}_{\mathcal{F},n}^2 :=& \frac{1}{n_0} \sum_{i=1}^{n_0} \Big(\int N_{\mathcal{F},n}^1(x, \hat{\theta}_n, z_i)^2 \, dB_n^1(x|z_i) - \big(\int N_{\mathcal{F},n}^1(x, \hat{\theta}_n, z_i) \, dB_n^1(x|z_i) \big)^2 \Big) \\
&+ \frac{1}{n_0} \sum_{i=1}^{n_0} \Big(\int N_{\mathcal{F},n}^0(x, \hat{\theta}_n, z_i)^2 \, dB_n^0(x|z_i) - \big(\int N_{\mathcal{F},n}^0(x, \hat{\theta}_n, z_i) \, dB_n^0(x|z_i) \big)^2 \Big).
\end{aligned}
$$

Lemma 3.2.5. *If* D1–D7 *hold, then we have*

$$
\hat{\sigma}_{\mathcal{F},n}^2 = \sigma_{\mathcal{F}}^2 + o_p(1).
$$

Proof. Note that $N_{\mathcal{F},n}(x, \hat{\theta}_n, z_i) \gamma_n(x|z_i)$ is one of the summands in $N_{\mathcal{F},n}^1(x, \hat{\theta}_n, z_i)$. Hence,

$$
\frac{1}{n_0} \sum_{i=1}^{n_0} \int N_{\mathcal{F},n}^2(x, \hat{\theta}_n, z_i) \gamma_n^2(x|z_i) \, dB_n^1(x|z_i)
$$

is one of the summands in $\hat{\sigma}_{\mathcal{F},n}^2$ and its corresponding part in $\sigma_{\mathcal{F}}^2$ is

$$
\frac{1}{n_0} \sum_{i=1}^{n_0} \int N_{\mathcal{F}}^2(x, \theta_{n_0}, z_i) \gamma^2(x|z_i) \, dB^1(x|z_i).
$$

In the following we will show that

$$
\frac{1}{n_0} \sum_{i=1}^{n_0} \int N_{\mathcal{F},n}^2(x, \hat{\theta}_n, z_i) \gamma_n^2(x|z_i) \, dB_n^1(x|z_i)
$$

$$=\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F}}^2(x,\theta_{n_0},z_i)\gamma^2(x|z_i)\,dB^1(x|z_i)+o_p(1). \tag{3.2.2}$$

First we write

$$\begin{aligned}
&\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F},n}^2(x,\hat\theta_n,z_i)\gamma_n^2(x|z_i)\,dB_n^1(x|z_i)\\
&\qquad-\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F}}^2(x,\theta_{n_0},z_i)\gamma^2(x|z_i)\,dB^1(x|z_i)\\
=&\Big(\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F},n}^2(x,\hat\theta_n,z_i)\gamma_n^2(x|z_i)\,dB_n^1(x|z_i)\\
&\qquad-\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F},n}^2(x,\theta_{n_0},z_i)\gamma_n^2(x|z_i)\,dB_n^1(x|z_i)\Big)\\
&+\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F},n}^2(x,\theta_{n_0},z_i)\big(\gamma_n^2(x|z_i)-\gamma^2(x|z_i)\big)\,dB_n^1(x|z_i)\\
&+\Big(\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F},n}^2(x,\theta_{n_0},z_i)\gamma^2(x|z_i)\,dB_n^1(x|z_i)\\
&\qquad-\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F}}^2(x,\theta_{n_0},z_i)\gamma^2(x|z_i)\,dB^1(x|z_i)\Big).
\end{aligned}\tag{3.2.3}$$

Note that for any $(x,z)\in(-\infty,\tau]\times[0,1]^d$

$$\gamma_n(x|z)=\exp\Big(\int_{-\infty}^{x-}\frac{1}{1-B_n(u|z)}dB_n^0(u|z)\Big)\le\exp\Big(\int_{-\infty}^{x-}\frac{1}{1-B_n(\tau|z)}dB_n^0(u|z)\Big).$$

Since n_0 is fixed, by the Glivenko-Cantelli theorem, $1-B_n(\tau|\cdot)$ is stochastically bounded away from zero uniformly on $[0,1]^d$. Thus, the function γ_n is bounded in probability. Analogously, the function γ is bounded as well. Under D7, we have then the derivative of

$$\frac{1}{n_0}\sum_{i=1}^{n_0}\int N_{\mathcal{F},n}^2(x,\cdot,z_i)\gamma_n^2(x|z_i)\,dB_n^1(x|z_i)$$

is stochastically bounded on Θ. Thus, under D1–D6, it follows from Lemma 3.2.2 that the first term on the right-hand side of (3.2.3) is equal to $o_p(1)$.

In the sequel, we show that the second term on the right-hand side of (3.2.3) is equal to $o_p(1)$ as well. Note that by definition $N_{\mathcal{F},n}^2$ is bounded. Hence, it suffices to show that

$$\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\left|\gamma_n^2(x|z_i)-\gamma^2(x|z_i)\right|dB_n^1(x|z_i)=o_p(1).$$

By Cauchy-Schwarz's inequality,

$$\begin{aligned}
&\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\left|\gamma_n^2(x|z_i)-\gamma^2(x|z_i)\right|dB_n^1(x|z_i)\\
=&\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\left|\gamma_n(x|z_i)-\gamma(x|z_i)\right|\cdot\left|\gamma_n(x|z_i)+\gamma(x|z_i)\right|dB_n^1(x|z_i)\\
\leq&\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\big(\gamma_n(x|z_i)-\gamma(x|z_i)\big)^2dB_n^1(x|z_i)\cdot\int_{-\infty}^{\tau}\big(\gamma_n(x|z_i)+\gamma(x|z_i)\big)^2dB_n^1(x|z_i).
\end{aligned}$$

Since γ_n and γ are stochastically bounded on $[-\infty,\tau]\times[0,1]^d$, thus

$$\int_{-\infty}^{\tau}\big(\gamma_n(x|\cdot)+\gamma(x|\cdot)\big)^2dB_n^1(x|\cdot)$$

is stochastically bounded on $[0,1]^d$. It remains to bound

$$\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\big(\gamma_n(x|z_i)-\gamma(x|z_i)\big)^2dB_n^1(x|z_i).$$

Note that by a Taylor expansion,

$$\gamma_n(x|z)-\gamma(x|z)=\tilde{\gamma}(x|z)\Big(\int_{-\infty}^{x-}\frac{1}{1-B_n(u|z)}dB_n^0(u|z)-\int_{-\infty}^{x-}\frac{1}{1-B(u|z)}dB^0(u|z)\Big)$$

where $\tilde{\gamma}_n(x|z)$ lies between $\gamma_n(x|z)$ and $\gamma(x|z)$. Let $C_{in}(z)$ be defined as in the proof of Lemma A.3.2, we can write then

$$\begin{aligned}
&\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\big(\gamma_n(x|z_i)-\gamma(x|z_i)\big)^2dB_n^1(x|z_i)\\
=&\frac{1}{n_0}\sum_{i=1}^{n_0}\sum_{j=1}^{n}\tilde{\gamma}^2(Y_j|z_i)\delta_j(z_i)\Delta_iI(Y_i\leq\tau)C_{jn}^2(z_i).
\end{aligned}\tag{3.2.4}$$

With similar arguments as in the proof of Lemma A.3.2, it can be shown that the last term equals to $o_p(1)$. Hence, the second term on the right-hand side of (3.2.3) is equal to $o_p(1)$.

By Cauchy-Schwarz's inequality and the boundedness of $N_{\mathcal{F},n}$, $N_{\mathcal{F}}$, and γ, there exists a constant $C>0$ such that the third term on the right-hand side of (3.2.3) is stochastically bounded by

$$C\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int\big(N_{\mathcal{F},n}(x,\theta_{n_0},z_i)-N_{\mathcal{F}}(x,\theta_{n_0},z_i)\big)^2dB_n^1(x|z_i).$$

Next we show that

$$\frac{1}{n_0}\sum_{i=1}^{n_0}\int\big(N_{\mathcal{F},n}(x,\theta_{n_0},z_i)-N_{\mathcal{F}}(x,\theta_{n_0},z_i)\big)^2\,dB_n^1(x|z_i)=o_p(1).$$

Note that there exists a constant $C>0$ such that

$$\begin{aligned}
&\frac{1}{n_0}\sum_{i=1}^{n_0}\int\big(N_{\mathcal{F},n}(x,\theta_{n_0},z_i)-N_{\mathcal{F}}(x,\theta_{n_0},z_i)\big)^2\,dB_n^1(x|z_i)\\
\le&C\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\Big(\int_{-\infty}^{x}\gamma_n(u|z_i)dB_n^1(u|z_i)-H(x|z_i)\Big)^2dB_n^1(x|z_i)\\
&+C\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\Big(\int_{x}^{\tau}\big(\int_{-\infty}^{u}\gamma_n(t|z_i)dB_n^1(t|z_i)-F(u|\theta,z_i)\big)\gamma_n(u|z_i)dB_n^1(u|z_i)\\
&\quad-\int_{x}^{\tau}\big(H(u|z_i)-F(u|\theta,z_i)\big)\gamma(u|z_i)dB^1(u|z_i)\Big)^2dB_n^1(x|z_i).
\end{aligned}\tag{3.2.5}$$

The first term on the right-hand side of (3.2.5) is bounded by

$$\begin{aligned}
&2C\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\Big(\int_{-\infty}^{x}\gamma_n(u|z_i)dB_n^1(u|z_i)-\int_{-\infty}^{x}\gamma(u|z_i)dB_n^1(u|z_i)\Big)^2dB_n^1(x|z_i)\\
&+2C\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\Big(\int_{-\infty}^{x}\gamma(u|z_i)dB_n^1(u|z_i)-\int_{-\infty}^{x}\gamma(u|z_i)dB^1(u|z_i)\Big)^2dB_n^1(x|z_i)\\
\le&2C\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\int_{-\infty}^{x}\big(\gamma_n(u|z_i)-\gamma(u|z_i)\big)^2dB_n^1(u|z_i)dB_n^1(x|z_i)\\
&+2C\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\Big(\int_{-\infty}^{x}\gamma(u|z_i)d\big(B_n^1(u|z_i)-B^1(u|z_i)\big)\Big)^2dB_n^1(x|z_i)\\
\le&2C\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\big(\gamma_n(u|z_i)-\gamma(u|z_i)\big)^2dB_n^1(u|z_i)\\
&+2C\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\Big(\int_{-\infty}^{x}\gamma(u|z_i)d\big(B_n^1(u|z_i)-B^1(u|z_i)\big)\Big)^2dB_n^1(x|z_i).
\end{aligned}\tag{3.2.6}$$

By (3.2.4), the first term on the right-hand side of (3.2.6) equals $o_p(1)$. Analogously to (2.2.12), by Corollary A.1.2, the second term on the right-hand side of (3.2.6) can be shown to be $o_p(1)$ as well. Hence, the first term on the right-hand side of (3.2.5) is equal to $o_p(1)$. With the same arguments, it can be shown that the second term on the right-hand side of (3.2.5) is equal to $o_p(1)$ as well. Therefore, the third term on the right-hand side of (3.2.3) equals $o_p(1)$ and (3.2.2) holds.

With similar arguments, under D1–D7, we can show the same results for other parts of $\hat{\sigma}^2_{\mathcal{F},n}$ and $\sigma^2_{\mathcal{F}}$. Consequently, the assertion follows. □

For the Model $\mathcal{G}$, let $C_{\mathcal{G}}$, $N^1_{\mathcal{G}}$, $N^0_{\mathcal{G}}$ and their estimates be defined accordingly. The variance and its estimator are defined as

$$\begin{aligned}
\sigma^2 :=& \frac{1}{n_0}\sum_{i=1}^{n_0}\sum_{r=0}^{1}\int \big(N^r_{\mathcal{F}}(x,\theta_{n_0},z_i) - N^r_{\mathcal{G}}(x,\gamma_{n_0},z_i)\big)^2\, dB^r(x|z_i)\\
&- \frac{1}{n_0}\sum_{i=1}^{n_0}\sum_{r=0}^{1}\Big(\int \big(N^r_{\mathcal{F}}(x,\theta_{n_0},z_i) - N^r_{\mathcal{G}}(x,\gamma_{n_0},z_i)\big)\, dB^r(x|z_i)\Big)^2,\\
\hat{\sigma}^2_n :=& \sum_{r=0}^{1}\int \big(N^r_{\mathcal{F},n}(x,\hat{\theta}_n,z) - N^r_{\mathcal{G},n}(x,\hat{\gamma}_n,z)\big)^2 dQ^r_n(x,z)\\
&- \frac{1}{n_0}\sum_{i=1}^{n_0}\sum_{r=0}^{1}\Big(\int \big(N^r_{\mathcal{F},n}(x,\hat{\theta}_n,z_i) - N^r_{\mathcal{G},n}(x,\hat{\gamma}_n,z_i)\big) dB^r_n(x|z_i)\Big)^2.
\end{aligned}$$

Analogously to the case without censoring, Theorem 3.2.4 and Lemma 3.2.5 imply the following two theorems on the asymptotic normality of test statistic T_n. The proofs are omitted.

Theorem 3.2.6. *If* D1–D7 *hold then*

$$\sqrt{n}\cdot\Big(T_n - \big(d_H(\mathcal{F}) - d_H(\mathcal{G})\big)\Big) \xrightarrow{d} \mathcal{N}(0,\sigma^2) \quad \textit{and} \quad \hat{\sigma}^2_n \to \sigma^2.$$

Theorem 3.2.7. *Let* D1–D7 *be satisfied.*

(1) If $\mathcal{H}^{\mathcal{F}}$ holds, then $\sqrt{n}\cdot T_n$ tends to $-\infty$ in probability.

(2) If $\mathcal{H}^{\mathcal{G}}$ holds, then $\sqrt{n}\cdot T_n$ tends to $+\infty$ in probability.

(3) If $\mathcal{H}^0$ holds, then $\sqrt{n}\cdot T_n \xrightarrow{d} \mathcal{N}(0,\sigma^2)$.

If $\sigma^2 > 0$ and $\mathcal{H}^0$ hold true, then

$$\frac{\sqrt{n}\cdot T_n}{\hat{\sigma}_n} \xrightarrow{d} \mathcal{N}(0,1).$$

The decision rules of our test are given as follows: for a given significance level α we will decide for the hypothesis $\mathcal{H}^0$, if $|\sqrt{n}\cdot T_n/\hat{\sigma}_n| \le z_{1-\alpha/2}$, where z_α denotes the α-quantile of a standard normal distribution. In the case of $\sqrt{n}\cdot T_n/\hat{\sigma}_n < -z_{1-\alpha/2}$ we reject $\mathcal{H}^0$ in favor of $\mathcal{H}^{\mathcal{F}}$. If $\sqrt{n}\cdot T_n/\hat{\sigma}_n > z_{1-\alpha/2}$, we reject $\mathcal{H}^0$ in favour of $\mathcal{H}^{\mathcal{G}}$. However, we propose to use the model with less parameters, ever if $\mathcal{H}^0$ is not rejected.

3.3 The Case with Number of Covariates Tending to Infinity

In this section, the Kaplan-Meier estimator is replaced by Beran's estimator. Define the weight function $w_{ni} : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^+$ by

$$w_{ni}(z,h) := \frac{K(\frac{z_i - z}{h})}{\sum_{k=1}^{n} K(\frac{z_k - z}{h})}.$$

where $K : \mathbb{R}^d \to \mathbb{R}^+$ is the kernel function and $h > 0$ the bandwidth. In our setting, the Beran's estimator is defined by

$$\hat{H}_n^{KM}(x|z) := 1 - \prod_{Y_{(i)} \leq x} \Big(1 - \frac{w_{n(i)}(z,h) \cdot \Delta_{(i)}}{1 - \sum_{j=1}^{i-1} w_{n(j)}(z,h)}\Big),$$

where $Y_{(1)} \leq ... \leq Y_{(n)}$ are the ordered $Y_1, .., Y_n$, $\delta_{(i)}$ and $w_{n(i)}(z,h)$ are the corresponding indicator function and weight of $Y_{(n)}$. Further, we define the kernel estimator for the joint distribution by

$$\hat{Q}_n^{KM}(x,z) := \frac{1}{n_0} \sum_{i=1}^{n_0} \hat{H}_n^{KM}(x|z_i) \cdot I(z_i \leq z).$$

The distance $d_H(\mathcal{F})$ defined in (3.1.2) can then be estimated by

$$\hat{d}_{H,n}(\mathcal{F}) := \int \big(\hat{H}_n^{KM}(x|z) - F(x|\hat{\theta}_n, z)\big)^2 \cdot I(x \leq \tau)\, d\hat{Q}_n^{KM}(x,z).$$

For the class $\mathcal{G}$, the estimator $\hat{d}_{H,n}(\mathcal{G})$ is defined in an analogous way. As test statistic we take the difference of the estimated distances

$$T_n := \hat{d}_{H,n}(\mathcal{F}) - \hat{d}_{H,n}(\mathcal{G}).$$

For $z \in [0,1]^d$, denote the kernel estimates of the sub-distributions by

$$\hat{B}_n^1(x|z) := \sum_{i=1}^{n} w_{ni}(z,h) \cdot \Delta_i \cdot I(Y_i \leq x),$$

$$\hat{B}_n^0(x|z) := \sum_{i=1}^{n} w_{ni}(z,h) \cdot (1 - \Delta_i) \cdot I(Y_i \leq x)$$

and the kernel estimate for B by

$$\hat{B}_n(x|z) := \sum_{i=1}^{n} w_{ni}(z,h) \cdot I(Y_i \leq x).$$

Note that we have

$$\hat{B}_n(x|z) = \hat{B}_n^1(x|z) + \hat{B}_n^0(x|z).$$

For the consistency of the kernel estimator, let the following assumptions hold true through out this section.

(i) The functions H and J have bounded derivative and Hessian matrix with respect to z. The functions $\|\partial^2 H/\partial z \partial x\|$ and $\|\partial J/\partial z \partial x\|$ are dominated by Lebesgue integrable functions independent of z.

(ii) As $n_0 \to \infty$, $h \to 0$, $n_0^{-1} h^{-d} \log(n_0) \to 0$ and $n_0 h^{2d} \to \infty$, .

(iii) Let K be a positive Lipschitz continuous function on $[-1, 1]^d$, zero otherwise. Further, for all $x \in \mathbb{R}^d$, $K(x) = K(|x|)$.

The assumptions on the competing models are stated as follows.

E1 For each $(\theta, z) \in \Theta \times [0, 1]^d$, the distribution $F(\cdot|\theta, z)$ has a density function $f(\cdot|\theta, z) : \mathbb{R} \to \mathbb{R}$. The functions $f(\cdot|\theta, z)$ and $1 - F(\cdot|\theta, z)$ are strictly positive $H(\cdot|z)$-a.s.

E2 The functions $\log f$ and $\log(1 - F)$ are three times continuously differentiable in θ on Θ.

E3 The function $\log f$ is dominated by a B^1 square integrable function independent of z. The function $\log(1 - F)$ is dominated by a B^0 square integrable function independent of z.

E4 For each $n_0 \in \mathbb{N}$, the function L_{f,n_0} reaches its maximum at θ_{n_0}, which are interior points of Θ.

E5 The functions $\|\partial \log f/\partial\theta\|^4$ and $\|\partial^2 \log f/\partial\theta^2\|^4$ are dominated by B^1-integrable functions independent of z. The functions $\|\partial \log(1 - F)/\partial\theta\|^4$ and $\|\partial^2 \log(1 - F)/\partial\theta^2\|^4$ are dominated by B^0-integrable functions independent of z.

E6 For any i, j, $k \in \{1, 2, ..., p\}$, the function $\partial^3 \log f/\partial\theta_i \partial\theta_j \partial\theta_k$ is dominated by a B^1 square integrable function independent of z. The function $\partial^3 \log(1 - F)/\partial\theta_i \partial\theta_j \partial\theta_k$ is dominated by a B^0 square integrable function independent of z.

E7 The function $\dot{F}$ and $\ddot{F}$ exist and they are bounded.

E8 The function $L_{f,\infty}$ has a unique maximizer on Θ at θ_*, which is an interior point of Θ. The Hessian matrix $\ddot{L}_{f,\infty}(\theta_*)$ is invertible with inverse $\ddot{L}^{-1}_{f,\infty}(\theta_*)$.

E9 The function $\|\partial\dot{F}/\partial z\|$ is dominated by a B^1 integrable function independent of z. The functions $\|\partial F/\partial z\|$ and $\|\partial^2 \log f/\partial z\partial\theta\|$ are dominated by B^1 square integrable functions independent of z. The function $\|\partial^2\log(1-F)/\partial z\partial\theta\|$ is dominated by a B^0 square integrable functions independent of z.

The following two lemmas can be shown analogously to Lemma A.2.7 and Lemma A.2.8. The proofs are omitted.

Lemma 3.3.1. *For each $r \in \{0,1\}$, let the function $\psi_r : \mathbb{R}\times[0,1]^d \to \mathbb{R}$ be an B^r square integrable function independent of z, then*

$$\sum_{r=0}^{1}\int \psi_r(x,z)dQ_n^r(x,z) = \sum_{r=0}^{1}\int \psi_r(x,z)dQ^r(x,z) + o_p(1).$$

Lemma 3.3.2. *For each $r \in \{0,1\}$, let $(\psi_{rn})_{n\in m\cdot\mathbb{N}}, (\tilde{\psi}_{rn})_{n\in m\cdot\mathbb{N}} : \mathbb{R}\times[0,1]^d \to \mathbb{R}$ be two sequence of functions. Assume that there exists a constant $\delta > 0$, such that for any $r \in \{0,1\}$ and $n \in m\cdot\mathbb{N}$, $|\psi_{rn}|^{2+\delta}$, $|\tilde{\psi}_{rn}|^{2+\delta}$ and $\|\partial\psi_{rn}/\partial z\|^2$ are dominated by the same B^r integrable function independent of z. Define for $n \in m\cdot\mathbb{N}$*

$$\begin{aligned}\sigma_n^2 :=& \frac{1}{n}\sum_{i=1}^{n}\sum_{r=0}^{1}\int \big(\psi_{rn}(x,z_i)+\tilde{\psi}_{rn}(x,z_i)\big)^2 dB^r(x|z_i)\\ &- \frac{1}{n}\sum_{i=1}^{n}\sum_{r=0}^{1}\Big(\int \big(\psi_{rn}(x,z_i)+\tilde{\psi}_{rn}(x,z_i)\big)dB^r(x|z_i)\Big)^2.\end{aligned}$$

If there exists a constant σ such that $\sigma_n^2 \to \sigma^2$, then

$$\begin{aligned}&\sqrt{n}\cdot\sum_{r=0}^{1}\int \psi_{rn}(x,z)d\big(\hat{Q}_n^r(x,z) - E[\hat{Q}_n^r(x,z)]\big)\\ &\quad + \sqrt{n}\cdot\sum_{r=0}^{1}\int \tilde{\psi}_{rn}(x,z)d\big(Q_n^r(x,z) - E[Q_n^r(x,z)]\big) \to N(0,\sigma^2).\end{aligned}$$

Analogously to the case without censoring, based on Lemma 3.3.1 and Lemma 3.3.2, the asymptotic behavior of the maximum likelihood estimator can be stated as follows. The proofs are omitted.

Lemma 3.3.3. *If* E1–E3, E5 *and* E8 *hold, then* $\|\hat\theta_n - \theta_*\| = o_p(1)$.

Lemma 3.3.4. *If* E1–E5 *and* E8 *hold, then* $\|\theta_{n_0} - \theta_*\| = o(1)$.

Corollary 3.3.5. *If* E1–E5 *and* E8 *hold, then* $\|\hat\theta_n - \theta_{n_0}\| = o_p(1)$.

Lemma 3.3.6. *If* E1–E6 *and* E8 *hold, then*

$$\sqrt{n} \cdot \|\hat\theta_n - \theta_{n_0}\| = O_p(1),$$
$$\sqrt{n} \cdot \big\|\hat\theta_n - \theta_{n_0} + \ddot L^{-1}_{f,\infty}(\theta_*) \cdot \dot{\hat L}_{f,n}(\theta_{n_0})\big\| = o_p(1).$$

In order to state the main theorems we introduce the functions $C_{\mathcal F} : \Theta \to \mathbb{R}^p$, and $N_{\mathcal F} : \mathbb{R} \times \Theta \times [0,1]^d \to \mathbb{R}$ with

$$C_{\mathcal F}(\theta) := \int \big(H(x|z) - F(x|\theta,z)\big) \cdot \dot F(x|\theta,z) \cdot I(x \le \tau)\, dQ(x,z),$$
$$N_{\mathcal F}(x,\theta,z) := \big(H(x|z) - F(x|\theta,z)\big)^2 \cdot I(x \le \tau) + 2\int_x^\tau \big(H(u|z) - F(u|\theta,z)\big) dH(u|z) \cdot I(x \le \tau).$$

For each $n \in m \cdot \mathbb{N}$ we define $\gamma_n, C_n : (-\infty, \tau_B] \times [0,1]^d \to \mathbb{R}$ and $N_{\mathcal F,n} : \mathbb{R} \times \Theta \times [0,1]^d \to \mathbb{R}$ with

$$\gamma_n(x|z) := \exp \int_{-\infty}^{x-} \frac{1}{1 - E\big[\hat B_n(u|z)\big]} dE\big[\hat B_n^0(u|z)\big],$$

$$C_n(x|z) := \int_{-\infty}^{x-} \frac{1}{\big(1 - E\big[\hat B_n(u|z)\big]\big)^2} dE\big[\hat B_n^0(u|z)\big],$$

$$N_{\mathcal F,n}(x,\theta,z) := \Big(\int_{-\infty}^{x} \gamma_n(u|z) dE[\hat B_n^1(u|z)] - F(x|\theta,z)\Big)^2 \cdot I(x \le \tau)$$
$$+ 2\int_x^\tau \Big(\int_{-\infty}^{u} \gamma_n(t|z) dE[\hat B_n^1(t|z)] - F(u|\theta,z)\Big)\gamma_n(u|z)\, dE[\hat B_n^1(u|z)] \cdot I(x \le \tau).$$

Further we denote $N^1_{\mathcal F,n}, N^0_{\mathcal F,n}, N^1_{\mathcal F}, N^0_{\mathcal F} : (-\infty, \tau_B] \times \Theta \times [0,1]^d \to \mathbb{R}$ with

$$N^1_{\mathcal F,n}(x,\theta,z) := N_{\mathcal F,n}(x,\theta,z)\gamma_n(x|z) + 2C^T_{\mathcal F}(\theta_*) \cdot \ddot L^{-1}_{f,\infty}(\theta_*) \cdot \frac{\partial \log f(x|\theta,z)}{\partial\theta} \cdot I(x \le \tau)$$

$$
\begin{aligned}
&-\int N_{\mathcal{F},n}(u,\theta,z)C_n(x\wedge u|z)\gamma_n(u|z)dE\big[\hat{B}_n^1(u|z)\big],\\
N_{\mathcal{F},n}^0(x,\theta,z) :=&\frac{1}{1-E\big[\hat{B}_n(x|z)\big]}\int I(x<u)\cdot N_{\mathcal{F},n}(u,\theta,z)\gamma_n(u|z)dE\big[\hat{B}_n^1(u|z)\big]\\
&+2C_{\mathcal{F}}^T(\theta_*)\cdot \ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\frac{\partial\log\big(1-F(x|\theta,z)\big)}{\partial\theta}\cdot I(x\le\tau)\\
&-\int N_{\mathcal{F},n}(u,\theta,z)C_n(x\wedge u|z)\gamma_n(u|z)dE\big[\hat{B}_n^1(u|z)\big],\\
N_{\mathcal{F}}^1(x,\theta,z) :=&N_{\mathcal{F}}(x,\theta,z)\gamma(x|z)+2C_{\mathcal{F}}^T(\theta_*)\cdot \ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\frac{\partial\log f(x|\theta,z)}{\partial\theta}\cdot I(x\le\tau)\\
&-\int N_{\mathcal{F}}(u,\theta,z)C(x\wedge u|z)dH(u|z),\\
N_{\mathcal{F}}^0(x,\theta,z) :=&\frac{1}{1-B(x|z)}\int I(x<u)\cdot N_{\mathcal{F}}(u,\theta,z)dH(u|z)\\
&+2C_{\mathcal{F}}^T(\theta_*)\cdot \ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\frac{\partial\log\big(1-F(x|\theta,z)\big)}{\partial\theta}\cdot I(x\le\tau)\\
&-\int N_{\mathcal{F}}(u,\theta,z)C(x\wedge u|z)dH(u|z).
\end{aligned}
$$

Denote the two constant $d_{H,n}(\mathcal{F})$ and $\sigma_{\mathcal{F}}^2$ as

$$
d_{H,n}(\mathcal{F}) := \int\Big(\int_{-\infty}^{x}\gamma_n(u|z)dE[\hat{B}_n^1(u|z)]-F(x|\theta_{n_0},z)\Big)^2\gamma_n(x|z)\cdot I(x\le\tau)dE[\hat{Q}_n^1(x,z)],
$$

$$
\begin{aligned}
\sigma_{\mathcal{F}}^2 :=&\int\Big(\int N_{\mathcal{F}}^1(x,\theta_*,z)^2\,dB^1(x|z)-\big(\int N_{\mathcal{F}}^1(x,\theta_*,z)\,dB^1(x|z)\big)^2\Big)dz\\
&+\int\Big(\int N_{\mathcal{F}}^0(x,\theta_*,z)^2\,dB^0(x|z)-\big(\int N_{\mathcal{F}}^0(x,\theta_*,z)\,dB^0(x|z)\big)^2\Big)dz.
\end{aligned}
$$

Theorem 3.3.7. *Let* E1–E9 *be satisfied, then we have*

$$
\sqrt{n}\cdot\big(\hat{d}_{H,n}(\mathcal{F})-d_{H,n}(\mathcal{F})\big)\xrightarrow{d}\mathcal{N}(0,\sigma_{\mathcal{F}}^2),
$$

and

$$
d_{H,n}(\mathcal{F})\to d_H(\mathcal{F}).
$$

Proof. Note that we can write

$$
\begin{aligned}
\hat{H}_n(x|z)-F(x|\hat{\theta}_n,z)=&\Big(\hat{H}_n(x|z)-\int_{-\infty}^{x}\gamma_n(u|z)dE[\hat{B}_n^1(u|z)]\Big)\\
&+\Big(\int_{-\infty}^{x}\gamma_n(u|z)dE[\hat{B}_n^1(u|z)]-F(x|\theta_{n_0},z)\Big)-\big(F(x|\hat{\theta}_n,z)-F(x|\theta_{n_0},z)\big).
\end{aligned}
$$

Analogously to Theorem 3.2.4, under E1–E9, based on Lemma 3.3.3, Lemma 3.3.6, Lemma A.4.3, Lemma A.4.4 and Lemma A.4.7, we can show

$$\begin{aligned}\sqrt{n}\cdot\hat{d}_{H,n}(\mathcal{F}) =& \sqrt{n}\cdot\int N_{\mathcal{F},n}(x,\theta_{n_0},z)\,d\hat{Q}_n^{KM}(x,z) + 2C_{\mathcal{F}}^T(\theta_*)\cdot\ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\dot{\hat{L}}_{f,n}(\theta_{n_0})\\ &-2\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\int_x^{\tau}\Big(\int_{-\infty}^{u}\gamma_n(t|z_i)dE[\hat{B}_n^1(t|z_i)] - F(u|\theta,z_i)\Big)\\ &\times\gamma_n(u|z_i)dE[\hat{B}_n^1(u|z_i)]\gamma_n(x|z_i)\,dE[\hat{B}_n^1(x|z_i)] + o_p(1)\end{aligned}$$

and

$$\begin{aligned}d_{H,n}(\mathcal{F}) =& \int N_{\mathcal{F},n}(x,\theta_{n_0},z)\gamma_n(x|z)\,dE[\hat{Q}_n^1(x,z)] + 2C_{\mathcal{F}}^T(\theta_*)\cdot\ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\dot{L}_{f,n_0}(\theta_{n_0})\\ &-2\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\int_{-\infty}^{\tau}\int_x^{\tau}\Big(\int_{-\infty}^{u}\gamma_n(t|z_i)dE[\hat{B}_n^1(t|z_i)] - F(u|\theta,z_i)\Big)\\ &\times\gamma_n(u|z_i)dE[\hat{B}_n^1(u|z_i)]\gamma_n(x|z_i)\,dE[\hat{B}_n^1(x|z_i)] + o_p(1).\end{aligned}$$

Consequently, we can write

$$\begin{aligned}&\sqrt{n}\cdot\big(\hat{d}_{H,n}(\mathcal{F}) - d_{H,n}(\mathcal{F})\big)\\ =&\sqrt{n}\cdot\int N_{\mathcal{F},n}(x,\theta_{n_0},z)\,d\hat{Q}_n^{KM}(x,z) - \sqrt{n}\cdot\int N_{\mathcal{F},n}(x,\theta_{n_0},z)\gamma_n(x|z)\,dE[\hat{Q}_n^1(x,z)] + o_p(1)\\ &+2\sqrt{n}\cdot C_{\mathcal{F}}^T(\theta_*)\cdot\ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\int\frac{\partial\log f(x|\theta_{n_0},z)}{\partial\theta}\cdot I(x\le\tau)d\big(Q_n^1(x,z) - E[Q_n^1(x,z)]\big)\\ &+2\sqrt{n}\cdot C_{\mathcal{F}}^T(\theta_*)\cdot\ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\int\frac{\partial\log\big(1-F(x|\theta_{n_0},z)\big)}{\partial\theta}\cdot I(x\le\tau)d\big(Q_n^0(x,z) - E[Q_n^0(x,z)]\big).\end{aligned}$$

By Lemma A.4.2 with $\psi_n(x,z) = N_{\mathcal{F},n}(x,\theta_{n_0},z)$ and similar arguments as in (3.2.1), we have that

$$\begin{aligned}&\sqrt{n}\cdot\int N_{\mathcal{F},n}(x,\theta_{n_0},z)\,d\hat{Q}_n^{KM}(x,z) - \sqrt{n}\cdot\int N_{\mathcal{F},n}(x,\theta_{n_0},z)\gamma_n(x|z)\,dE[\hat{Q}_n^1(x,z)]\\ =&\sqrt{n}\cdot\int N_{\mathcal{F},n}(x,\theta_{n_0},z)\gamma_n(x|z)\,d\big(\hat{Q}_n^1(x,z) - E[\hat{Q}_n^1(x,z)]\big)\\ &-\sqrt{n}\cdot\int\int N_{\mathcal{F},n}(u,\theta_{n_0},z)C_n(x\wedge u|z)\gamma_n(u|z)dE\big[\hat{B}_n^1(u|z)\big]d\big(\hat{Q}_n^1(x,z) - E[\hat{Q}_n^1(x,z)]\big)\\ &+\sqrt{n}\cdot\int\int\frac{I(x<u)\cdot N_{\mathcal{F},n}(u,\theta_{n_0},z)\gamma_n(u|z)}{1-E\big[\hat{B}_n(x|z)\big]}dE\big[\hat{B}_n^1(u|z)\big]\,d\big(\hat{Q}_n^0(x,z) - E[\hat{Q}_n^0(x,z)]\big)\\ &-\sqrt{n}\cdot\int\int N_{\mathcal{F},n}(u,\theta_{n_0},z)C_n(x\wedge u|z)\gamma_n(u|z)dE\big[\hat{B}_n^1(u|z)\big]d\big(\hat{Q}_n^0(x,z) - E[\hat{Q}_n^0(x,z)]\big).\end{aligned} \tag{3.3.1}$$

In the sequel, we show that

$$\begin{aligned}
&\sqrt{n}\cdot\int N_{\mathcal{F},n}(x,\theta_{n_0},z)\gamma_n(x|z)\,d\big(\hat{Q}_n^1(x,z)-E[\hat{Q}_n^1(x,z)]\big)\\
=&\sqrt{n}\cdot\int N_{\mathcal{F}}(x,\theta_{n_0},z)\gamma(x|z)\,d\big(\hat{Q}_n^1(x,z)-E[\hat{Q}_n^1(x,z)]\big)+o_p(1). \qquad (3.3.2)
\end{aligned}$$

Note that we have

$$\begin{aligned}
&\big|N_{\mathcal{F},n}(x,\theta_{n_0},z)\gamma_n(x|z)-N_{\mathcal{F}}(x,\theta_{n_0},z)\gamma(x|z)\big|\\
\leq&\big|N_{\mathcal{F},n}(x,\theta_{n_0},z)\big(\gamma_n(x|z)-\gamma(x|z)\big)+\big(N_{\mathcal{F},n}(x,\theta_{n_0},z)-N_{\mathcal{F}}(x,\theta_{n_0},z)\big)\gamma(x|z)\big|\\
\leq&\big|N_{\mathcal{F},n}(x,\theta_{n_0},z)\big|\cdot\big|\gamma_n(x|z)-\gamma(x|z)\big|+\big|N_{\mathcal{F},n}(x,\theta_{n_0},z)-N_{\mathcal{F}}(x,\theta_{n_0},z)\big|\cdot\big|\gamma(x|z)\big|.
\end{aligned}$$

Analogously to (2.3.5), by Lemma A.4.6, there exists a constant $C>0$, such that for all $(x,z)\in(-\infty,\tau]\times[0,1]^d$,

$$\big|N_{\mathcal{F},n}(x,\theta_{n_0},z)-N_{\mathcal{F}}(x,\theta_{n_0},z)\big|\leq Ch.$$

Therefore, by Lemma A.4.5 and the boundedness of $N_{\mathcal{F},n}$ and $\gamma(x|z)$, there exists a constant $C>0$, such that for all $(x,z)\in(-\infty,\tau]\times[0,1]^d$,

$$\big|N_{\mathcal{F},n}(x,\theta_{n_0},z)\gamma_n(x|z)-N_{\mathcal{F}}(x,\theta_{n_0},z)\gamma(x|z)\big|\leq Ch. \qquad (3.3.3)$$

Hence, analogously to (2.3.6), Equality (3.3.2) holds. With the same arguments, it can be shown that

$$\begin{aligned}
&-\sqrt{n}\cdot\int\int N_{\mathcal{F},n}(u,\theta_{n_0},z)C_n(x\wedge u|z)\gamma_n(u|z)dE\big[\hat{B}_n^1(u|z)\big]d\big(\hat{Q}_n^1(x,z)-E[\hat{Q}_n^1(x,z)]\big)\\
&+\sqrt{n}\cdot\int\int\frac{I(x<u)\cdot N_{\mathcal{F},n}(u,\theta_{n_0},z)\gamma_n(u|z)}{1-E\big[\hat{B}_n(x|z)\big]}dE\big[\hat{B}_n^1(u|z)\big]\,d\big(\hat{Q}_n^0(x,z)-E[\hat{Q}_n^0(x,z)]\big)\\
&-\sqrt{n}\cdot\int\int N_{\mathcal{F},n}(u,\theta_{n_0},z)C_n(x\wedge u|z)\gamma_n(u|z)dE\big[\hat{B}_n^1(u|z)\big]d\big(\hat{Q}_n^0(x,z)-E[\hat{Q}_n^0(x,z)]\big)\\
=&-\sqrt{n}\cdot\int\int N_{\mathcal{F}}(u,\theta_{n_0},z)C(x\wedge u|z)dH(u|z)d\big(\hat{Q}_n^1(x,z)-E[\hat{Q}_n^1(x,z)]\big)\\
&+\sqrt{n}\cdot\int\int\frac{I(x<u)\cdot N_{\mathcal{F}}(u,\theta_{n_0},z)}{1-B(x|z)}dH(u|z)\,d\big(\hat{Q}_n^0(x,z)-E[\hat{Q}_n^0(x,z)]\big)\\
&-\sqrt{n}\cdot\int\int N_{\mathcal{F}}(u,\theta_{n_0},z)C(x\wedge u|z)dH(u|z)d\big(\hat{Q}_n^0(x,z)-E[\hat{Q}_n^0(x,z)]\big)+o_p(1).
\end{aligned}$$

Consequently, we get

$$\sqrt{n}\cdot\big(\hat{d}_{H,n}(\mathcal{F})-d_{H,n}(\mathcal{F})\big)$$

$$
\begin{aligned}
=&\sqrt{n}\cdot\int N_{\mathcal{F}}(x,\theta_{n_0},z)\gamma(x|z)\,d\big(\hat{Q}_n^1(x,z)-E[\hat{Q}_n^1(x,z)]\big)\\
&-\sqrt{n}\cdot\int\int N_{\mathcal{F}}(u,\theta_{n_0},z)C(x\wedge u|z)dH(u|z)d\big(\hat{Q}_n^1(x,z)-E[\hat{Q}_n^1(x,z)]\big)\\
&+\sqrt{n}\cdot\int\int \frac{I(x<u)\cdot N_{\mathcal{F}}(u,\theta_{n_0},z)}{1-B(x|z)}dH(u|z)\,d\big(\hat{Q}_n^0(x,z)-E[\hat{Q}_n^0(x,z)]\big)\\
&-\sqrt{n}\cdot\int\int N_{\mathcal{F}}(u,\theta_{n_0},z)C(x\wedge u|z)dH(u|z)d\big(\hat{Q}_n^0(x,z)-E[\hat{Q}_n^0(x,z)]\big)\\
&+2\sqrt{n}\cdot C_{\mathcal{F}}^T(\theta_*)\cdot\ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\int\frac{\partial\log f(x|\theta_{n_0},z)}{\partial\theta}\cdot I(x\le\tau)d\big(Q_n^1(x,z)-E[Q_n^1(x,z)]\big)\\
&+2\sqrt{n}\cdot C_{\mathcal{F}}^T(\theta_*)\cdot\ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\int\frac{\partial\log\big(1-F(x|\theta_{n_0},z)\big)}{\partial\theta}\cdot I(x\le\tau)d\big(Q_n^0(x,z)-E[Q_n^0(x,z)]\big)\\
&+o_p(1).
\end{aligned}
$$

Next we show that the conditions of the Lemma 3.3.2 are fulfilled, with

$$\psi_{1n}(x,z)=N_{\mathcal{F}}(x,\theta_{n_0},z)\gamma(x|z)+\int N_{\mathcal{F}}(u,\theta_{n_0},z)C(x\wedge u|z)dH(u|z),$$

$$
\begin{aligned}
\psi_{0n}(x,z)=&\frac{1}{1-B(x|z)}\int I(x<u)\cdot N_{\mathcal{F}}(u,\theta_{n_0},z)dH(u|z)\\
&+\int N_{\mathcal{F}}(u,\theta_{n_0},z)C(x\wedge u|z)dH(u|z),
\end{aligned}
$$

$$\tilde{\psi}_{1n}(x,z)=2\sqrt{n}\cdot C_{\mathcal{F}}^T(\theta_*)\cdot\ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\frac{\partial\log f(x|\theta_{n_0},z)}{\partial\theta}\cdot I(x\le\tau),$$

$$\tilde{\psi}_{0n}(x,z)=2\sqrt{n}\cdot C_{\mathcal{F}}^T(\theta_*)\cdot\ddot{L}_{f,\infty}^{-1}(\theta_*)\cdot\frac{\partial\log\big(1-F(x|\theta_{n_0},z)\big)}{\partial\theta}\cdot I(x\le\tau).$$

By definition the functions $\psi_{1n},\psi_{0n},\tilde{\psi}_{1n}$ and $\tilde{\psi}_{0n}$ are all bounded on $(-\infty,\tau]\times[0,1]^d$.

Further note that

$$
\begin{aligned}
\frac{N_{\mathcal{F}}(x,\theta_{n_0},z)}{\partial z}=&\big(H(x|z)-F(x|\theta_{n_0},z)\big)\Big(\frac{\partial H(x|z)}{\partial z}-\frac{\partial F(x|\theta_{n_0},z)}{\partial z}\Big)\cdot I(x\le\tau)\\
&+2\int_x^\tau\Big(\frac{\partial H(x|z)}{\partial z}-\frac{\partial F(x|\theta_{n_0},z)}{\partial z}\Big)dH(u|z)\cdot I(x\le\tau)\\
&+2\int_x^\tau\big(H(x|z)-F(x|\theta_{n_0},z)\big)\frac{\partial^2 H(u|z)}{\partial z\partial u}du\cdot I(x\le\tau),\\
\frac{\gamma(x|z)}{\partial z}=&-\frac{1}{\big(1-J(x|z)\big)^2}\cdot\frac{\partial J(x|z)}{\partial z}
\end{aligned}
\qquad(3.3.4)
$$

and the derivative of $\int N_{\mathcal{F}}(u,\theta_{n_0},z)C(x\wedge u|z)dH(u|z)$ with respect to z can be written as

$$\int \frac{\partial N_{\mathcal{F}}(u,\theta_{n_0},z)}{\partial z}C(x\wedge u|z)dH(u|z)+\int N_{\mathcal{F}}(u,\theta_{n_0},z)\frac{\partial C(x\wedge u|z)}{\partial z}dH(u|z)$$
$$+\int N_{\mathcal{F}}(u,\theta_{n_0},z)C(x\wedge u|z)\frac{\partial^2 H(u|z)}{\partial z\partial u}du$$

where

$$\begin{aligned}\frac{\partial C(x\wedge u|z)}{\partial z} &= -\int \frac{I(t<x\wedge u)}{\big(1-B(t|z)\big)^3}\frac{\partial B(t|z)}{\partial z}dB^0(t|z)\\ &\quad -\int \frac{I(t<x\wedge u)}{\big(1-B(t|z)\big)^2}\frac{\partial H(t|z)}{\partial z}dJ(t|z)\\ &\quad +\int \frac{I(t<x\wedge u)\big(1-H(t|z)\big)}{\big(1-B(t|z)\big)^2}\frac{\partial^2 J(t|z)}{\partial z\partial t}dt.\end{aligned}$$

Hence, under E9 and Assumption (i), there exists a constant $C>0$ such that

$$\Big\|\frac{\psi_{1n}(x,z)}{\partial z}\Big\| \le C\cdot\Big(1+\Big\|\frac{\partial F(x|\theta_{n_0},z)}{\partial z}\Big\|\Big)\cdot I(x\le\tau).$$

Thus, under E9 the derivative of for any $n\in m\cdot\mathbb{N}$, the functions $\|\psi_{1n}(x,z)/\partial z\|^2$ are dominated by the same B^1 integrable function. With the same arguments, we can show for any $n\in m\cdot\mathbb{N}$, the function $\|\psi_{0n}(x,z)/\partial z\|^2$ with $n\in m\cdot\mathbb{N}$ are dominated by the same B^0 integrable function.

Further, we note that for $r\in\{0,1\}$

$$N^r_{\mathcal{F}}(x,\theta_{n_0},z_i)=\psi_{rn}(x,z)+\tilde{\psi}_{rn}(x,z),$$

und

$$\begin{aligned}\Big\|\frac{\partial N^1_{\mathcal{F}}(x,\theta,z_i)^2}{\partial\theta}\Big\| &= |N^1_{\mathcal{F}}(x,\theta,z_i)|\cdot\Big\|\frac{\partial N^1_{\mathcal{F}}(x,\theta,z_i)}{\partial\theta}\Big\|\\ &\le C\cdot\Big(1+\Big\|\frac{\partial\log f(x|\theta,z)}{\partial\theta}\Big\|\Big)\cdot\Big(1+\Big\|\frac{\partial^2\log f(x|\theta,z)}{\partial\theta^2}\Big\|\Big)\\ \Big\|\frac{\partial N^0_{\mathcal{F}}(x,\theta,z_i)^2}{\partial\theta}\Big\| &= |N^0_{\mathcal{F}}(x,\theta,z_i)|\cdot\Big\|\frac{\partial N^0_{\mathcal{F}}(x,\theta,z_i)}{\partial\theta}\Big\|\\ &\le C\cdot\Big(1+\Big\|\frac{\partial\log\big(1-F(x|\theta,z)\big)}{\partial\theta}\Big\|\Big)\cdot\Big(1+\Big\|\frac{\partial^2\log\big(1-F(x|\theta,z)\big)}{\partial\theta^2}\Big\|\Big).\end{aligned}$$

Thus, under E5 and D7, the function

$$\frac{1}{n}\sum_{i=1}^n\Big(\int N^1_{\mathcal{F}}(x,\cdot,z_i)^2\,dB^1(x|z_i)-\big(\int N^1_{\mathcal{F}}(x,\cdot,z_i)\,dB^1(x|z_i)\big)^2\Big)$$

$$+\frac{1}{n}\sum_{i=1}^{n}\Big(\int N_{\mathcal{F}}^{0}(x,\cdot,z_i)^2\,dB^0(x|z_i)-\big(\int N_{\mathcal{F}}^{0}(x,\cdot,z_i)\,dB^0(x|z_i)\big)^2\Big)$$

has a bounded derivative with respect to θ on Θ. Hence, it follows from Lemma 3.3.4 that

$$\frac{1}{n}\sum_{i=1}^{n}\Big(\int N_{\mathcal{F}}^{1}(x,\theta_{n_0},z_i)^2\,dB^1(x|z_i)-\big(\int N_{\mathcal{F}}^{1}(x,\theta_{n_0},z_i)\,dB^1(x|z_i)\big)^2\Big)$$
$$+\frac{1}{n}\sum_{i=1}^{n}\Big(\int N_{\mathcal{F}}^{0}(x,\theta_{n_0},z_i)^2\,dB^0(x|z_i)-\big(\int N_{\mathcal{F}}^{0}(x,\theta_{n_0},z_i)\,dB^0(x|z_i)\big)^2\Big)\to\sigma_{\mathcal{F}}^2.$$

Therefore, the first part of the assertion follows from Lemma 3.3.2

Based on Lemma A.4.6, the second part of the assertion can be shown analogously as in Theorem 2.3.5. □

For the estimation of the variance $\sigma_{\mathcal{F}}^2$, we define for each $n\in m\cdot\mathbb{N}$, $\hat{\gamma}_n,\hat{C}_n:(-\infty,\tau_B]\times[0,1]^d\to\mathbb{R}$, $\hat{C}_{\mathcal{F},n}:\Theta\to\mathbb{R}^p$ and $N_{\mathcal{F},n}:\mathbb{R}\times\Theta\times[0,1]^d\to\mathbb{R}$ with

$$\hat{\gamma}_n(x|z):=\exp\int_{-\infty}^{x-}\frac{1}{1-\hat{B}_n(u|z)}d\hat{B}_n^0(u|z),$$

$$\hat{C}_n(x|z):=\int_{-\infty}^{x-}\frac{1}{\big(1-\hat{B}_n(u|z)\big)^2}d\hat{B}_n^0(u|z),$$

$$\hat{C}_{\mathcal{F},n}(\theta)=\int\Big(\int_{-\infty}^{x}\hat{\gamma}_n(u|z)d\hat{B}_n^1(u|z)-F(x|\theta,z)\Big)\dot{F}(x|\theta,z)\hat{\gamma}_n(x|z)I(x\le\tau)d\hat{Q}_n^1(x,z),$$

$$\begin{aligned}\hat{N}_{\mathcal{F},n}(x,\theta,z):=&\Big(\int_{-\infty}^{x}\hat{\gamma}_n(u|z)d\hat{B}_n^1(u|z)-F(x|\theta,z)\Big)^2\cdot I(x\le\tau)\\&+2\int_x^{\tau}\Big(\int_{-\infty}^{u}\hat{\gamma}_n(t|z)d\hat{B}_n^1(t|z)-F(u|\theta,z)\Big)\hat{\gamma}_n(u|z)d\hat{B}_n^1(u|z)\cdot I(x\le\tau).\end{aligned}$$

Further we denote $\hat{N}_{\mathcal{F},n}^1,\hat{N}_{\mathcal{F},n}^0:(-\infty,\tau_B]\times\Theta\times[0,1]^d\to\mathbb{R}$ with

$$\begin{aligned}\hat{N}_{\mathcal{F},n}^1(x,\theta,z):=&\hat{N}_{\mathcal{F},n}(u,\theta,z)\hat{\gamma}_n(x|z)+2C_{\mathcal{F},n}^T(\hat{\theta}_n)\cdot\ddot{L}_{f,n}^{-1}(\hat{\theta}_n)\cdot\frac{\partial\log f(x|\theta,z)}{\partial\theta}\cdot I(x\le\tau)\\&-\int\hat{N}_{\mathcal{F},n}(u,\theta,z)\hat{C}_n(x\wedge u|z)\gamma_n(u|z)d\hat{B}_n^1(u|z),\\ \hat{N}_{\mathcal{F},n}^0(x,\theta,z):=&\frac{1}{1-\hat{B}_n(x|z)}\int I(x<u)\cdot\hat{N}_{\mathcal{F},n}(u,\theta,z)\hat{\gamma}_n(u|z)d\hat{B}_n^1(u|z)\\&+2C_{\mathcal{F},n}^T(\hat{\theta}_n)\cdot\ddot{L}_{f,n}^{-1}(\hat{\theta}_n)\cdot\frac{\partial\log\big(1-F(x|\theta,z)\big)}{\partial\theta}\cdot I(x\le\tau)\\&-\int\hat{N}_{\mathcal{F},n}(u,\theta,z)\hat{C}_n(x\wedge u|z)\hat{\gamma}_n(u|z)d\hat{B}_n^1(u|z).\end{aligned}$$

In the next lemma, we show that $\sigma^2_{\mathcal{F}}$ can be consistently estimated by

$$\hat{\sigma}^2_{\mathcal{F},n} := \frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int \hat{N}^1_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)^2\, d\hat{B}^1_n(x|z_i) - \big(\int \hat{N}^1_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)\, d\hat{B}^1_n(x|z_i)\big)^2\Big) + \frac{1}{n_0}\sum_{i=1}^{n_0}\Big(\int \hat{N}^0_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)^2\, d\hat{B}^0_n(x|z_i) - \big(\int \hat{N}^0_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)\, d\hat{B}^0_n(x|z_i)\big)^2\Big).$$

Lemma 3.3.8. *If* E1–E9 *hold, then we have*

$$\hat{\sigma}^2_{\mathcal{F},n} = \sigma^2_{\mathcal{F}} + o_p(1).$$

Proof. As in the proof of the Lemma 3.2.5, we show that

$$\frac{1}{n_0}\sum_{i=1}^{n_0}\int \hat{N}^2_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)\cdot\hat{\gamma}^2_n(x|z_i)\, d\hat{B}^1_n(x|z_i) = \frac{1}{n_0}\sum_{i=1}^{n_0}\int N^2_{\mathcal{F}}(x,\theta_*,z_i)\cdot\gamma^2(x|z_i)\, dB^1(x|z_i) + o_p(1). \tag{3.3.5}$$

First we write

$$\begin{aligned}
&\frac{1}{n_0}\sum_{i=1}^{n_0}\int \hat{N}^2_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)\cdot\hat{\gamma}^2_n(x|z_i)\, d\hat{B}^1_n(x|z_i) - \frac{1}{n_0}\sum_{i=1}^{n_0}\int N^2_{\mathcal{F}}(x,\theta_*,z_i)\cdot\gamma^2(x|z_i)\, dB^1(x|z_i)\\
=&\Big(\frac{1}{n_0}\sum_{i=1}^{n_0}\int \hat{N}^2_{\mathcal{F},n}(x,\hat{\theta}_n,z_i)\cdot\hat{\gamma}^2_n(x|z_i)\, d\hat{B}^1_n(x|z_i) - \frac{1}{n_0}\sum_{i=1}^{n_0}\int N^2_{\mathcal{F},n}(x,\theta_*,z_i)\cdot\gamma^2_n(x|z_i)\, dE[\hat{B}^1_n(x|z_i)]\Big)\\
&+\Big(\frac{1}{n_0}\sum_{i=1}^{n_0}\int N^2_{\mathcal{F},n}(x,\theta_*,z_i)\cdot\gamma^2_n(x|z_i)\, dE[\hat{B}^1_n(x|z_i)] - \frac{1}{n_0}\sum_{i=1}^{n_0}\int N^2_{\mathcal{F}}(x,\theta_*,z_i)\cdot\gamma^2(x|z_i)\, dE[\hat{B}^1_n(x|z_i)]\Big)\\
&+\Big(\frac{1}{n_0}\sum_{i=1}^{n_0}\int N^2_{\mathcal{F}}(x,\theta_*,z_i)\cdot\gamma^2(x|z_i)\, dE[\hat{B}^1_n(x|z_i)] - \frac{1}{n_0}\sum_{i=1}^{n_0}\int N^2_{\mathcal{F}}(x,\theta_*,z_i)\cdot\gamma^2(x|z_i)\, dB^1(x|z_i)\Big)\\
=:&\sigma_{1n}+\sigma_{2n}+\sigma_{3n}.
\end{aligned}$$

Analogously to Lemma 3.2.5, under E1–E8, it can be shown under Lemma 3.3.3 and Corollary A.2.10 that $\sigma_{1n} = o_p(1)$. By (3.3.3), we obtain $\sigma_{2n} = o(1)$. Further, by (3.3.4), under E9 the derivative of the function $N^2_{\mathcal{F}} \cdot \gamma^2$ with respect to z is dominated by B^1 integrable function. Thus, with the same arguments used in Lemma A.2.12, it can be shown that $\sigma_{3n} = o(1)$. Hence, (3.3.5) holds.

By the convergence of Riemann sum we get under E3 and E5,

$$\frac{1}{n_0} \sum_{i=1}^{n_0} \int \hat{N}^2_{\mathcal{F},n}(x, \hat{\theta}_n, z_i) \cdot \hat{\gamma}^2_n(x|z_i) \, d\hat{B}^1_n(x|z_i)$$
$$= \int \int N^2_{\mathcal{F}}(x, \theta_*, z) \cdot \gamma^2(x|z) \, dB^1(x|z) dz + o_p(1)$$

With the same arguments, similar results can be shown under E1-E9 for other terms of $\hat{\sigma}^2_{\mathcal{F},n}$ and $\sigma^2_{\mathcal{F}}$. □

For the Model $\mathcal{G}$, let $C_{\mathcal{G}}$, $N^1_{\mathcal{G}}$, $N^0_{\mathcal{G}}$ and their estimates be defined accordingly. The variance and its estimator are defined as

$$\sigma^2 := \sum_{r=0}^{1} \int \int \left(N^r_{\mathcal{F}}(x, \theta_*, z) - N^r_{\mathcal{G}}(x, \gamma_*, z)\right)^2 dB^r(x|z) dz$$
$$- \sum_{r=0}^{1} \int \left(\int \left(N^r_{\mathcal{F}}(x, \theta_*, z) - N^r_{\mathcal{G}}(x, \gamma_*, z)\right) dB^r(x|z) \right)^2 dz,$$
$$\hat{\sigma}^2_n := \sum_{r=0}^{1} \int \left(\hat{N}^r_{\mathcal{F},n}(x, \hat{\theta}_n, z) - \hat{N}^r_{\mathcal{G},n}(x, \hat{\gamma}_n, z)\right)^2 d\hat{Q}^r_n(x, z)$$
$$- \frac{1}{n_0} \sum_{i=1}^{n_0} \sum_{r=0}^{1} \left(\int \left(\hat{N}^r_{\mathcal{F},n}(x, \hat{\theta}_n, z_i) - \hat{N}^r_{\mathcal{G},n}(x, \hat{\gamma}_n, z_i)\right) d\hat{B}^r_n(x|z_i) \right)^2.$$

Analogously to the case without censoring, Theorem 3.3.7 and Lemma 3.3.8 imply the following two theorems on the asymptotic normality of test statistic T_n. The proofs are omitted.

Theorem 3.3.9. *If* E1–E9 *hold then*

$$\sqrt{n} \cdot \Big(T_n - \big(d_{H,n}(\mathcal{F}) - d_{H,n}(\mathcal{G})\big)\Big) \xrightarrow{d} \mathcal{N}(0, \sigma^2) \quad \text{and} \quad \hat{\sigma}^2_n \to \sigma^2.$$

Theorem 3.3.10. *Let* E1–E9 *be satisfied.*

(1) If $\mathcal{H}^{\mathcal{F}}$ holds, then $\sqrt{n} \cdot T_n$ tends to $-\infty$ in probability.

(2) If $\mathcal{H}^{\mathcal{G}}$ holds, then $\sqrt{n} \cdot T_n$ tends to $+\infty$ in probability.

(3) If $\mathcal{H}^0$ *holds and* $\sqrt{n} \cdot \big(d_{H,n}(\mathcal{F}) - d_{H,n}(\mathcal{G})\big) = o(1)$*, then*

$$\sqrt{n} \cdot T_n \xrightarrow{d} \mathcal{N}(0, \sigma^2).$$

For the case $d = 1$ the following lemma can be shown analogously to Lemma 2.3.9 by Lemma A.4.6 . The proof is omitted.

Lemma 3.3.11. *If* $d = 1$*,* $n_0 h^4 \to 0$ *and the function*

$$\Big\| \int \frac{\partial \log f(x|\theta_*, \cdot)}{\partial \theta} \cdot \big(1 - J(x|\cdot)\big) \cdot \frac{\partial^2 H(x|\cdot)}{\partial z \partial x} dx + \int \frac{\partial \log \big(1 - F(x|\theta_*, \cdot)\big)}{\partial \theta} \cdot \big(1 - H(x|\cdot)\big) \cdot \frac{\partial^2 J(x|\cdot)}{\partial z \partial x} dx \Big\|$$

is bounded on $[0, 1]$*, then under* E1–E6, E8 *and* E9*,*

$$\sqrt{n} \cdot d_{H,n}(\mathcal{F}) = \sqrt{n} \cdot d_H(\mathcal{F}) + o(1).$$

The decision rule of the test for the case $d = 1$ can then be formulated as in the case with fixed n_0 and $m \to \infty$ in Section 2.3. For a given significance level α, we will decide for the hypothesis $\mathcal{H}^0$, if $|\sqrt{n} \cdot T_n/\hat{\sigma}_n| \leq z_{1-\alpha/2}$, where z_α denotes the α-quantile of a standard normal distribution. In the case of $\sqrt{n} \cdot T_n/\hat{\sigma}_n < -z_{1-\alpha/2}$ we reject $\mathcal{H}^0$ in favor of $\mathcal{H}^{\mathcal{F}}$. If $\sqrt{n} \cdot T_n/\hat{\sigma}_n > z_{1-\alpha/2}$, we reject $\mathcal{H}^0$ in favour of $\mathcal{H}^{\mathcal{G}}$.

For the case $d > 1$, the Equality $\sqrt{n} \cdot \|\theta_* - \theta_{n_0}\| = o(1)$ does not hold in general. A one-sided test can then be carried out with

$$\mathcal{H}_a^0: \quad d_H(\mathcal{F}) - d_H(\mathcal{G}) < a$$

against

$$\mathcal{H}_a^1: \quad d_H(\mathcal{F}) - d_H(\mathcal{G}) \geq a,$$

where a is a constant. Given a significance level α, we reject the hypothesis $\mathcal{H}_a^0$ in favour of $\mathcal{H}_a^1$, in the case of $\sqrt{n} \cdot (T_n - a)/\hat{\sigma}_n > z_{1-\alpha}$, otherwise $\mathcal{H}_a^0$ will be accepted.

Chapter 4

Case Study

In the framework of a collaboration with the Institute of Design and Production in Precision Engineering at the University of Stuttgart endurance tests on DC motors (12V-motor type) with brushes were run under the predetermined load levels of $2.5, 3.75, 5, 6.25, 7.5\ mNm$, see Bobrowski et al.(2015) for details. For each load level $m = 16$ lifetimes were observed. By transforming the load levels linearly on to $[0, 1]$, the values become $0.2, 0.4, 0.6, 0.8$ and 1.

In this chapter, we apply our test to this data set. The model class $\mathcal{F}$ is set to be the family of Weibull distributions with constant shape parameter and scale parameter as a linear function of z, i.e.

$$F(x|a_0, b_0, b_1, z) = \begin{cases} 1 - \exp\big(-(x/(b_0 + b_1 z))^{a_0}\big) & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

and

$$\mathcal{F} := \{F(x|a_0, b_0, b_1, z) : a_0, b_0, b_0 + b_1 > 0, z \in [0, 1]\}.$$

The class $\mathcal{G}$ consists of Weibull distributions with constant scale parameter and shape parameter as a linear function of z, i.e.

$$G(x|c_0, c_1, d_0, z) = \begin{cases} 1 - \exp\big(-(x/d_0)^{(c_0 + c_1 z)}\big) & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

and

$$\mathcal{G} := \{G(x|c_0, c_1, d_0, z) : c_0, d_0, c_0 + c_1 > 0, z \in [0, 1]\}.$$

The compactness of the parameter sets Θ and Γ can be realized by assumptions like $\epsilon_1 \leq a_0 \leq \epsilon_2$ and $\epsilon_1 \leq a_0 + a_1 \leq \epsilon_2$ for suitable constants $\epsilon_1, \epsilon_2 > 0$.

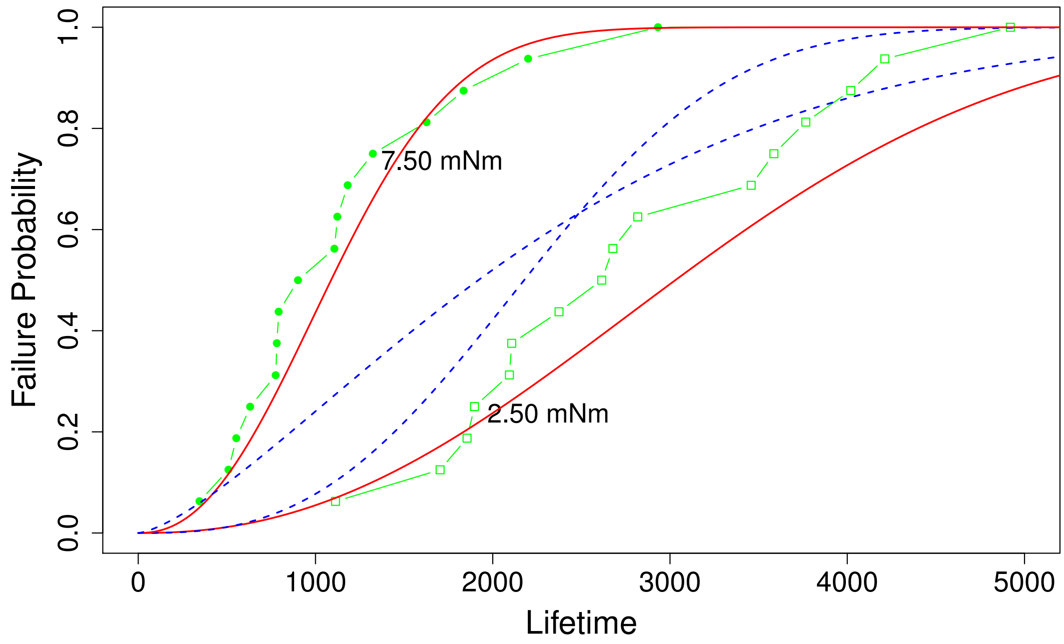

Figure 4.1: The red solid and the blue dashed curves are fitted Weibull distributions from class $\mathcal{F}$ and $\mathcal{G}$, respectively. The green solid curves with points represent the empirical distribution function.

However, this is not a real restriction for practical applications. Note that with the above settings, all necessary conditions for the test are satisfied.

Since $m = 16$ and $n_0 = 5$, empirical distribution functions are used in the test. The calculated p-value for the test statistic $\sqrt{n} \cdot T_n/\hat{\sigma}_n$ is approximately 0.001. Therefore, the null hypothesis is rejected in favour of model $\mathcal{F}$, i.e. it is preferred to model the shape parameter as a constant and the scale parameter as a linear function of the load level rather than the other way round.

In Figure 4.1 the empirical distributions and the fitted distribution functions are shown for load levels 2.5 and 7.5 mNm. It can be seen that the empirical distributions and the solid lines of the Weibull distributions fitted from class $\mathcal{F}$ are close to each other, whereas the dotted lines of distributions fitted from class $\mathcal{G}$ show bad coincidence.

Chapter 5

Simulation Studies

In this chapter we report some Monte Carlo simulation results to evaluate the performance of the proposed model selection tests with moderate sample size. In Section 5.1 the performance of the test is shown comparing the two Weibull model classes as defined in the case study. Section 5.2 deals with an example with two dimensional covariate.

The simulations are conducted as follows. For different combinations of m, n_0, 1000 samples are generated. If m is large we used the empirical function in the test statistic T_n and the estimator for variance $\hat{\sigma}_n$ (as in the case $m \to \infty$, n_0 fixed). For the case with large n_0, the kernel estimator with uniform kernel function $K(x) = \frac{1}{2} \cdot I(|x| \leq 1)$ for the first simulation and

$$K(x_1, x_2) = \frac{1}{4} \cdot I(|x_1| \leq 1) \cdot I(|x_2| \leq 1)$$

for the second simulation is plugged in (as in the case $n_0 \to \infty$, m fixed). For the bandwidth, the minimizer of the function

$$\frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j \neq i}^{n} \big(I(X_i \leq X_j) - \hat{H}_{-i}(X_j|z_i) \big)^2$$

with respect to h is used as proposed in Li et al.(2013), where $\hat{H}_{-i}(\cdot|z_i)$ is the leave-one-out kernel estimator of $H(\cdot|z_i)$.

We also give the results for the case if one of the model classes has to be chosen. In this case the sign of test statistic is indicative. If the sign is negative, model $\mathcal{F}$ should be chosen, otherwise $\mathcal{G}$. In this case, our test can be seen as a model selection procedure.

5.1 Comparing Two Weibull Classes

In this simulation, we assume that the covariate is one dimensional ($d = 1$). The competing model classes are the ones defined in the case study (Chapter 4). The underlying distribution H is set to be Weibull distribution function with shape parameter $(1-p) \cdot 2.26 + p \cdot (2-1.5z)$ and scale parameter $(1-p) \cdot (3563 - 2284z) + p \cdot 2485$ for $0 \leq p \leq 1$. Note that for $p = 0$ the distribution H lies in the class $\mathcal{F}$, for $p = 1$ the distribution H lies in the class $\mathcal{G}$ and for $0 < p < 1$ the distribution H lies neither in $\mathcal{F}$ nor in $\mathcal{G}$. The underlying distribution H was chosen in the way that the simulated data set lies in the similar region as the motor data from the case study.

Tests are conducted for $p = 0, 0.5, 1$ with a moderate data size ($n = 200$). While test decisions are made using the asymptotic critical values as described in Chapter 2. In Table 5.1 and Table 5.2 the percentages of rejection of $\mathcal{H}_0$ in favour of model class $\mathcal{F}$ or $\mathcal{G}$ at significance levels $\alpha = 0.1, 0.2$ along with the test decisions by the sign of test statistic are recorded. In Table 5.1 empirical distribution functions are used, while the kernel estimators for distribution functions are used in Table 5.2. The test performs very well for the case $p = 0$,

	$m \to \infty$		$\alpha = 0.1$		$\alpha = 0.2$		sign	
p	n_0	m	$\mathcal{F}$	$\mathcal{G}$	$\mathcal{F}$	$\mathcal{G}$	$\mathcal{F}$	$\mathcal{G}$
0	20	10	100	0	100	0	100	0
	10	20	100	0	100	0	100	0
0.5	20	10	36.2	0	59.2	0.2	97.3	2.7
	10	20	29.0	0.1	53.0	0.2	96.9	3.1
1	20	10	0	83.9	0	92.3	0.2	99.8
	10	20	0	93.5	0	95.9	0	100

Table 5.1: Performance of the test comparing two Weibull classes ($d = 1$) with empirical distribution functions.

while for the case $p = 1$ the test is not as good as for the case $p = 0$. The reason is that if $p = 1$, the scale parameter of the function H equals $2 - 1.5z$, which changes slowly with respect to z. Thus, it can also be estimated well by a constant scale parameter. However, if $p = 0$, the shape parameter of the function H equals $3563 - 2284z$, which can not be estimated well by a constant.

	$n_0 \to \infty$		$\alpha = 0.1$		$\alpha = 0.2$		sign	
p	n_0	m	$\mathcal{F}$	$\mathcal{G}$	$\mathcal{F}$	$\mathcal{G}$	$\mathcal{F}$	$\mathcal{G}$
0	200	1	98.2	0	99.4	0	99.9	0.1
	50	4	98.8	0	99.9	0	100	0
	20	10	98.5	0	99.0	0	99.9	0.1
	10	20	99.2	0	99.7	0	100	0
0.5	200	1	9.1	0	23.6	0.1	85.3	14.7
	50	4	9.5	0	22.9	0.5	86.8	13.2
	20	10	10.8	0	26.3	0.2	89.0	11.0
	10	20	9.6	0.1	26.5	0.3	86.8	13.4
1	200	1	0	61.0	0	72.9	2.2	97.8
	50	4	0	66.1	0	77.2	1.0	99.0
	20	10	0	69.8	0	79.5	1.2	98.8
	10	20	0	74.7	0	81.3	1.0	99.0

Table 5.2: Performance of the test comparing two Weibull classes ($d = 1$) with kernel estimators for distribution functions.

The combination of m and n_0 seems to have little influence on the outcome. However, if n_0 is large, the empirical distribution function should be preferred in the test, since Table 5.1 shows a better performance of the test than Table 5.2 for the cases $n_0 = 20$ and $m = 10$ or $n_0 = 10$ and $m = 20$. In all cases the sign of T_n is a very good indicator if one of the models has to be chosen.

5.2 Comparing Two Weibull Classes with Two Dimensional Covariate

In this simulation, we assume that the covariate $z := (z^0, z^1) \in [0,1]^2$ is two dimensional. The class $\mathcal{F}$ is set to be a family of Weibull distributions with shape parameter as linear function of the first covariate and scale parameter as linear function of the second covariate, i.e.

$$F(x|a_0, a_1, b_0, b_1, z) = \begin{cases} 1 - \exp\Big(- \big(x/(b_0 + b_1 z^1)\big)^{a_0 + a_1 z^0} \Big) & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

and

$$\mathcal{F} := \{F(x|a_0, a_1, b_0, b_1, z) : a_0, b_0, a_0 + a_1, b_0 + b_1 > 0, z \in [0,1]^2\}.$$

The class $\mathcal{G}$ consists of Weibull distributions with shape parameter as linear function of the second covariate and scale parameter as linear function of the first covariate, i.e.

$$G(x|c_0, c_1, d_0, d_1, z) = \begin{cases} 1 - \exp\Big(- \big(x/(d_0 + d_1 z^0)\big)^{(c_0 + c_1 z^1)} \Big) & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

and

$$\mathcal{G} := \{G(x|c_0, c_1, d_0, d_1, z) : c_0, d_0, c_0 + c_1, d_0 + d_1 > 0, z \in [0,1]^2\}.$$

For the compactness of the parameter sets see the remark in the case study. The underlying distribution H is assumed to be a Weibull distribution with shape parameter $0.2+(1-p)\cdot z^0+p\cdot z^1$ and scale parameter $0.2+p\cdot z^0+(1-p)\cdot z^1$ for $0 \leq p \leq 1$. Again for $p = 0$ the distribution H lies in the model $\mathcal{F}$, for $p = 1$ the distribution H lies in the model $\mathcal{G}$ and for $0 < p < 1$ the distribution H lies neither in $\mathcal{F}$ nor in $\mathcal{G}$.

The simulations are conducted with $n = 100$ in the same way as in Section 5.1. However, since $d = 2$, one-sided tests are conducted. The results are shown for the case $p = 0, 0.5, 1$.

In Table 5.3 the empirical distribution functions are used, while in Table 5.4 the kernel estimators for the distribution functions are used with the bandwidth h calculated by the cross-validation method proposed Li et al.(2013).

	$m \to \infty$		$\alpha = 0.1$		$\alpha = 0.2$		sign	
p	n_0	m	$\mathcal{F}$	$\mathcal{G}$	$\mathcal{F}$	$\mathcal{G}$	$\mathcal{F}$	$\mathcal{G}$
0	25	4	33.6	0.1	49.6	0.2	93.0	7.0
	4	25	45.7	0	65.2	0.2	96.7	3.3
0.5	25	4	1.6	2.0	4.3	5.7	49.8	50.2
	4	25	1.0	2.0	4.5	5.3	49.7	50.3
1	25	4	0.3	34.2	0.4	49.5	8.3	91.7
	4	25	0	46.1	0	65.1	3.1	96.9

Table 5.3: Performance of the test comparing two Weibull classes ($d = 2$) with empirical distribution functions.

For $\alpha = 0.1, 0.2$ the column $\mathcal{F}$ gives the percentage of rejection of the null hypothesis $\mathcal{H}_a^0 : d_{\mathcal{F}}(H) - d_{\mathcal{G}}(H) \geq 0$ in favour of the alternative hypothesis $\mathcal{H}_a^1 : d_{\mathcal{F}}(H) - d_{\mathcal{G}}(H) < 0$ i.e. the model class $\mathcal{F}$ offers a better goodness-of-fit. While the column $\mathcal{G}$ gives the percentage of rejection of the null hypothesis $\mathcal{H}_a^0 : d_{\mathcal{F}}(H) - d_{\mathcal{G}}(H) \leq 0$ in favour of the alternative hypothesis $\mathcal{H}_a^1 : d_{\mathcal{F}}(H) - d_{\mathcal{G}}(H) > 0$. i.e. the model class $\mathcal{G}$ offers a better goodness-of-fit. The columns "sign" present the percentage of decisions for each model class for the case that one model has to be chosen.

	$n_0 \to \infty$		$\alpha = 0.1$		$\alpha = 0.2$		sign	
p	n_0	m	$\mathcal{F}$	$\mathcal{G}$	$\mathcal{F}$	$\mathcal{G}$	$\mathcal{F}$	$\mathcal{G}$
0	100	1	19.9	0	29.9	0	85.6	14.4
	25	4	20.4	0	30.4	0	82.8	17.2
0.5	100	1	0.8	1.2	3.3	2.6	50.3	49.7
	25	4	0.4	0.6	2.3	2.3	49.1	50.9
1	100	1	0	21.1	0	31.8	13.9	86.1
	25	4	0	7.2	0.1	12.7	44.5	55.5

Table 5.4: Performance of the test comparing two Weibull classes ($d = 2$) with kernel estimators for distribution functions and h computed by cross-validation.

Due to the small sample size ($n = 100$), we noticed in the simulations that the bandwidth calculated by the cross-validation method is pretty bad for estimation of the asymptotic variance of the test statistic. Thus, Table 5.4

	$n_0 \to \infty$		$\alpha = 0.1$		$\alpha = 0.2$		sign	
p	n_0	m	$\mathcal{F}$	$\mathcal{G}$	$\mathcal{F}$	$\mathcal{G}$	$\mathcal{F}$	$\mathcal{G}$
0	100	1	54.2	0	69.1	0.1	98.5	1.5
	25	4	43.6	0	62.3	0	97.0	3.0
0.5	100	1	1.3	1.2	5.0	3.4	50.9	49.1
	25	4	1.1	1.1	3.7	3.2	49.7	50.3
1	100	1	0	53.3	0	69.3	2.3	97.7
	25	4	0	45.5	0	62.6	3.3	96.7

Table 5.5: Performance of the test comparing two Weibull classes ($d = 2$) with kernel estimators for distribution functions and fixed $h = 0.2$.

shows a relatively poor performance of the tests. Lack of data is a common problem in the practice especially for multidimensional data. In this case, we propose to used a fixed bandwidth h depending on the data size. In Table 5.5 we give the results for the tests with fixed bandwidth $h = 0.2$. It can be seen that the tests performed better than them in Table 5.4.

Chapter 6

Conclusion

In this thesis, we proposed model selection tests from two competing parametric distribution model classes in a fixed design setting. The measure for the goodness-of-fit of a distribution model class to the underlying distribution is defined based on the Cramér-von Mises distance and the maximum likelihood theory. The model class with smaller distance is chosen to be the better fitted model. Model selection test procedures are derived from the asymptotic normality of the test statistics, which is defined as the difference of the estimated distances.

We handled two cases i.e. the case with a fixed number of covariate values and the number of observations at each covariate value tending to infinity and the case the other way round. The covariate is assumed to be multi-dimensional.

In the first case, the distance between the underlying distributions and the candidating model classes is estimated based on the empirical distribution function at each covariate value. Under a number of regularity assumptions, we showed that $\sqrt{n} \cdot T_n$ is asymptotically normally distributed under $\mathcal{H}^0$, while under the alternative hypothesises it tends to infinity or minus infinity in probability. Hence, our test is consistent. In addition the asymptotic variance can be estimated consistently by a plug-in estimator. Based on these results, the decision rules for the test are formulated.

In the case with the number of covariate values tending to infinity, the empirical distribution function is replaced by the kernel estimator of the distribution function. Similar results were shown for the situation of a one dimensional covariate ($d = 1$). For the case $d \geq 2$, a one-sided-test was proposed.

Further the proposed tests were generalized to the case with right random censoring, where the Kaplan-Meier estimator and Beran estimator were used in place of the empirical distribution function and the kernel estimator. Similar results as in the case without censoring were obtained.

The performance of the tests was reported for some examples in simulation studies. In addition we applied our tests to observed lifetimes of motors in a case study.

The tests proposed in this thesis can be modified or extended in various aspects. First, if there are more than two competing model classes in consideration, the model selection tests can be carried out pairwise.

Secondly, the case with m and n_0 both tending to infinity can be investigated. For instance n_0 can be assumed to be a function of m such that as $m \to \infty$, it holds $n_0 \to \infty$ as well.

Thirdly, in the test statistic T_n the empirical distribution function, kernel estimator for distribution function, the Kaplan-Meier estimator and Beran estimator can also be smoothed with respect to x (double kernel).

Fourthly, the Cramér-von Mises estimator

$$\underset{\theta \in \Theta}{\operatorname{argmax}} \Big(-\int \big(H_n(x|z) - F(x|\theta, z)\big)^2 \, dQ_n(x, z) \Big),$$

which corresponds to the distance measure, and other distance measures like those introduced in Section 1.1 can also be used to describe the goodness-of-fit of a model class to the underlying distribution. By the standardization of the difference of two distances we would expect also the asymptotic normality property of the appropriate test statistics. However, different distance definitions can lead to different decisions as shown in the simulation studies.

Last but not the least, the proposed tests can also be extended to the model selection between semi-parametric models such as Cox-model, which is often applied in the survival analysis. The asymptotic theorems of the partial likelihood estimator for the parameter in Cox-Model was proven by Struthers and Kalbfleisch (1986) and Lin and Wei (1989). Similarly, the asymptotic behaviour can also be shown for the Breslow estimator. Therefore, the corresponding theorems could be established analogously as for parametric model classes proposed in this thesis.

Appendix A

In this Appendix we show some auxiliary lemmas, based on which the theorems in this thesis are proven. Section A1– Section A4 are corresponding to Section 2.2, Section 2.3, Section 3.2 and Section 3.3, respectively.

A.1 Appendix of Section 2.2 ($m \to \infty, n_0$ Fixed)

In this section, let $z_1, ..., z_n \in \mathbb{R}^d$ with $n = n_0 \cdot m$ be the covariate values as defined in Section 1.1. For each $i \in \{1, ..., k\}$ and $j \in \{1, ..., n\}$ with $k \in \mathbb{N}$, let X_{ij} be a real valued random variables with distribution function $H_i(\cdot|z_j)$. It is assumed that for each $i, a \in \{1, ..., k\}$ and $j, b \in \{1, ..., n\}$, if $j \neq b$, X_{ij} and X_{ab} are independent.

The empirical distribution function at $z \in [0, 1]^d$ is denoted by

$$H_{in}(x|z) := \frac{1}{m} \sum_{j=1}^{n} \delta_j(z) I(X_{ij} \leq x),$$

where $\delta_j(z) = 1$ if $z_j = z$, otherwise, $\delta_j(z) = 0$. Let $\psi : \mathbb{R}^k \times [0, 1]^d \to \mathbb{R}$ be a function. For simplicity of notation, we denote for each $i \in \{1, ..., k\}$,

$$\begin{aligned} P_{in}(z) :=& \sqrt{n} \int \cdots \int \psi(x_1, ..., x_k, z) dH_1(x_1|z) \cdots dH_{i-1}(x_{i-1}|z) \\ & \times dH_{in}(x_i|z) dH_{i+1}(x_{i+1}|z) \cdots dH_k(x_k|z), \\ \bar{P}_{in} :=& \frac{1}{n_0} \sum_{j=1}^{n_0} P_{in}(z_j). \end{aligned}$$

Similarly, we denote

$$P_n(z) := \sqrt{n} \int \cdots \int \psi(x_1, ..., x_k, z) dH_{1n}(x_1|z) \cdots dH_{kn}(x_k|z),$$

$$P_{0n}(z) := \sqrt{n}\int\cdots\int \psi(x_1,...,x_k,z)dH_1(x_1|z)\cdots dH_k(x_k|z),$$

$$\bar{P}_n := \frac{1}{n_0}\sum_{j=1}^{n_0} P_n(z_j), \qquad \bar{P}_{0n} := \frac{1}{n_0}\sum_{j=1}^{n_0} P_{0n}(z_j).$$

Lemma A.1.1. *Let $C > 0$ be a constant such that*

$$E[\psi^2(X_{1i_1},...,X_{ki_k},z)] \le C$$

for all $i_1,...,i_k \in \{1,...,n\}$ and $z \in \{z_1,...,z_{n_0}\}$. Then as $m \to \infty$ and n_0 stays fixed,

$$\bar{P}_n = \sum_{i=1}^{k}\bar{P}_{in} - (k-1)\bar{P}_{0n} + o_p(1).$$

Proof. First we show that

$$Var[\bar{P}_n - \sum_{i=1}^{k}\bar{P}_{in} + (k-1)\bar{P}_{0n}] = o(1).$$

In this proof, we denote the following sets

$$I := \big\{(i_1,...,i_k,j_1,...,j_k) : i_1,...,i_k,j_1,...,j_k \in \{1,...,n\}\big\},$$

$$I_1 := \big\{(i_1,...,i_k,j_1,...,j_k) \in I : \big(\cup_{a=1}^{k}\{i_a\}\big) \cap \big(\cup_{b=1}^{k}\{j_b\}\big) = \emptyset\big\},$$

$$I_2 := \{(i_1,...,i_k,j_1,...,j_k) \in I\backslash I_1 : |\cup_{a=1}^{k}\{i_a\} \cup_{b=1}^{k}\{j_b\}| \le 2k-2\},$$

$$I_3 := \{(i_1,...,i_k,j_1,...,j_k) \in I\backslash I_1 : |\cup_{a=1}^{k}\{i_a\} \cup_{b=1}^{k}\{j_b\}| = 2k-1\},$$

$$J := \big\{(i_1,...,i_k) : i_1,...,i_k \in \{1,...,n\}\big\},$$

$$J_1 := \big\{(i_1,...,i_k) \in J : |\cup_{a=1}^{k}\{i_a\}| \le k-1\big\},$$

where for a finite set S, $|S|$ denotes the number of elements in the set S. Note that I_1, I_2 and I_3 are disjoint and

$$I = I_1 \cup I_2 \cup I_3.$$

For each a, $b \in \{1,...,k\}$, we denote further

$$I_{ab} := \big\{(i_1,...,i_k,j_1,...,j_k) \in I_3 : i_a = j_b\big\}$$

$$\bar{I}_{ab} := \big\{(i_1,...,i_k,j_1,...,j_k) \in I_2 : i_a = j_b\big\}.$$

Let $z, \tilde{z}$ be two arbitrary covariate values in $\{z_1, ..., z_{n_0}\}$. For any set $\tilde{I} \subset I$, denote

$$C_{\tilde{I},n}(z,\tilde{z}) := n_0 m^{-2k+1} \sum_{\tilde{I}} \delta_{i_1}(z) \cdots \delta_{i_k}(z) \delta_{j_1}(\tilde{z}) \cdots \delta_{j_k}(\tilde{z}) \\ \times Cov\big[\psi(X_{1i_1}, ..., X_{ki_k}, z), \psi(X_{1j_1}, ..., X_{kj_k}, \tilde{z})\big].$$

By definition, we have

$$P_n(z) = \sqrt{n} m^{-k} \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \delta_{i_1}(z) \cdots \delta_{i_k}(z) \psi(X_{1i_1}, ..., X_{ki_k}, z).$$

Thus, we can write

$$Cov[P_n(z), P_n(\tilde{z})] = C_{I,n}(z,\tilde{z}) = C_{I_1,n}(z,\tilde{z}) + C_{I_2,n}(z,\tilde{z}) + C_{I_3,n}(z,\tilde{z}). \quad \text{(A.1.1)}$$

If $(i_1, ..., i_k, j_1, ..., j_k) \in I_1$, by assumption $\psi(X_{1i_1}, ..., X_{ki_k}, z)$ and $\psi(X_{1j_1}, ..., X_{kj_k}, \tilde{z})$ are independent, thus

$$Cov\big[\psi(X_{1i_1}, ..., X_{ki_k}, z), \psi(X_{1j_1}, ..., X_{kj_k}, \tilde{z})\big] = 0.$$

Therefore, the first term on the right-hand side of (A.1.1)

$$C_{I_1,n}(z,\tilde{z}) = 0. \quad \text{(A.1.2)}$$

For the second term, note that

$$I_2 = \bigcup_{a=1}^{k} \bigcup_{b=1}^{k} \bar{I}_{ab}.$$

Thus,

$$\big|C_{I_2,n}(z,\tilde{z})\big| \leq \sum_{a=1}^{k} \sum_{b=1}^{k} \big|C_{\bar{I}_{ab},n}(z,\tilde{z})\big|. \quad \text{(A.1.3)}$$

In the sequel, we will show for $a = b = 1$,

$$\big|C_{\bar{I}_{ab},n}(z,\tilde{z})\big| \leq C(2k-1)(k-1) \cdot n_0 m^{-1} =: c_n. \quad \text{(A.1.4)}$$

Note that by Cauchy-Schwarz's inequality and the square integrability of ψ, for all $(i_1, ..., i_k, j_1, ..., j_k) \in I$,

$$\Big|Cov\big[\psi(X_{1i_1}, ..., X_{ki_k}, z), \psi(X_{1j_1}, ..., X_{kj_k}, \tilde{z})\big]\Big|$$

$$\leq \Big(Var\big[\psi(X_{1i_1}, ..., X_{ki_k}, z)\big] \cdot Var\big[\psi(X_{1j_1}, ..., X_{kj_k}, \tilde{z})\big]\Big)^{1/2}$$
$$\leq \Big(E\big[\psi^2(X_{1i_1}, ..., X_{ki_k}, z)\big] \cdot E\big[\psi^2(X_{1j_1}, ..., X_{kj_k}, \tilde{z})\big]\Big)^{1/2} \leq C. \tag{A.1.5}$$

Consequently,

$$\big|C_{\bar{I}_{11},n}(z, \tilde{z})\big| \leq Cn_0 m^{-2k+1} \sum_{\bar{I}_{11}} \delta_{i_1}(z) \cdots \delta_{i_k}(z) \delta_{j_1}(\tilde{z}) \cdots \delta_{j_k}(\tilde{z}). \tag{A.1.6}$$

Further note that if $(i_1, ..., i_k, j_1, ..., j_k) \in \bar{I}_{11}$, then at least two numbers out of $i_1, ..., i_k, j_2, ..., j_k$ are equal. Hence, there are at most

$$\binom{2k-1}{2} \cdot m^{2k-2}$$

elements in $\bar{I}_{11}$ with

$$\delta_{ni_1}(z) \cdots \delta_{ni_k}(z) \delta_{nj_1}(\tilde{z}) \cdots \delta_{nj_k}(\tilde{z}) \neq 0.$$

Hence, by (A.1.6),

$$\big|C_{\bar{I}_{11},n}(z, \tilde{z})\big| \leq Cn_0 m^{-2k+1} \sum_{\bar{I}_{11}} \delta_{i_1}(z) \cdots \delta_{i_k}(z) \delta_{j_1}(\tilde{z}) \cdots \delta_{j_k}(\tilde{z})$$
$$\leq Cn_0 m^{-2k+1} \cdot \binom{2k-1}{2} \cdot m^{2k-2} = c_n. \tag{A.1.7}$$

With the same arguments, it can be shown that (A.1.4) holds true for any a, $b \in \{1, 2, ..., k\}$. By (A.1.3), we obtain then

$$\big|C_{I_2,n}(z, \tilde{z})\big| \leq \sum_{a=1}^{k} \sum_{b=1}^{k} \big|C_{\bar{I}_{ab},n}(z, \tilde{z})\big| \leq k^2 c_n. \tag{A.1.8}$$

For the third term on the right-hand side of (A.1.1), note that

$$C_{I_3,n}(z, \tilde{z}) = \sum_{a=1}^{k} \sum_{b=1}^{k} C_{I_{ab},n}(z, \tilde{z}). \tag{A.1.9}$$

In the following we will show that for $a = b = 1$.

$$\big|C_{I_{ab},n}(z, \tilde{z}) - Cov[P_{an}(z), P_{bn}(\tilde{z})]\big| \leq c_n.$$

With the same arguments, it can be shown that this inequality holds true for any $a, b \in \{1, ..., k\}$. Denote the function $\tilde{\psi} : \mathbb{R} \times [0, 1]^d \to \mathbb{R}$ with

$$\tilde{\psi}(x, z) := \int \cdots \int \psi(x, x_2, ..., x_k, z) dH_2(x_2|z) \cdots dH_k(x_k|z).$$

Note that if $(i_1, ..., i_k, j_1, ..., j_k) \in I_{11}$ with

$$\delta_{i_1}(z) \cdots \delta_{i_k}(z)\delta_{j_1}(\tilde{z}) \cdots \delta_{j_k}(\tilde{z}) \neq 0$$

then it must be holds that

$$z_{i_2} = ... = z_{i_k} = z \quad \text{and} \quad z_{j_2} = ... = z_{j_k} = \tilde{z}$$

and $X_{1i_1}, ..., X_{ki_k}, X_{1j_2}, ..., X_{kj_k}$ are independent. Therefore, by Fubini's theorem for any $(i_1, ..., i_k, j_1, ..., j_k) \in I_{11}$ with

$$\delta_{i_1}(z) \cdots \delta_{i_k}(z)\delta_{j_1}(\tilde{z}) \cdots \delta_{j_k}(\tilde{z}) \neq 0,$$

we have

$$\begin{aligned}
&Cov\big[\psi(X_{1i_1}, X_{2i_2}, ..., X_{ki_k}, z), \psi(X_{1j_1}, X_{2j_2}, ..., X_{kj_k}, \tilde{z})\big] \\
=&Cov\big[\psi(X_{1i_1}, X_{2i_2}, ..., X_{ki_k}, z), \psi(X_{1i_1}, X_{2j_2}, ..., X_{kj_k}, \tilde{z})\big] \\
=&E\big[\psi(X_{1i_1}, X_{2i_2}, ..., X_{ki_k}, z) \cdot \psi(X_{1i_1}, X_{2j_2}, ..., X_{kj_k}, \tilde{z})\big] \\
&\quad - E\big[\psi(X_{1i_1}, X_{2i_2}, ..., X_{ki_k}, z)\big] \cdot E\big[\psi(X_{1i_1}, X_{2j_2}, ..., X_{kj_k}, \tilde{z})\big] \\
=&E\big[\tilde{\psi}(X_{1i_1}, z) \cdot \tilde{\psi}(X_{1i_1}, \tilde{z})\big] - E\big[\tilde{\psi}(X_{1i_1}, z)\big] \cdot E\big[\tilde{\psi}(X_{1i_1}, \tilde{z})\big] \\
=&Cov\big[\tilde{\psi}(X_{1i_1}, z), \tilde{\psi}(X_{1i_1}, \tilde{z})\big]. \qquad \text{(A.1.10)}
\end{aligned}$$

Therefore,

$$C_{I_{11},n}(z, \tilde{z}) = n_0 m^{-2k+1} \sum_{I_{11}} \delta_{i_1}(z) \cdots \delta_{i_k}(z)\delta_{j_1}(\tilde{z}) \cdots \delta_{j_k}(\tilde{z}) Cov\big[\tilde{\psi}(X_{1i_1}, z), \tilde{\psi}(X_{1i_1}, \tilde{z})\big].$$

Further, by the independence of X_{1i_1} and X_{1j_1} provided $i_1 \neq j_1$,

$$\begin{aligned}
&Cov\big[P_{1n}(z), P_{1n}(\tilde{z})\big] \\
=&nCov\Big[\int \tilde{\psi}(x, z) dH_{1n}(x|z), \int \tilde{\psi}(x, \tilde{z}) dH_{1n}(x|\tilde{z})\Big] \\
=&nm^{-2} Cov\Big[\sum_{i_1=1}^{n} \delta_{i_1}(z)\tilde{\psi}(X_{1i_1}, z), \sum_{j_1=1}^{n} \delta_{j_1}(\tilde{z})\tilde{\psi}(X_{1j_1}, \tilde{z})\Big] \\
=&n_0 m^{-1} \sum_{i_1=1}^{n} \delta_{i_1}(z)\delta_{i_1}(\tilde{z}) Cov\big[\tilde{\psi}(X_{1i_1}, z), \tilde{\psi}(X_{1i_1}, \tilde{z})\big] \\
=&n_0 m^{-2k+1} \sum_{I_{11} \cup \bar{I}_{11}} \delta_{i_1}(z) \cdots \delta_{i_k}(z)\delta_{j_1}(\tilde{z}) \cdots \delta_{j_k}(\tilde{z}) Cov\big[\tilde{\psi}(X_{1i_1}, z), \tilde{\psi}(X_{1i_1}, \tilde{z})\big]
\end{aligned}$$

where the last term follows from the definition of $\delta_{i_1}(z), ..., \delta_{i_k}(z), \delta_{j_1}(\tilde{z}), ..., \delta_{j_k}(\tilde{z})$. Note further that by (A.1.5) and (A.1.10)

$$Cov\big[\tilde{\psi}(X_{1i_1}, z), \tilde{\psi}(X_{1i_1}, \tilde{z})\big] \leq C. \tag{A.1.11}$$

Hence,

$$\begin{aligned}
&\Big|C_{I_{11},n}(z,\tilde{z}) - Cov\big[P_{1n}(z), P_{1n}(\tilde{z})\big]\Big| \\
\leq& n_0 m^{-2k+1} \sum_{\bar{I}_{11}} \delta_{i_1}(z)\cdots\delta_{i_k}(z)\delta_{j_1}(\tilde{z})\cdots\delta_{j_k}(\tilde{z}) \Big|Cov\big[\tilde{\psi}(X_{1i_1}, z), \tilde{\psi}(X_{1i_1}, \tilde{z})\big]\Big| \\
\leq& C n_0 m^{-2k+1} \sum_{\bar{I}_{11}} \delta_{i_1}(z)\cdots\delta_{i_k}(z)\delta_{j_1}(\tilde{z})\cdots\delta_{j_k}(\tilde{z}) \leq c_n, \qquad \text{(A.1.12)}
\end{aligned}$$

where the last step follows from (A.1.7). Therefore, by (A.1.9),

$$\begin{aligned}
&\Big|C_{I_3,n}(z,\tilde{z}) - \sum_{a=1}^{k}\sum_{b=1}^{k} Cov[P_{an}(z), P_{bn}(\tilde{z})]\Big| \\
\leq& \sum_{a=1}^{k}\sum_{b=1}^{k} \Big|C_{I_{ab},n}(z,\tilde{z}) - Cov[P_{an}(z), P_{bn}(\tilde{z})]\Big| \leq k^2 c_n. \qquad \text{(A.1.13)}
\end{aligned}$$

By (A.1.1), (A.1.2), (A.1.8) and (A.1.13), we obtain

$$\begin{aligned}
&\Big|Cov[P_n(z), P_n(\tilde{z})] - \sum_{a=1}^{k}\sum_{b=1}^{k} Cov[P_{an}(z), P_{bn}(\tilde{z})]\Big| \\
\leq& \Big|C_{I_1,n}(z,\tilde{z})\Big| + \Big|C_{I_2,n}(z,\tilde{z})\Big| + \Big|C_{I_3,n}(z,\tilde{z}) - \sum_{a=1}^{k}\sum_{b=1}^{k} Cov[P_{an}(z), P_{bn}(\tilde{z})]\Big| \leq 2k^2 c_n
\end{aligned}$$

Hence, by definition,

$$\begin{aligned}
&\Big|Var[\bar{P}_n] - \sum_{a=1}^{k}\sum_{b=1}^{k} Cov[\bar{P}_{an}, \bar{P}_{bn}]\Big| \\
=& \Big|\frac{1}{n_0^2}\sum_{i=1}^{n_0}\sum_{j=1}^{n_0} Cov[P_n(z_i), P_n(z_j)] - \frac{1}{n_0^2}\sum_{i=1}^{n_0}\sum_{j=1}^{n_0}\sum_{a=1}^{k}\sum_{b=1}^{k} Cov[P_{an}(z_i), P_{bn}(z_j)]\Big| \\
\leq& \frac{1}{n_0^2}\sum_{i=1}^{n_0}\sum_{j=1}^{n_0}\Big|Cov[P_n(z_i), P_n(z_j)] - \sum_{a=1}^{k}\sum_{b=1}^{k} Cov[P_{an}(z_i), P_{bn}(z_j)]\Big| \leq 2k^2 c_n = o(1),
\end{aligned}$$

where the last step follows from $c_n \to 0$ as $m \to \infty$ and n_0 stays fixed. Thus, we get

$$Var[\bar{P}_n] = \sum_{a=1}^{k}\sum_{b=1}^{k} Cov[\bar{P}_{an}, \bar{P}_{bn}] + o(1).$$

Analogously, we can show that

$$Cov[\bar{P}_n, \sum_{j=1}^{k} \bar{P}_{jn}] = \sum_{i=1}^{k}\sum_{j=1}^{k} Cov[\bar{P}_{in}, \bar{P}_{jn}] + o(1).$$

Consequently,

$$\begin{aligned}
&Var[\bar{P}_n - \sum_{i=1}^{k} \bar{P}_{in} + (k-1)\bar{P}_{0n}] \\
=&Var[\bar{P}_n - \sum_{i=1}^{k} \bar{P}_{in}] \\
=&Var[\bar{P}_n] - 2Cov[\bar{P}_n, \sum_{i=1}^{k} \bar{P}_{in}] + \sum_{i=1}^{k}\sum_{j=1}^{k} Cov[\bar{P}_{in}, \bar{P}_{jn}] = o(1). \qquad (A.1.14)
\end{aligned}$$

Next, we show that

$$E\big[\bar{P}_n - \sum_{i=1}^{k} \bar{P}_{in} + (k-1)\bar{P}_{0n}\big] = o(1).$$

For any set $\tilde{J} \subset J$, denote

$$E_{\tilde{J},n}(z) := \sqrt{n}m^{-k} \sum_{\tilde{J}} \delta_{i_1}(z)\cdots\delta_{i_k}(z)E\big[\psi(X_{1i_1}, ..., X_{ki_k}, z)\big].$$

We can write then

$$E[P_n(z)] = E_{J,n}(z) = E_{J_1,n}(z) + E_{J\backslash J_1,n}(z). \qquad (A.1.15)$$

Note that if $(i_1, ..., i_k) \in J\backslash J_1$ with

$$\delta_{i_1}(z)\cdots\delta_{i_k}(z) \neq 0$$

it must hold that

$$z_{i_1} = ... = z_{i_k} = z$$

and $X_{1i_1}, ..., X_{ki_k}$ are independent, thus

$$\sqrt{n}E\big[\psi(X_{1i_1}, ..., X_{ki_k}, z)\big] = P_{0n}(z). \qquad (A.1.16)$$

Hence, the second term on the right-hand side of (A.1.15)

$$E_{J\backslash J_1,n}(z) = m^{-k} \sum_{J\backslash J_1} \delta_{i_1}(z)\cdots\delta_{i_k}(z)P_{0n}(z).$$

Further, note that

$$P_{0n}(z) = m^{-k} \sum_J \delta_{i_1}(z) \cdots \delta_{i_k}(z) P_{0n}(z).$$

Therefore, we can write

$$\begin{aligned}
&\big|E[P_n(z)] - P_{0n}(z)\big| \\
\leq &\big|E_{J_1,n}(z)\big| + \big|E_{J\setminus J_1,n}(z) - P_{0n}(z)\big| \\
\leq &m^{-k} \sum_{J_1} \delta_{i_1}(z) \cdots \delta_{i_k}(z) \cdot \Big(\big|\sqrt{n} E\big[\psi(X_{1i_1}, ..., X_{ki_k}, z)\big]\big| + \big|P_{0n}(z)\big|\Big). \quad \text{(A.1.17)}
\end{aligned}$$

Note that by definition there are at most

$$\binom{k}{2} \cdot m^{k-1}$$

elements in J_1 with

$$\delta_{ni_1}(z) \cdots \delta_{ni_k}(z) \neq 0.$$

Further, by Jensen's inequality,

$$E\big[\psi(X_{1i_1}, ..., X_{ki_k}, z)\big] \leq \Big(E\big[\psi^2(X_{1i_1}, ..., X_{ki_k}, z)\big]\Big)^{1/2} \leq C^{1/2}.$$

Analogously, by (A.1.16),

$$|P_{0n}(z)| \leq C^{1/2}\sqrt{n}.$$

Hence, by (A.1.17), we obtain

$$\big|E[P_n(z)] - P_{0n}(z)\big| \leq 2C^{1/2}\sqrt{n} m^{-k} \sum_{J_1} \delta_{i_1}(z) \cdots \delta_{i_k}(z) \leq C^{1/2} k(k-1)\sqrt{n} m^{-1}.$$

Consequently, as $m \to \infty$,

$$\begin{aligned}
\big|E[\bar{P}_n] - \bar{P}_{0n}\big| = &\frac{1}{n_0} \sum_{i=1}^{n_0} \big|E[P_n(z_i)] - P_{0n}(z_i)\big| \\
\leq &C^{1/2} k(k-1)\sqrt{n} m^{-1} = o(1). \quad \text{(A.1.18)}
\end{aligned}$$

In the sequel, we show that for $i = 1$,

$$E[\bar{P}_{in}] = \bar{P}_{0n}.$$

With the same arguments, it can be shown that the equality holds true for any $i \in \{1, ..., k\}$. By the definition of z and $\delta_j(z)$,

$$\begin{aligned}E[P_{1n}(z)] =& E\Big[\sqrt{n}\int \tilde{\psi}(x,z)dH_{1n}(x|z)\Big]\\ =& \sqrt{n}\cdot\frac{1}{m}\sum_{j=1}^{n}\delta_j(z)E\big[\tilde{\psi}(X_{1j},z)\big]\\ =& \sqrt{n}\cdot\frac{1}{m}\sum_{j=1}^{n}\delta_j(z)\int \tilde{\psi}(x,z)dH_1(x|z_j) = P_{0n}(z).\end{aligned}$$

Thus,

$$E[\bar{P}_{1n}] = \frac{1}{n_0}\sum_{i=1}^{n_0}E[P_{1n}(z_i)] = \frac{1}{n_0}\sum_{i=1}^{n_0}P_{0n}(z_i) = \bar{P}_{0n}.$$

Therefore, by (A.1.18)

$$E\big[\bar{P}_n - \sum_{i=1}^{k}\bar{P}_{in} + (k-1)\bar{P}_{0n}\big] = o(1). \tag{A.1.19}$$

The assertion follows then from Chebyshev's inequality by (A.1.14) and (A.1.19). □

Lemma A.1.1 can also be proven as in Lemma 2.2 in Stute (1995), however, we used Chebyshev's inequality, so that it can be easily extended to the case with kernel estimator (Lemma A.2.9) under weak conditions.

Corollary A.1.2. *Under the assumptions of Lemma A.1.1, we have*

$$n^{-1/2}\bar{P}_n = n^{-1/2}\bar{P}_{0n} + o_p(1).$$

Proof. Let $\tilde{\psi} : \mathbb{R}\times[0,1]^d \to \mathbb{R}$ be defined as in the proof of Lemma A.1.1. Analogously to Lemma 2.2.1, we have

$$\begin{aligned}n^{-1/2}P_{1n} =& \frac{1}{n_0}\sum_{i=1}^{n_0}\int \tilde{\psi}(x,z_i)dH_{1n}(x|z_i)\\ =& \frac{1}{n_0}\sum_{i=1}^{n_0}\int \tilde{\psi}(x,z_i)dH_1(x|z_i) + o_p(1) = n^{-1/2}P_{0n} + o_p(1).\end{aligned}$$

With the same arguments, it can be shown that for any $i \in \{1, ..., k\}$

$$n^{-1/2}\bar{P}_{in} = n^{-1/2}\bar{P}_{0n} + o_p(1).$$

Hence, by Lemma A.1.1, we obtain

$$n^{-1/2}\bar{P}_n = n^{-1/2}\sum_{i=1}^{k}\bar{P}_{in} - (k-1)n^{-1/2}\bar{P}_{0n} + o_p(1) = n^{-1/2}\bar{P}_{0n} + o_p(1).$$

□

A.2 Appendix of Section 2.3 ($n_0 \to \infty, m$ Fixed)

In this section, we denote for $h < 1/2$, the set $S_h := (h, 1-h]^d$. Define further

$$I_0 := \Big\{ (\frac{i_1}{\bar{n}_0 h}, ..., \frac{i_d}{\bar{n}_0 h}) : -\lceil \bar{n}_0 h \rceil \leq i_1, ..., i_d \leq \lceil \bar{n}_0 h \rceil, i_1, ..., i_d \in \mathbb{Z} \Big\}$$

$$I_0' := \Big\{ (\frac{i_1}{\bar{n}_0 h}, ..., \frac{i_d}{\bar{n}_0 h}) : 0 \leq i_1, ..., i_d \leq \lceil \bar{n}_0 h \rceil, i_1, ..., i_d \in \mathbb{Z} \Big\}$$

Lemma A.2.1. *If $h < 1/2$ then for any $z_i \in S_h$,*

$$\sum_{k=1}^{n} K(\frac{z_k - z_i}{h}) = m \sum_{z_k' \in I_0} K(z_k')$$

Proof. Let

$$z_i := (\frac{i_1}{\bar{n}_0}, ..., \frac{i_d}{\bar{n}_0}) \in S_h,$$

then by the definition of S_h,

$$\bar{n}_0 h < i_j \leq \bar{n}_0 - \bar{n}_0 h,$$

for $1 \leq j \leq d$, thus,

$$i_j - \lceil \bar{n}_0 h \rceil \geq 1 \quad \text{and} \quad i_j + \lceil \bar{n}_0 h \rceil \leq \bar{n}_0.$$

By Assumption (iii), $K(x) = 0$ if $\|x\| > 1$, thus

$$\begin{aligned}
\sum_{k=1}^{n} K(\frac{z_k - z_i}{h}) =& m \sum_{k_1=1}^{\bar{n}_0} ... \sum_{k_d=1}^{\bar{n}_0} K\Big(\frac{k_1 - i_1}{\bar{n}_0 h}, ..., \frac{k_d - i_d}{\bar{n}_0 h}\Big) \\
=& m \sum_{k_1=i_1-\lceil \bar{n}_0 h \rceil}^{i_1+\lceil \bar{n}_0 h \rceil} ... \sum_{k_d=i_d-\lceil \bar{n}_0 h \rceil}^{i_d+\lceil \bar{n}_0 h \rceil} K\Big(\frac{k_1 - i_1}{\bar{n}_0 h}, ..., \frac{k_d - i_d}{\bar{n}_0 h}\Big) \\
=& m \sum_{k_1'=-\lceil \bar{n}_0 h \rceil}^{\lceil \bar{n}_0 h \rceil} ... \sum_{k_d'=-\lceil \bar{n}_0 h \rceil}^{\lceil \bar{n}_0 h \rceil} K\Big(\frac{k_1'}{\bar{n}_0 h}, ..., \frac{k_d'}{\bar{n}_0 h}\Big) = m \sum_{z_k' \in I_0} K(z_k').
\end{aligned}$$

□

Lemma A.2.2. *If $h < 1/2$, then for any $j \in \{1, ..., n\}$,*

$$\sum_{i=1}^{n} w_{nj}(z_i, h) \leq 2^d.$$

Proof. We show first for any $i \in \{1, ..., n\}$

$$m \sum_{z'_k \in I'_0} K(z'_k) \leq \sum_{k=1}^{n} K(\frac{z_k - z_i}{h}) \leq m \sum_{z'_k \in I_0} K(z'_k). \tag{A.2.1}$$

Let

$$z_i := (\frac{i_1}{\bar{n}_0}, ..., \frac{i_d}{\bar{n}_0})$$

with $i_1, ..., i_d \in \{1, ..., \bar{n}_0\}$. Since $K(x) = 0$ for $\|x\| > 1$, we get

$$\begin{aligned}
\sum_{k=1}^{n} K(\frac{z_k - z_i}{h}) =& m \sum_{k_1=1}^{\bar{n}_0} ... \sum_{k_d=1}^{\bar{n}_0} K\Big(\frac{k_1 - i_1}{\bar{n}_0 h}, ..., \frac{k_d - i_d}{\bar{n}_0 h}\Big) \\
\leq& m \sum_{k_1=i_1-\lceil \bar{n}_0 h \rceil}^{i_1+\lceil \bar{n}_0 h \rceil} ... \sum_{k_d=i_d-\lceil \bar{n}_0 h \rceil}^{i_d+\lceil \bar{n}_0 h \rceil} K\Big(\frac{k_1 - i_1}{\bar{n}_0 h}, ..., \frac{k_d - i_d}{\bar{n}_0 h}\Big) \\
=& m \sum_{k'_1=-\lceil \bar{n}_0 h \rceil}^{\lceil \bar{n}_0 h \rceil} ... \sum_{k'_d=-\lceil \bar{n}_0 h \rceil}^{\lceil \bar{n}_0 h \rceil} K\Big(\frac{k'_1}{\bar{n}_0 h}, ..., \frac{k'_d}{\bar{n}_0 h}\Big) = m \sum_{z'_k \in I_0} K(z'_k).
\end{aligned}$$

For the other inequality of (A.2.1), we denote that

$$I_{z_i} := \big\{j : i_j \leq \lceil \bar{n}_0 h \rceil, j \in \{1, ..., d\}\big\} \quad \text{and} \quad I_{z_i}^c = \{1, ..., d\} \backslash I_{z_i}.$$

By Assumption (iii), $K(x) = K(|x|)$ for all $x \in \mathbb{R}^d$, thus, for $h < 1/2$,

$$\begin{aligned}
\sum_{k=1}^{n} K(\frac{z_k - z_i}{h}) =& m \sum_{k_1=1}^{\bar{n}_0} ... \sum_{k_d=1}^{\bar{n}_0} K\Big(\frac{k_1 - i_1}{\bar{n}_0 h}, ..., \frac{k_d - i_d}{\bar{n}_0 h}\Big) \\
\geq& m \sum_{j \in I_{z_i}} \sum_{k_j=i_j}^{i_j+\lceil \bar{n}_0 h \rceil} \sum_{l \in I_{z_i}^c} \sum_{k_l=i_l-\lceil \bar{n}_0 h \rceil}^{i_l} K\Big(\frac{k_1 - i_1}{\bar{n}_0 h}, ..., \frac{k_d - i_d}{\bar{n}_0 h}\Big) \\
=& m \sum_{j \in I_{z_i}} \sum_{k_j=i_j}^{i_j+\lceil \bar{n}_0 h \rceil} \sum_{l \in I_{z_i}^c} \sum_{k_l=i_l}^{i_l+\lceil \bar{n}_0 h \rceil} K\Big(\frac{k_1 - i_1}{\bar{n}_0 h}, ..., \frac{k_d - i_d}{\bar{n}_0 h}\Big) \\
=& m \sum_{k_1=i_1}^{i_1+\lceil \bar{n}_0 h \rceil} ... \sum_{k_d=i_d}^{i_d+\lceil \bar{n}_0 h \rceil} K\Big(\frac{k_1 - i_1}{\bar{n}_0 h}, ..., \frac{k_d - i_d}{\bar{n}_0 h}\Big) \\
=& m \sum_{k'_1=0}^{\lceil \bar{n}_0 h \rceil} ... \sum_{k'_d=0}^{\lceil \bar{n}_0 h \rceil} K\Big(\frac{k'_1}{\bar{n}_0 h}, ..., \frac{k'_d}{\bar{n}_0 h}\Big) = m \sum_{z'_k \in I'_0} K(z'_k).
\end{aligned}$$

Thus, Inequality (A.2.1) holds. Hence, for any $j \in \{1, ..., n\}$,

$$\sum_{i=1}^{n} w_{nj}(z_i, h) = \sum_{i=1}^{n} \frac{K(\frac{z_j - z_i}{h})}{\sum_{k=1}^{n} K(\frac{z_k - z_i}{h})} \leq \frac{m \sum_{z_i' \in I_0} K(z_i')}{m \sum_{z_k' \in I_0'} K(z_k')} \leq 2^d.$$

□

Lemma A.2.3. *Let* $(\psi_n)_{n \in m \cdot \mathbb{N}} : [0,1]^d \to \mathbb{R}$ *be a sequence of uniformly bounded functions, i.e. there exists a constant* $C > 0$ *such that* $|\psi_n(z)| < C$ *for all* $z \in [0,1]^d$ *and* $n \in m \cdot \mathbb{N}$*, then*

$$\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{nj}(z_i, h) \psi_n(z_j) = \frac{1}{n} \sum_{j=1}^{n} \psi_n(z_j) + o(1).$$

Proof. Since $h \to 0$ and the assertion deals with a convergence property, we assume $h < 1/4$ in this proof. We write first

$$\begin{aligned} &\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{nj}(z_i, h) \psi_n(z_j) \\ =&\frac{1}{n} \sum_{z_i \notin S_h} \sum_{j=1}^{n} w_{nj}(z_i, h) \psi_n(z_j) + \frac{1}{n} \sum_{z_i \in S_h} \sum_{z_j \in S_{2h}} w_{nj}(z_i, h) \psi_n(z_j) \\ &+ \frac{1}{n} \sum_{z_i \in S_h} \sum_{z_j \notin S_{2h}} w_{nj}(z_i, h) \psi_n(z_j) \\ =:&Q_{1n} + Q_{2n} + Q_{3n}. \end{aligned}$$

By the uniform boundedness of the functions $(\psi_n)_{n \in m \cdot \mathbb{N}}$, there exists a constant $C > 0$, such that

$$|Q_{1n}| \leq C \cdot \frac{1}{n} \sum_{z_i \notin S_h} \sum_{j=1}^{n} w_{nj}(z_i, h) = C \cdot \frac{1}{n} \sum_{z_i \notin S_h} 1.$$

Note that S_h has at least $m \cdot (\bar{n}_0 - 2 \cdot \lceil \bar{n}_0 h \rceil - 1)^d$ points of $z_1, ..., z_n$ in it. Thus, by Assumption (ii) that $h \to 0$, we get

$$\begin{aligned} \frac{1}{n} \sum_{z_i \notin S_h} 1 \leq& \frac{1}{n} \cdot \left(n - m \cdot (\bar{n}_0 - 2 \cdot \lceil \bar{n}_0 h \rceil - 1)^d\right) \\ =& 1 - \left(1 - \frac{2 \cdot \lceil \bar{n}_0 h \rceil}{\bar{n}_0} - \frac{1}{\bar{n}_0}\right)^d = o(1). \end{aligned} \tag{A.2.2}$$

Hence, $Q_{1n} = o(1)$. Analogously, by Lemma A.2.2 and the uniform boundedness of the functions ψ_n, there exists a constant $C > 0$ such that

$$|Q_{3n}| \leq C \cdot \frac{1}{n} \sum_{z_j \notin S_{2h}} \sum_{i=1}^{n} w_{nj}(z_i, h) \leq C \cdot \frac{1}{n} \sum_{z_j \notin S_{2h}} 2^d = o(1),$$

consequently, $Q_{3n} = o(1)$ as well. Note that

$$Q_{2n} = \frac{1}{n} \sum_{z_j \in S_{2h}} \sum_{z_i \in S_h} \frac{K(\frac{z_j - z_i}{h})}{\sum_{k=1}^{n} K(\frac{z_k - z_i}{h})} \psi_n(z_j),$$

by Lemma A.2.1, we can write

$$Q_{2n} = \frac{1}{n} \sum_{z_j \in S_{2h}} \sum_{z_i \in S_h} \frac{K(\frac{z_j - z_i}{h})}{m \sum_{z'_k \in I_0} K(z'_k)} \cdot \psi_n(z_j).$$

Analogously to Lemma A.2.1, by Assumption (iii) $K(|x|) = K(x)$, for all $x \in [0,1]^d$, if $h < 1/4$, we can show for each $z_j \in S_{2h}$,

$$\sum_{z_i \in S_h} K(\frac{z_j - z_i}{h}) = \sum_{z_i \in S_h} K(\frac{z_i - z_j}{h}) = m \sum_{z'_i \in I_0} K(z'_i).$$

Hence,

$$\begin{aligned} Q_{2n} =& \frac{1}{n} \sum_{z_j \in S_{2h}} \frac{m \sum_{z'_i \in I_0} K(z'_i)}{m \sum_{z'_k \in I_0} K(z'_k)} \cdot \psi_n(z_j) \\ =& \frac{1}{n} \sum_{z_j \in S_{2h}} \psi_n(z_j) = \frac{1}{n} \sum_{j=1}^{n} \psi_n(z_j) - \frac{1}{n} \sum_{z_j \notin S_{2h}} \psi_n(z_j) = \frac{1}{n} \sum_{j=1}^{n} \psi_n(z_j) + o(1) \end{aligned} \tag{A.2.3}$$

where the last step can be shown analogously to (A.2.2). Therefore, the assertion follows. □

Lemma A.2.4. *Under the conditions of Lemma A.2.3, we get*

$$\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} w_{nj}(z_i, h) w_{nj}(z_k, h) \psi_n(z_j) = \frac{1}{n} \sum_{j=1}^{n} \psi_n(z_j) + o(1)$$

Proof. As in the proof of Lemma A.2.3, we assume $h < 1/4$. By the uniform boundedness of functions $(\psi_n)_{n \in m \cdot \mathbb{N}}$, there exists a $C > 0$, such that

$$\begin{aligned} &\Big| \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} w_{nj}(z_i, h) w_{nj}(z_k, h) \psi_n(z_j) \\ &\qquad - \frac{1}{n} \sum_{z_i \in S_h} \sum_{z_j \in S_{2h}} \sum_{z_k \in S_h} w_{nj}(z_i, h) w_{nj}(z_k, h) \psi_n(z_j) \Big| \\ \le& C \frac{1}{n} \sum_{z_i \notin S_h} \sum_{j=1}^{n} \sum_{k=1}^{n} w_{nj}(z_i, h) w_{nj}(z_k, h) \end{aligned}$$

$$
\begin{aligned}
&+C\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{z_k\notin S_h} w_{nj}(z_i,h)w_{nj}(z_k,h)\\
&+C\frac{1}{n}\sum_{i=1}^{n}\sum_{z_j\notin S_{2h}}\sum_{k=1}^{n} w_{nj}(z_i,h)w_{nj}(z_k,h)\\
=:&\bar{Q}_{1n}+\bar{Q}_{2n}+\bar{Q}_{3n}.
\end{aligned}
$$

Analogously to (A.2.3), we have

$$
\begin{aligned}
&\frac{1}{n}\sum_{z_i\in S_h}\sum_{z_j\in S_{2h}}\sum_{z_k\in S_h} w_{nj}(z_i,h)w_{nj}(z_k,h)\psi_n(z_j)\\
=&\frac{1}{n}\sum_{z_j\in S_{2h}}\sum_{z_i\in S_h}\sum_{z_k\in S_h}\frac{K(\frac{z_j-z_i}{h})}{m\sum_{z'_s\in I_0}K(z'_s)}\cdot\frac{K(\frac{z_j-z_k}{h})}{m\sum_{z'_l\in I_0}K(z'_l)}\cdot\psi_n(z_j)\\
=&\frac{1}{n}\sum_{z_j\in S_{2h}}\frac{\sum_{z'_i\in I_0}K(z'_i)}{\sum_{z'_s\in I_0}K(z'_s)}\cdot\frac{\sum_{z'_k\in I_0}K(z'_k)}{\sum_{z'_l\in I_0}K(z'_l)}\cdot\psi_n(z_j)\\
=&\frac{1}{n}\sum_{z_j\in S_{2h}}\psi_n(z_j)=\frac{1}{n}\sum_{j=1}^{n}\psi_n(z_j)+o(1).
\end{aligned}
$$

Further, by Lemma A.2.2,

$$
\begin{aligned}
\bar{Q}_{1n}=&C\cdot\frac{1}{n}\sum_{z_i\notin S_h}\sum_{j=1}^{n}w_{nj}(z_i,h)\Big(\sum_{k=1}^{n}w_{nj}(z_k,h)\Big)\\
\le&2^d\cdot C\cdot\frac{1}{n}\sum_{z_i\notin S_h}\sum_{j=1}^{n}w_{nj}(z_i,h)=2^d\cdot C\cdot\frac{1}{n}\sum_{z_i\notin S_h}1=o(1).
\end{aligned}
$$

where the last step follows from (A.2.2). With similar arguments, it can be shown that $\bar{Q}_{2n}$ and $\bar{Q}_{3n}$ are both equal to $o(1)$ as well. Hence, the assertion follows. □

Lemma A.2.5. *For any $j\in\{1,...,n\}$, $r\in\mathbb{N}$ and $z\in[0,1]^d$,*

$$w_{nj}(z,h)\cdot\left\|z-z_j\right\|^r\le w_{nj}(z,h)\cdot h^r.$$

Proof. Notice that for any $j\in\{1,...,n\}$, $w_{nj}(z,h)$ is always non-negative, hence if $\|z-z_j\|\le h$,

$$w_{nj}(z,h)\cdot\|z-z_j\|^r\le w_{nj}(z,h)\cdot h^r.$$

If $\|z-z_j\|>h$, by Assumption (iii) $K(x)=0$ for $\|x\|>1$, we get

$$w_{nj}(z,h)=\frac{K(\frac{z_j-z}{h})}{\sum_{k=1}^{n}K(\frac{z_k-z}{h})}=0.$$

Therefore,

$$w_{nj}(z,h) \cdot \|z - z_j\|^r = 0 = w_{nj}(z,h) \cdot h^r.$$

□

Lemma A.2.6. *There exists a constant $C > 0$ such that for all $(x,z) \in \mathbb{R} \times [0,1]^d$ and eventual all $n \in m \cdot \mathbb{N}$,*

$$\big|E[\hat{H}_n(x|z)] - H(x|z)\big| \leq Ch, \tag{A.2.4}$$

and for all $(x, z_i) \in \mathbb{R} \times S_h$ and eventual all $n \in m \cdot \mathbb{N}$,

$$\big|E[\hat{H}_n(x|z_i)] - H(x|z_i)\big| \leq Ch^2 \tag{A.2.5}$$

Proof. By a Taylor expansion, we get

$$\begin{aligned}
E[\hat{H}_n(x|z)] - H(x|z) =& \sum_{j=1}^n w_{nj}(z,h) H(x|z_j) - H(x|z) \\
=& \sum_{j=1}^n w_{nj}(z,h) \Big(\frac{\partial H(x|z)}{\partial z}\Big)^T \cdot (z_j - z) \\
&+ \sum_{j=1}^n w_{nj}(z,h)(z_j - z)^T \cdot \frac{\partial^2 H(x|\tilde{z}_j)}{\partial z^2} \cdot (z_j - z)
\end{aligned}$$

where $\tilde{z}_j$ lies between z and z_j. By the boundedness of the partial derivative and the Hessian matrix of H with respect to z (Assumption (i)) and Lemma A.2.5, there exists a constant $C > 0$ independent of x, z and z_j, such that

$$\begin{aligned}
&\Big|\sum_{j=1}^n w_{nj}(z,h)(z_j - z)^T \cdot \frac{\partial^2 H(x|\tilde{z}_j)}{\partial z^2} \cdot (z_j - z)\Big| \\
\leq& C \sum_{j=1}^n w_{nj}(z,h) \big\|z_j - z\big\|^2 \leq Ch^2 \cdot \sum_{j=1}^n w_{nj}(z,h) = Ch^2.
\end{aligned}$$

Analogously, for all $(x,z) \in \mathbb{R} \times [0,1]^d$ and eventual all $n \in m \cdot \mathbb{N}$, there exists a $C > 0$ such that

$$\begin{aligned}
&\Big|\sum_{j=1}^n w_{nj}(z,h)\Big(\frac{\partial H(x|z)}{\partial z}\Big)^T \cdot (z_j - z)\Big| \\
\leq& C \sum_{j=1}^n w_{nj}(z,h) \|z_j - z\| \leq Ch \cdot \sum_{j=1}^n w_{nj}(z,h) = Ch.
\end{aligned}$$

Hence, by Assumption (ii) $h \to 0$, (A.2.4) holds.

If $z = z_i \in S_h$, analogously to Lemma A.2.1 we can show that

$$\sum_{j=1}^{n} K(\frac{z_j - z_i}{h}) \cdot (z_j - z_i) = m \sum_{z'_j \in I_0} K(z'_j) \cdot z'_j.$$

By the definition of I_0 and Assumption (iii) $K(|x|) = K(x)$ for all $x \in \mathbb{R}^d$, thus

$$m \sum_{z'_j \in I_0} K(z'_j) \cdot z'_j = m \sum_{-z'_j \in I_0} K(-z'_j) \cdot (-z'_j) = -m \sum_{z'_j \in I_0} K(z'_j) \cdot z'_j.$$

Hence, for $z_i \in S_h$,

$$\sum_{j=1}^{n} K(\frac{z_j - z_i}{h}) \cdot (z_j - z_i) = m \sum_{z'_j \in I_0} K(z'_j) \cdot z'_j = 0.$$

Consequently, for $z_i \in S_h$,

$$\begin{aligned} &\sum_{j=1}^{n} w_{nj}(z_i, h) \Big(\frac{\partial H(x|z_i)}{\partial z} \Big)^T \cdot (z_j - z_i) \\ =& \Big(\frac{\partial H(x|z_i)}{\partial z} \Big)^T \cdot \sum_{j=1}^{n} w_{nj}(z_i, h)(z_j - z_i) \\ =& \Big(\frac{\partial H(x|z_i)}{\partial z} \Big)^T \cdot \frac{\sum_{j=1}^{n} K(\frac{z_j - z_i}{h})(z_j - z_i)}{\sum_{k=1}^{n} K(\frac{z_k - z_i}{h})} = 0. \end{aligned}$$

Hence, (A.2.5) holds. □

Lemma A.2.7. *Let $\psi : \mathbb{R} \times [0,1]^d \to \mathbb{R}$ be an H square integrable function independent of z, then*

$$\int \psi(x,z) dQ_n(x,z) - \int \psi(x,z) dQ(x,z) = o_p(1). \qquad \text{(A.2.6)}$$

Proof. By definition

$$\begin{aligned} E\Big[\int \psi(x,z) dQ_n(x,z) \Big] =& \frac{1}{n} \sum_{i=1}^{n} E\big[\psi(X_i, z_i) \big] \\ =& \frac{1}{n} \sum_{i=1}^{n} \int \psi(x, z_i) dH(x|z_i) = \frac{1}{n_0} \sum_{i=1}^{n_0} \int \psi(x, z_i) dH(x|z_i). \end{aligned}$$

Hence, by the convergence of the Riemann sum,

$$E\Big[\int \psi(x,z) dQ_n(x,z) \Big] \to \int \psi(x,z) dQ(x,z).$$

And the variance

$$\begin{aligned} Var\Big[\int \psi(x,z)dQ_n(x,z)\Big] =& \frac{1}{n^2}\sum_{i=1}^n Var\big[\psi(X_i,z_i)\big] \\ \le& \frac{1}{n^2}\sum_{i=1}^n E\big[\psi^2(X_i,z_i)\big] \\ =& \frac{1}{n^2}\sum_{i=1}^n \int \psi^2(x,z_i)dH(x|z_i) \to 0. \end{aligned}$$

Thus, the assertion follows from Chebyshev's inequality. □

Lemma A.2.8. *Let $(\psi_{1n})_{n\in m\cdot\mathbb{N}}, (\psi_{2n})_{n\in m\cdot\mathbb{N}} : \mathbb{R}\times[0,1]^d \to \mathbb{R}$ be two sequences of functions. Assume that there exists a constant $\delta > 0$, such that for any $n \in m\cdot\mathbb{N}$, $|\psi_{1n}|^{2+\delta}$, $|\psi_{2n}|^{2+\delta}$ and $\|\partial\psi_{1n}/\partial z\|^2$ are dominated by the same H integrable function independent of z. Define for $n \in m\cdot\mathbb{N}$*

$$\begin{aligned} \sigma_n^2 :=& \frac{1}{n}\sum_{i=1}^n \int \big(\psi_{1n}(x,z_i)+\psi_{2n}(x,z_i)\big)^2 dH(x|z_i) \\ &- \frac{1}{n}\sum_{i=1}^n \Big(\int \big(\psi_{1n}(x,z_i)+\psi_{2n}(x,z_i)\big)dH(x|z_i)\Big)^2. \end{aligned}$$

If there exits a constant σ such that $\sigma_n^2 \to \sigma^2$, then

$$\begin{aligned} &\sqrt{n}\int \psi_{1n}(x,z)d\big(\hat{Q}_n(x,z) - E[\hat{Q}_n(x,z)]\big) \\ &\qquad + \sqrt{n}\int \psi_{2n}(x,z)d\big(Q_n(x,z) - E[Q_n(x,z)]\big) \to N(0,\sigma^2). \end{aligned}$$

Proof. We denote for each $i \in \{1,...,n\}$,

$$\tilde{\psi}_n(X_i,z_i) := \sum_{j=1}^n w_{ni}(z_j,h)\psi_{1n}(X_i,z_j) + \psi_{2n}(X_i,z_i).$$

Note that $\tilde{\psi}_n(X_1,z_1),...,\tilde{\psi}_n(X_n,z_n)$ are independent,

$$\begin{aligned} &\sqrt{n}\int \psi_{1n}(x,z)d\hat{Q}_n(x,z) + \sqrt{n}\int \psi_{2n}(x,z)dQ_n(x,z) \\ =& \sqrt{n}\cdot\frac{1}{n}\sum_{i=1}^n\Big(\sum_{j=1}^n w_{ni}(z_j,h)\psi_{1n}(X_i,z_j) + \psi_{2n}(X_i,z_i)\Big) = \sqrt{n}\cdot\frac{1}{n}\sum_{i=1}^n \tilde{\psi}_n(X_i,z_i) \end{aligned}$$

and

$$\sqrt{n}\int \psi_{1n}(x,z)dE[\hat{Q}_n(x,z)] + \sqrt{n}\int \psi_{2n}(x,z)dE[Q_n(x,z)] = \sqrt{n}\cdot\frac{1}{n}\sum_{i=1}^n E[\tilde{\psi}_n(X_i,z_i)].$$

In the sequel, we show first that

$$\frac{1}{\sigma_n^{2+\delta}}\sum_{i=1}^{n}E\Big[\Big|n^{-1/2}\tilde{\psi}_n(X_i,z_i)-E\big[n^{-1/2}\tilde{\psi}_n(X_i,z_i)\big]\Big|^{2+\delta}\Big]=o(1). \tag{A.2.7}$$

By the Minkowski's inequality,

$$\begin{aligned}
&\frac{1}{\sigma_n^{2+\delta}}\sum_{i=1}^{n}E\Big[\Big|n^{-1/2}\tilde{\psi}_n(X_i,z_i)-E\big[n^{-1/2}\tilde{\psi}_n(X_i,z_i)\big]\Big|^{2+\delta}\Big]\\
=&\frac{n^{-\delta/2}}{\sigma_n^{2+\delta}}\frac{1}{n}\sum_{i=1}^{n}E\Big[\Big|\tilde{\psi}_n(X_i,z_i)-E\big[\tilde{\psi}_n(X_i,z_i)\big]\Big|^{2+\delta}\Big]\\
\leq&\frac{n^{-\delta/2}}{\sigma_n^{2+\delta}}\frac{1}{n}\sum_{i=1}^{n}\Big(\big(E\big[\big|\tilde{\psi}_n(X_i,z_i)\big|^{2+\delta}\big]\big)^{\frac{1}{2+\delta}}+\big(E\big[\big|E[\tilde{\psi}_n(X_i,z_i)]\big|^{2+\delta}\big]\big)^{\frac{1}{2+\delta}}\Big)^{2+\delta}\\
=&\frac{n^{-\delta/2}}{\sigma_n^{2+\delta}}\frac{1}{n}\sum_{i=1}^{n}\Big(\big(E\big[\big|\tilde{\psi}_n(X_i,z_i)\big|^{2+\delta}\big]\big)^{\frac{1}{2+\delta}}+\big|E\big[\tilde{\psi}_n(X_i,z_i)\big]\big|\Big)^{2+\delta}.
\end{aligned}$$

Note that by Jensen's inequality for any $i\in\{1,...,n\}$,

$$\big|E\big[\tilde{\psi}_n(X_i,z_i)\big]\big|\leq E\big[\big|\tilde{\psi}_n(X_i,z_i)\big|^{2+\delta}\big]^{\frac{1}{2+\delta}}.$$

Hence,

$$\begin{aligned}
&\frac{1}{\sigma_n^{2+\delta}}\sum_{i=1}^{n}E\Big[\Big|n^{-1/2}\tilde{\psi}_n(X_i,z_i)-E\big[n^{-1/2}\tilde{\psi}_n(X_i,z_i)\big]\Big|^{2+\delta}\Big]\\
\leq&2^{2+\delta}\frac{n^{-\delta/2}}{\sigma_n^{2+\delta}}\frac{1}{n}\sum_{i=1}^{n}E\big[\big|\tilde{\psi}_n(X_i,z_i)\big|^{2+\delta}\big]\\
=&2^{2+\delta}\frac{n^{-\delta/2}}{\sigma_n^{2+\delta}}\frac{1}{n}\sum_{i=1}^{n}\int\big|\tilde{\psi}_n(x,z_i)\big|^{2+\delta}dH(x|z_i).
\end{aligned} \tag{A.2.8}$$

Further by the Minkowski's inequality, for any $i\in\{1,...,n\}$,

$$\begin{aligned}
&\Big(\int\big|\tilde{\psi}_n(x,z_i)\big|^{2+\delta}dH(x|z_i)\Big)^{\frac{1}{2+\delta}}\\
=&\Big(\int\big|\sum_{j=1}^{n}w_{ni}(z_j,h)\psi_{1n}(x,z_j)+\psi_{2n}(x,z_i)\big|^{2+\delta}dH(x|z_i)\Big)^{\frac{1}{2+\delta}}\\
\leq&\sum_{j=1}^{n}\Big(\int\big|w_{ni}(z_j,h)\psi_{1n}(x,z_j)\big|^{2+\delta}dH(x|z_i)\Big)^{\frac{1}{2+\delta}}+\Big(\int\big|\psi_{2n}(x,z_i)\big|^{2+\delta}dH(x|z_i)\Big)^{\frac{1}{2+\delta}}\\
=&\sum_{j=1}^{n}w_{ni}(z_j,h)\Big(\int\big|\psi_{1n}(x,z_j)\big|^{2+\delta}dH(x|z_i)\Big)^{\frac{1}{2+\delta}}+\Big(\int\big|\psi_{2n}(x,z_i)\big|^{2+\delta}dH(x|z_i)\Big)^{\frac{1}{2+\delta}}.
\end{aligned}$$

By the assumptions of this lemma, there exists a constant $C > 0$, such that the last term is bounded by

$$C \cdot \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ni}(z_j, h) + C = 2C.$$

Therefore, Inequality (A.2.7) follows from (A.2.8).

Next we show that

$$Var\Big[\sqrt{n} \int \psi_{1n}(x,z) d\hat{Q}_n(x,z) + \sqrt{n} \int \psi_{2n}(x,z) dQ_n(x,z)\Big] = \sigma^2 + o(1).$$

Note that

$$\begin{aligned}
& Cov\Big[\sqrt{n} \int \psi_{1n}(x,z) d\hat{Q}_n(x,z), \sqrt{n} \int \psi_{2n}(x,z) dQ_n(x,z)\Big] \\
=& \frac{1}{n} Cov\Big[\sum_{i=1}^{n} \sum_{j=1}^{n} w_{nj}(z_i,h) \psi_{1n}(X_j, z_i), \sum_{k=1}^{n} \psi_{2n}(X_k, z_k)\Big] \\
=& \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} w_{nj}(z_i,h) Cov\Big[\psi_{1n}(X_j, z_i), \psi_{2n}(X_k, z_k)\Big] \\
=& \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{nj}(z_i,h) Cov\Big[\psi_{1n}(X_j, z_i), \psi_{2n}(X_j, z_j)\Big] \\
=& \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{nj}(z_i,h) \Big(\int \psi_{1n}(x, z_i) \psi_{2n}(x, z_j) dH(x|z_j) \\
& \qquad - \int \psi_{1n}(x, z_i) dH(x|z_j) \int \psi_{2n}(x, z_j) dH(x|z_j) \Big).
\end{aligned}$$

Note that by Cauchy-Schwarz's inequality and Lemma A.2.5, there exist a constant $C > 0$ and z_{ij} lying between z_i and z_j such that

$$\begin{aligned}
& \Big| \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{nj}(z_i,h) \int \big(\psi_{1n}(x,z_i) - \psi_{1n}(x,z_j)\big) \psi_{2n}(x,z_j) dH(x|z_j) \Big| \\
\le & \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{nj}(z_i,h) \int \big|\psi_{1n}(x,z_i) - \psi_{1n}(x,z_j)\big| \cdot |\psi_{2n}(x,z_j)| dH(x|z_j) \\
\le & p \cdot \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{nj}(z_i,h) \|z_i - z_j\| \int \Big\| \frac{\psi_{1n}(x, z_{ij})}{\partial z} \Big\| \cdot |\psi_{2n}(x,z_j)| dH(x|z_j) \\
\le & p \cdot \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{nj}(z_i,h) \|z_i - z_j\| \Big(\int \Big\| \frac{\psi_{1n}(x, z_{ij})}{\partial z} \Big\|^2 dH(x|z_j) \cdot \int |\psi_{2n}(x,z_j)|^2 dH(x|z_j) \Big)^{1/2} \\
\le & C \cdot \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{nj}(z_i,h) \|z_i - z_j\| \le Ch \cdot \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{nj}(z_i,h) = Ch.
\end{aligned}$$

Therefore, we get

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{nj}(z_i,h)\int \psi_{1n}(x,z_i)\psi_{2n}(x,z_j)dH(x|z_j)\\
=&\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{nj}(z_i,h)\int \psi_{1n}(x,z_j)\psi_{2n}(x,z_j)dH(x|z_j)+o(1)\\
=&\frac{1}{n}\sum_{j=1}^{n}\int \psi_{1n}(x,z_j)\psi_{2n}(x,z_j)dH(x|z_j)+o(1) \qquad \text{(A.2.9)}
\end{aligned}
$$

where the last step follows from Lemma A.2.3 with

$$\psi_n(z)=\int \psi_{1n}(x,z)\psi_{2n}(x,z)dH(x|z).$$

With similar arguments, we can show further

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{nj}(z_i,h)\int \psi_{1n}(x,z_i)dH(x|z_j)\int \psi_{2n}(x,z_j)dH(x|z_j)\\
=&\frac{1}{n}\sum_{j=1}^{n}\int \psi_{1n}(x,z_j)dH(x|z_j)\cdot\int \psi_{2n}(x,z_j)dH(x|z_j)+o(1).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&Cov\Big[\sqrt{n}\int \psi_{1n}(x,z)d\hat{Q}_n(x,z),\sqrt{n}\int \psi_{2n}(x,z)dQ_n(x,z)\Big]\\
=&\frac{1}{n}\sum_{j=1}^{n}\Big(\int \psi_{1n}(x,z_j)\psi_{2n}(x,z_j)dH(x|z_j)\\
&\qquad -\int \psi_{1n}(x,z_j)dH(x|z_j)\cdot\int \psi_{2n}(x,z_j)dH(x|z_j)\Big)+o(1).
\end{aligned}
$$

Analogously,

$$
\begin{aligned}
&Var\Big[\sqrt{n}\cdot\int \psi_{1n}(x,z)d\hat{Q}_n(x,z)\Big]\\
=&Var\Big[\sqrt{n}\cdot\frac{1}{n}\sum_{i=1}^{n}\int \psi_{1n}(x,z_i)d\hat{H}_n(x|z_i)\Big]\\
=&\frac{1}{n}\cdot Cov\Big[\sum_{i=1}^{n}\sum_{j=1}^{n} w_{nj}(z_i,h)\psi_{1n}(X_j,z_i),\sum_{k=1}^{n}\sum_{l=1}^{n} w_{nl}(z_k,h)\psi_{1n}(X_l,z_k)\Big]\\
=&\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} w_{nj}(z_i,h)w_{nj}(z_k,h)Cov\Big[\psi_{1n}(X_j,z_i),\psi_{1n}(X_j,z_k)\Big]\\
=&\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} w_{nj}(z_i,h)w_{nj}(z_k,h)\Big(\int \psi_{1n}(x,z_i)\psi_{1n}(x,z_k)dH(x|z_j)
\end{aligned}
$$

$$
\begin{aligned}
&\qquad - \int \psi_{1n}(x, z_i) dH(x|z_j) \int \psi_{1n}(x, z_k) dH(x|z_j)\Big) \\
=&\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} w_{nj}(z_i, h) w_{nj}(z_k, h) \Big(\int \psi_{1n}(x, z_j) \psi_{1n}(x, z_j) dH(x|z_j) \\
&\qquad - \int \psi_{1n}(x, z_j) dH(x|z_j) \int \psi_{1n}(x, z_j) dH(x|z_j)\Big) \\
=&\frac{1}{n} \sum_{j=1}^{n} \Big(\int \psi_{1n}^2(x, z_j) dH(x|z_j) - \Big(\int \psi_{1n}(x, z_j) dH(x|z_j)\Big)^2 \Big) + o(1),
\end{aligned}
\tag{A.2.10}
$$

where the last step follows Lemma A.2.4 with

$$
\psi_n(z) = \int \psi_{1n}^2(x, z) dH(x|z) - \Big(\int \psi_{1n}(x, z) dH(x|z)\Big)^2.
$$

Further, by definition,

$$
\begin{aligned}
&Var\Big[\sqrt{n} \int \psi_{2n}(x, z) dQ_n(x, z)\Big] \\
=&\frac{1}{n} \sum_{i=1}^{n} \big(E[\psi_{2n}^2(X_i, z_i)] - E[\psi_{2n}(X_i, z_i)]^2\big) \\
=&\frac{1}{n} \sum_{i=1}^{n} \Big(\int \psi_{2n}^2(x, z_i) dH(x|z_i) - \Big(\int \psi_{2n}(x, z_i) dH(x|z_i)\Big)^2 \Big)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&Var\Big[\sqrt{n} \int \psi_{1n}(x, z) d\hat{Q}_n(x, z) + \sqrt{n} \int \psi_{2n}(x, z) dQ_n(x, z)\Big] \\
=&Var\Big[\sqrt{n} \int \psi_{1n}(x, z) d\hat{Q}_n(x, z)\Big] + Var\Big[\sqrt{n} \int \psi_{2n}(x, z) dQ_n(x, z)\Big] \\
&+ 2Cov\Big[\sqrt{n} \int \psi_{1n}(x, z) d\hat{Q}_n(x, z), \sqrt{n} \int \psi_{2n}(x, z) dQ_n(x, z)\Big] \\
=&\sigma_n^2 + o(1) = \sigma^2 + o(1).
\end{aligned}
$$

Therefore, the assertion follows from Lyapunov's central limit theorem. □

Next we extend Lemma A.1.1 to the case with kernel estimator for distribution function. Let (X_{ij}, z_j) for $i \in \{1, ..., k\}, j \in \{1, ..., n\}$ be defined as in Section A.1. The distribution function of X_{ij} is denoted by $H_i(\cdot|z_j)$. For each $i \in \{1, ..., k\}$, the kernel estimator for distribution function at $z \in [0, 1]^d$ is defined as

$$
\hat{H}_{in}(x|z) := \sum_{j=1}^{n} w_{nj}(z, h) I(X_{ij} \leq x).
$$

Further, for a sequence of functions $(\psi_n)_{n\in m\cdot\mathbb{N}} : \mathbb{R}^k \times [0,1]^d \to \mathbb{R}$ we denote for each $i \in \{1, ..., k\}$,

$$\begin{aligned}\hat{P}_{in}(z) :=& \sqrt{n}\int\cdots\int \psi_n(x_1, ..., x_k, z) dE\big[\hat{H}_{1n}(x_1|z)\big]\cdots dE\big[\hat{H}_{(i-1)n}(x_{i-1}|z)\big]\\ &\times d\hat{H}_{in}(x_i|z)dE\big[\hat{H}_{(i+1)n}(x_{i+1}|z)\big]\cdots dE\big[\hat{H}_{kn}(x_k|z)\big],\\ \bar{\hat{P}}_{in} :=& \frac{1}{n_0}\sum_{j=1}^{n_0}\hat{P}_{in}(z_j).\end{aligned}$$

Similarly, we define

$$\begin{aligned}\hat{P}_n(z) :=& \sqrt{n}\int\cdots\int \psi_n(x_1, ..., x_k, z) d\hat{H}_{1n}(x_1|z)\cdots d\hat{H}_{kn}(x_k|z),\\ \hat{P}_{0n}(z) :=& \sqrt{n}\int\cdots\int \psi_n(x_1, ..., x_k, z) dE\big[\hat{H}_{1n}(x_1|z)\big]\cdots dE\big[\hat{H}_{kn}(x_k|z)\big],\\ \bar{\hat{P}}_n :=& \frac{1}{n_0}\sum_{j=1}^{n_0}\hat{P}_n(z_j), \qquad \bar{\hat{P}}_{0n} := \frac{1}{n_0}\sum_{j=1}^{n_0}\hat{P}_{0n}(z_j).\end{aligned}$$

Lemma A.2.9. *Let C be a positive number such that*

$$E[\psi_n^2(X_{1i_1}, ..., X_{ki_k}, z)] \le C$$

for all $i_1, ..., i_k \in \{1, ..., n\}$, $n \in m\cdot\mathbb{N}$ and $z \in [0,1]^d$. Then

$$\bar{\hat{P}}_n = \sum_{i=1}^{k}\bar{\hat{P}}_{in} - (k-1)\bar{\hat{P}}_{0n} + o_p(1).$$

Proof. As in the proof of Lemma A.1.1, first we show that

$$Var\big[\bar{\hat{P}}_n - \sum_{i=1}^{k}\bar{\hat{P}}_{in} + (k-1)\bar{\hat{P}}_{0n}\big] = o(1). \tag{A.2.11}$$

Assume that $z, \tilde{z}$ are two arbitrary covariate values in $\{z_1, ..., z_{n_0}\}$. Let the sets I, I_1, I_2, I_3, J, J_1 I_{ab} and $\bar{I}_{ab}$ for $a, b \in \{1, ..., k\}$ be defined as in the proof of Lemma A.1.1. For any set $\tilde{I} \subset I$, we denote

$$\begin{aligned}C_{\tilde{I},n}(z,\tilde{z}) :=& n\sum_{\tilde{I}} w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)w_{nj_1}(\tilde{z},h)\cdots w_{nj_k}(\tilde{z},h)\\ &\times Cov\big[\psi_n(X_{1i_1}, ..., X_{ki_k}, z), \psi_n(X_{1j_1}, ..., X_{kj_k}, \tilde{z})\big].\end{aligned}$$

Note that by (A.2.1), for all $i \in \{1, ..., n\}$,

$$w_{ni}(z,h) = \frac{K(\frac{z_i - z}{h})}{\sum_{k=1}^{n}K(\frac{z_k - z}{h})} \le \frac{K(\frac{z_i - z}{h})}{m\sum_{z'_k\in I'_0}K(z'_k)}.$$

Since K is bounded and

$$\frac{1}{n_0 h^d}\sum_{z'_k\in I'_0} K(z'_k)\to\int_{[0,1]^d} K(x)dx,$$

for n large enough, there exists a constant $\tilde{C}>0$ such that

$$w_{ni}(z,h)\le\frac{\tilde{C}}{n_0h^d}. \tag{A.2.12}$$

Define

$$c_n := Cn(2k-1)(k-1)\cdot(2\lceil\bar{n}_0h\rceil+1)^{d(2k-2)}m^{2k-2}\big(\frac{\tilde{C}}{n_0h^d}\big)^{2k}.$$

Note the by Assumption (ii) $n_0h^{2d}\to\infty$ and $\bar{n}_0h\to\infty$, thus,

$$c_n = C\tilde{C}^{2k}(2k-1)(k-1)m^{2k-1}\cdot\big(\frac{2\lceil\bar{n}_0h\rceil+1}{\bar{n}_0h}\big)^{d(2k-2)}\cdot n_0^{-1}h^{-2d}=o(1).$$

In the following, we show that for $a=b=1$ and n large enough

$$\left|C_{\bar{I}_{ab},n}(z,\tilde{z})\right|\le c_n. \tag{A.2.13}$$

and

$$\left|C_{I_{ab},n}(z,\tilde{z})-Cov[\hat{P}_{an}(z),\hat{P}_{bn}(\tilde{z})]\right|\le c_n. \tag{A.2.14}$$

Note that if $(i_1,...,i_k,j_1,...,j_k)\in\bar{I}_{11}$ with

$$w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)w_{nj_1}(\tilde{z},h)\cdots w_{nj_k}(\tilde{z},h)\neq 0,$$

by Assumption (iii) $K(x)=0$ if $\|x\|>1$, it must hold that

$$\|z_{i_1}-z\|\le h,...,\|z_{i_k}-z\|\le h.$$

$$\|z_{j_1}-\tilde{z}\|\le h,...,\|z_{j_k}-\tilde{z}\|\le h.$$

and at least two numbers out of the $2k-1$ numbers $i_1,...,i_k,j_2,...,j_k$ are equal. Hence, there are at most

$$\binom{2k-1}{2}\cdot(2\lceil\bar{n}_0h\rceil+1)^{d(2k-2)}\cdot m^{2k-2}$$

elements in $\bar{I}_{11}$ with

$$w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)w_{nj_1}(\tilde{z},h)\cdots w_{nj_k}(\tilde{z},h)\neq 0.$$

Consequently, analogously to (A.1.7)

$$\begin{aligned}\left|C_{\bar{I}_{11},n}(z,\tilde{z})\right| \leq & Cn\sum_{\bar{I}_{11}} w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)w_{nj_1}(\tilde{z},h)\cdots w_{nj_k}(\tilde{z},h)\\ \leq & Cn\binom{2k-1}{2}\cdot(2\lceil\bar{n}_0h\rceil+1)^{d(2k-2)}\cdot m^{2k-2}\Big(\frac{\tilde{C}}{n_0h^d}\Big)^{2k}=c_n,\end{aligned}$$

i.e. Inequality (A.2.13) holds.

For Inequality (A.2.14), denote the function $\tilde{\psi}_n:\mathbb{R}\times[0,1]^{kd}\to\mathbb{R}$ with

$$\tilde{\psi}_n(x,z_{i_1},z_{i_2},...,z_{i_k}):=\int\cdots\int\psi_n(x,x_2,...,x_k,z_{i_1})dH_2(x_2|z_{i_2})\cdots dH_k(x_k|z_{i_k}).$$

Note that if $(i_1,...,i_k,j_1,...,j_k)\in I_{11}$, by Fubini's theorem and independence of $X_{1i_1},...,X_{ki_k},X_{1j_2},...,X_{kj_k}$

$$\begin{aligned}&Cov\big[\psi_n(X_{1i_1},X_{2i_2},...,X_{ki_k},z),\psi_n(X_{1j_1},X_{2j_2},...,X_{kj_k},\tilde{z})\big]\\ =&Cov\big[\psi_n(X_{1i_1},X_{2i_2},...,X_{ki_k},z),\psi_n(X_{1i_1},X_{2j_2},...,X_{kj_k},\tilde{z})\big]\\ =&E\big[\psi_n(X_{1i_1},X_{2i_2},...,X_{ki_k},z)\cdot\psi_n(X_{1i_1},X_{2j_2},...,X_{kj_k},\tilde{z})\big]\\ &\quad-E\big[\psi_n(X_{1i_1},X_{2i_2},...,X_{ki_k},z)\big]\cdot E\big[\psi_n(X_{1i_1},X_{2j_2},...,X_{kj_k},\tilde{z})\big]\\ =&E\big[\tilde{\psi}_n(X_{1i_1},z,z_{i_2},...,z_{i_k})\cdot\tilde{\psi}_n(X_{1i_1},\tilde{z},z_{j_2},...,z_{j_k})\big]\\ &\quad-E\big[\tilde{\psi}_n(X_{1i_1},z,z_{i_2},...,z_{i_k})\big]\cdot E\big[\tilde{\psi}_n(X_{1i_1},\tilde{z},z_{j_2},...,z_{j_k})\big]\\ =&Cov\big[\tilde{\psi}_n(X_{1i_1},z,z_{i_2},...,z_{i_k}),\tilde{\psi}_n(X_{1i_1},\tilde{z},z_{j_2},...,z_{j_k})\big].\end{aligned}$$

Therefore,

$$\begin{aligned}C_{I_{11},n}(z,\tilde{z})=&n\sum_{I_{11}} w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)w_{nj_1}(\tilde{z},h)\cdots w_{nj_k}(\tilde{z},h)\\ &\times Cov\big[\tilde{\psi}_n(X_{1i_1},z,z_{i_2},...,z_{i_k}),\tilde{\psi}_n(X_{1i_1},\tilde{z},z_{j_2},...,z_{j_k})\big].\end{aligned}$$

Further, by the independence of X_{1i_1} and X_{1j_1} providing $i_1\neq j_1$,

$$\begin{aligned}&Cov\big[\hat{P}_{1n}(z),\hat{P}_{1n}(\tilde{z})\big]\\ =&n\sum_{i_k=1}^{n}\cdots\sum_{i_1=1}^{n}\sum_{j_k=1}^{n}\cdots\sum_{j_1=1}^{n} w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)w_{nj_1}(\tilde{z},h)\cdots w_{nj_k}(\tilde{z},h)\\ &\times Cov\big[\tilde{\psi}_n(X_{1i_1},z,z_{i_2},...,z_{i_k}),\tilde{\psi}_n(X_{1j_1},\tilde{z},z_{j_2},...,z_{j_k})\big]\\ =&n\sum_{I_{11}\cup\bar{I}_{11}} w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)w_{nj_1}(\tilde{z},h)\cdots w_{nj_k}(\tilde{z},h)\\ &\times Cov\big[\tilde{\psi}_n(X_{1i_1},z,z_{i_2},...,z_{i_k}),\tilde{\psi}_n(X_{1i_1},\tilde{z},z_{j_2},...,z_{j_k})\big].\end{aligned}$$

Thus, analogously to (A.1.12)

$$
\begin{aligned}
&\Big|C_{I_{11},n}(z,\tilde z) - Cov\big[\hat P_{1n}(z), \hat P_{1n}(\tilde z)\big]\Big| \\
=&\Big|n\sum_{\bar I_{11}} w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)w_{nj_1}(\tilde z,h)\cdots w_{nj_k}(\tilde z,h) \\
&\times Cov\big[\tilde\psi_n(X_{1i_1}, z, z_{i_2},...,z_{i_k}), \tilde\psi_n(X_{1i_1}, \tilde z, z_{j_2},...,z_{j_k})\big]\Big| \\
\le& n\sum_{\bar I_{11}} w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)w_{nj_1}(\tilde z,h)\cdots w_{nj_k}(\tilde z,h) \\
&\times \Big|Cov\big[\tilde\psi_n(X_{1i_1}, z, z_{i_2},...,z_{i_k}), \tilde\psi_n(X_{1i_1}, \tilde z, z_{j_2},...,z_{j_k})\big]\Big| \\
\le& Cn\sum_{\bar I_{11}} w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)w_{nj_1}(\tilde z,h)\cdots w_{nj_k}(\tilde z,h) \le c_n,
\end{aligned}
$$

i.e. Inequality (A.2.14) holds. Then with the same arguments used in the proof of Lemma A.1.1, Equality (A.2.11) follows.

Next, we show that

$$E\big[\bar{\hat P}_n - \sum_{i=1}^{k} \bar{\hat P}_{in} + (k-1)\bar{\hat P}_{0n}\big] = o(1). \tag{A.2.15}$$

For any set $\tilde J \subset J$, denote

$$E_{\tilde J,n}(z) := \sqrt{n}\sum_{\tilde J} w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)E\big[\psi_n(X_{1i_1},...,X_{ki_k},z)\big].$$

We can write then

$$E[\hat P_n(z)] = E_{J,n}(z) = E_{J_1,n}(z) + E_{J\backslash J_1,n}(z). \tag{A.2.16}$$

Note that if $(i_1,...,i_k) \in J\backslash J_1$, by Fubini's theorem and independence of $X_{1i_1},...,X_{ki_k}$,

$$E\big[\psi_n(X_{1i_1},...,X_{ki_k},z)\big] = E\big[\tilde\psi_n(X_{1i_1}, z, z_{i_2},...,z_{i_k})\big]. \tag{A.2.17}$$

Hence, the second term on the right-hand side of (A.2.16)

$$E_{J\backslash J_1,n}(z) = \sqrt{n}\sum_{J\backslash J_1} w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)E\big[\tilde\psi_n(X_{1i_1}, z, z_{i_2},...,z_{i_k})\big].$$

Further, by definition,

$$\hat P_{0n}(z) = \sqrt{n}\sum_{J} w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)E\big[\tilde\psi_n(X_{1i_1}, z, z_{i_2},...,z_{i_k})\big].$$

Therefore,

$$
\begin{aligned}
&\big|E[\hat{P}_n(z)] - \hat{P}_{0n}(z)\big| \\
\leq& \big|E_{J_1,n}(z)\big| + \big|E_{J\setminus J_1,n}(z) - \hat{P}_{0n}(z)\big| \\
\leq& \sqrt{n}\sum_{J_1} w_{ni_1}(z,h)\cdots w_{ni_k}(z,h) \\
&\quad \times \Big(\big|E\big[\psi_n(X_{1i_1},...,X_{ki_k},z)\big]\big| + \big|E\big[\tilde{\psi}_n(X_{1i_1},z,z_{i_2},...,z_{i_k})\big]\big|\Big). \qquad \text{(A.2.18)}
\end{aligned}
$$

Note that if $(i_1,...,i_k) \in J_1$ with

$$w_{ni_1}(z,h)\cdots w_{ni_k}(z,h) \neq 0,$$

it must hold that

$$\|z_{i_1} - z\| \leq h, ..., \|z_{i_k} - z\| \leq h.$$

and at least two numbers out of the k numbers $i_1,...,i_k$ are equal. Hence, there are at most

$$\binom{k}{2} \cdot (2\lceil \bar{n}_0 h\rceil + 1)^{d(k-1)} \cdot m^{k-1}$$

elements in J_1 with

$$w_{ni_1}(z,h)\cdots w_{ni_k}(z,h) \neq 0.$$

Further, by Jensen's inequality, for all $i_1,...,i_k \in \{1,...,n\}$,

$$E\big[\psi_n(X_{1i_1},...,X_{ki_k},z)\big] \leq \Big(E\big[\psi_n^2(X_{1i_1},...,X_{ki_k},z)\big]\Big)^{1/2} \leq C^{1/2},$$

by (A.2.17), we get

$$\big|E\big[\tilde{\psi}_n(X_{1i_1},z,z_{i_2},...,z_{i_k})\big]\big| \leq C^{1/2}.$$

Thus,

$$
\begin{aligned}
\big|E[\hat{P}_n(z)] - \hat{P}_{0n}(z)\big| \leq& 2C^{1/2}\sqrt{n}\sum_{J_1} w_{ni_1}(z,h)\cdots w_{ni_k}(z,h) \\
\leq& k(k-1)C^{1/2}\sqrt{n}\cdot(2\lceil \bar{n}_0 h\rceil + 1)^{d(k-1)} \cdot m^{k-1}\Big(\frac{\tilde{C}}{n_0 h^d}\Big)^k
\end{aligned}
$$

Consequently, by Assumption (ii) $n_0 h^{2d} \to \infty$, we get as $n_0 \to \infty$,

$$\big|E[\bar{\hat{P}}_n] - \bar{\hat{P}}_{0n}\big| \leq \frac{1}{n_0}\sum_{i=1}^{n_0}\big|E[\hat{P}_n(z_i)] - \hat{P}_{0n}(z_i)\big|$$

$$\leq k(k-1)C^{1/2}\sqrt{n}\cdot(2\lceil\bar{n}_0 h\rceil+1)^{d(k-1)}\cdot m^{k-1}\big(\frac{\tilde{C}}{n_0 h^d}\big)^k$$
$$\leq k(k-1)C^{1/2}\tilde{C}^k\cdot m^{k-1/2}\big(\frac{2\lceil\bar{n}_0 h\rceil+1}{\bar{n}_0 h}\big)^{d(k-1)}n_0^{-1/2}h^{-1}=o(1).$$

By definition, we have

$$E[\hat{P}_{1n}(z)]=\sqrt{n}\sum_J w_{ni_1}(z,h)\cdots w_{ni_k}(z,h)E\big[\tilde{\psi}_n(X_{1i_1},z,z_{i_2},...,z_{i_k})\big]=\hat{P}_{0n}(z).$$

Thus,

$$E[\bar{\hat{P}}_{1n}]=\bar{\hat{P}}_{0n}.$$

With the same arguments, we can show that for any $i\in\{1,...,k\}$,

$$E[\bar{\hat{P}}_{in}]=\bar{\hat{P}}_{0n}.$$

Therefore, Equality (A.2.15) follows. The assertion follows then from Chebyshev's inequality by (A.2.11) and (A.2.15). □

Corollary A.2.10. *Under the assumptions of Lemma A.2.9, we have*

$$n^{-1/2}\bar{\hat{P}}_n=n^{-1/2}\bar{\hat{P}}_{0n}+o_p(1).$$

Proof. Analogously to Corollary A.1.2, the assertion can be shown by Lemma A.2.9. □

Lemma A.2.11. *Let $(\psi_n)_{n\in m\cdot\mathbb{N}}:\mathbb{R}\times[0,1]^d\to\mathbb{R}$ be a sequence of functions. If the functions $(\psi_n)_{n\in m\cdot\mathbb{N}}$ are dominated by the same H square integrable function independent of z, then*

$$\int\psi_n(x,z)d\hat{Q}_n(x,z)=\int\psi_n(x,z)dE[\hat{Q}_n(x,z)]+o_p(1)=O_p(1).$$

Proof. Note that by (A.2.10) and the assumption of this lemma, there exists a constant $C>0$ such that,

$$\begin{aligned}&Var\Big[\int\psi_n(x,z)d\hat{Q}_n(x,z)\Big]\\ \leq&\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n w_{nj}(z_i,h)w_{nj}(z_k,h)\Big|\int\psi_n(x,z_i)\psi_n(x,z_k)dH(x|z_j)\\ &\qquad-\int\psi_n(x,z_i)dH(x|z_j)\int\psi_n(x,z_k)dH(x|z_j)\Big|\end{aligned}$$

$$\leq C\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n w_{nj}(z_i,h)w_{nj}(z_k,h) = C\cdot\frac{1}{n^2}\sum_{j=1}^n 1 + o(1),$$

where the last step follows from Lemma A.2.4 with $\psi_n(z) = 1$ for all $n \in m\cdot\mathbb{N}$ and $z \in [0,1]^d$. Thus,

$$Var\Big[\int \psi_n(x,z)d\hat{Q}_n(x,z)\Big] = o(1).$$

Hence, it follows from Chebyshev's inequality that

$$\int \psi_n(x,z)d\hat{Q}_n(x,z) = \int \psi_n(x,z)dE[\hat{Q}_n(x,z)] + o_p(1).$$

Further, there exists a constant $C > 0$ such that,

$$\begin{aligned}\Big|\int \psi_n(x,z)dE\big[\hat{Q}_n(x,z)\big]\Big| \leq &\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n w_{nj}(z_i,h)\Big|\int \psi_n(x,z_i)dH(x|z_j)\Big| \\ \leq &C\cdot\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n w_{nj}(z_i,h) = C.\end{aligned}$$

Hence,

$$\int \psi_n(x,z)dE[\hat{Q}_n(x,z)] + o_p(1) = O_p(1).$$

□

Lemma A.2.12. *Let $\psi : \mathbb{R}\times[0,1]^d \to \mathbb{R}$ be an H integrable function. If $\partial\psi/\partial z$ is dominated by H intergrable function independent of z, then*

$$\int \psi(x,z)dE[\hat{Q}_n(x,z)] = \int \psi(x,z)dQ(x,z) + o(1)$$

Proof. By a Taylor expansion, for each $i,j \in \{1,...,n\}$, there exists z_{ij} lying between z_i and z_j, such that

$$\begin{aligned}&\int \psi(x,z)dE[\hat{Q}_n(x,z)] \\ =&\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n w_{nj}(z_i)\int \psi(x,z_i)dH(x|z_j) \\ =&\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n w_{nj}(z_i)\int \Big(\psi(x,z_j) + \Big(\frac{\psi(x,z_{ij})}{\partial z}\Big)^T\cdot(z_i - z_j)\Big)dH(x|z_j) \\ =&\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n w_{nj}(z_i)\int \psi(x,z_j)dH(x|z_j)\end{aligned}$$

$$+\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{nj}(z_i)\int\Big(\frac{\psi(x,z_{ij})}{\partial z}\Big)^T dH(x|z_j)\cdot(z_i-z_j). \tag{A.2.19}$$

By Lemma A.2.3, the first term on the right-hand side of (A.2.19) can be written as

$$\frac{1}{n}\sum_{j=1}^{n}\int\psi(x,z_j)dH(x|z_j)+o(1)=\int\psi(x,z)dQ(x,z)+o(1).$$

By the assumption of this lemma, there exists a $C>0$ such that the second term on the right-hand side of (A.2.19) is bounded by

$$C\cdot\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{nj}(z_i)\|z_i-z_j\|\le Ch\cdot\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{nj}(z_i)=Ch$$

where the inequality follows from Lemma A.2.5. Hence, the assertion follows. □

Lemma A.2.13. *Assume that d=1 and ψ is a Lipschitz continuous function on $[0,1]$, then*

$$\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\psi(z_i)-\sqrt{n}\cdot\int_0^1\psi(z)dz=o(1)$$

Proof. By the mean value theorem for integration, there exist z_i' lying between z_{i-1} and z_i for each $i\in\{1,...,n_0\}$, such that

$$\int_0^1\psi(z)dz=\sum_{i=1}^{n_0}\int_{z_{i-1}}^{z_i}\psi(z)dz=\frac{1}{n_0}\sum_{i=1}^{n_0}\psi(z_i'),$$

where $z_0=0$. By the Lipschitz continuity of the function ψ, there exists a constant $C>0$, such that

$$\begin{aligned}&\Big|\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\psi(z_i)-\sqrt{n}\cdot\int_0^1\psi(z)dz\Big|\\
=&\Big|\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\psi(z_i)-\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\psi(z_i')\Big|\\
\le&\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\big|\psi(z_i)-\psi(z_i')\big|\\
\le&C\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}|z_i-z_i'|\le C\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n_0}\frac{1}{n_0}=C\sqrt{n}\cdot\frac{1}{n_0}=o(1).\end{aligned}$$

□

A.3 Appendix of Section 3.2 ($m \to \infty, n_0$ Fixed)

Lemma A.3.1. *Let* $\psi : \mathbb{R}\times[0,1]^d \to \mathbb{R}$ *be a bounded function with* $\psi(x,z) = 0$*, if* $x > \tau$*, then*

$$\begin{aligned}
&\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\int \psi(x,z_k)\, dH_n^{KM}(x|z_k)\\
=&\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\int \psi(x,z_k)\gamma(x|z_k)\, dB_n^1(x|z_k)\\
&+\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\int\int \frac{I(x<u)\psi(u,z_k)}{1-B(x|z_k)}\, dH(u|z_k)dB_n^0(x|z_k)\\
&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\int\int \psi(u,z_k)C(x\wedge u|z_k)\, dH(u|z_k)dB_n(x|z_k) + o_p(1).
\end{aligned}$$

Proof. It follows directly from Theorem 1.1 of Stute (1995) that for each $k \in \{1,...,n_0\}$,

$$\begin{aligned}
&\sqrt{n}\cdot\int \psi(x,z_k)\, dH_n^{KM}(x|z_k)\\
=&\sqrt{n}\cdot\int \psi(x,z_k)\gamma(x|z_k)\, dB_n^1(x|z_k)\\
&+\sqrt{n}\cdot\int\int \frac{I(x<u)\psi(u,z_k)}{1-B(x|z_k)}\, dH(u|z_k)dB_n^0(x|z_k)\\
&-\sqrt{n}\cdot\int\int \psi(u,z_k)C(x\wedge u|z_k)\, dH(u|z_k)dB_n(x|z_k) + o_p(1).
\end{aligned}$$

Hence, the assertion follows since n_0 is fixed. □

Lemma A.3.2. *Let* $\psi : \mathbb{R}^2\times[0,1]^d \to \mathbb{R}$ *be a bounded function with* $\psi(x,y,z) = 0$*, if* $\max(x,y) > \tau$*. Then we have*

$$\begin{aligned}
\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\int\int \psi(x,y,z_k)\,\big[&dH_n^{KM}(x|z_k)dH_n^{KM}(y|z_k) - dH_n^{KM}(x|z_k)dH(y|z_k)\\
&- dH(x|z_k)dH_n^{KM}(y|z_k) + dH(x|z_k)dH(y|z_k)\big] = o_p(1).
\end{aligned}$$

Proof. Note that by definition $H_n^{KM}(\cdot|z)$ is a step function with possible jumps at $Y_1,...,Y_n$. The mass attached to ith order statistic $Y_{(i)}$ is equal to

$$1-\prod_{j=1}^{i}(1-\frac{\delta_{(j)}(z)\Delta_{(j)}}{m-\sum_{k=1}^{j-1}\delta_{(k)}(z)}) - \Big(1-\prod_{j=1}^{i-1}(1-\frac{\delta_{(j)}(z)\Delta_{(j)}}{m-\sum_{k=1}^{j-1}\delta_{(k)}(z)})\Big)$$

$$=\prod_{j=1}^{i-1}\big(1-\frac{\delta_{(j)}(z)\Delta_{(j)}}{m-\sum_{k=1}^{j-1}\delta_{(k)}(z)}\big)-\prod_{j=1}^{i}\big(1-\frac{\delta_{(j)}(z)\Delta_{(j)}}{m-\sum_{k=1}^{j-1}\delta_{(k)}(z)}\big)$$

$$=\Big(1-\big(1-\frac{\delta_{(i)}(z)\Delta_{(i)}}{m-\sum_{k=1}^{i-1}\delta_{(k)}(z)}\big)\Big)\cdot\prod_{j=1}^{i-1}\big(1-\frac{\delta_{(j)}(z)\Delta_{(j)}}{m-\sum_{k=1}^{j-1}\delta_{(k)}(z)}\big)$$

$$=\frac{\delta_{(i)}(z)\Delta_{(i)}}{m-\sum_{k=1}^{i-1}\delta_{(k)}(z)}\cdot\prod_{j=1}^{i-1}\big(1-\frac{\delta_{(j)}(z)\cdot\Delta_{(j)}}{m-\sum_{k=1}^{j-1}\delta_{(k)}(z)}\big)$$

$$=\frac{\delta_{(i)}(z)\Delta_{(i)}}{m-\sum_{k=1}^{i-1}\delta_{(k)}(z)}\cdot\prod_{j=1}^{i-1}\Big(\frac{m-\sum_{k=1}^{j-1}\delta_{(k)}(z)-\delta_{(j)}(z)\Delta_{(j)}}{m-\sum_{k=1}^{j-1}\delta_{(k)}(z)}\Big)$$

$$=\frac{1}{m}\cdot\delta_{(i)}(z)\Delta_{(i)}\prod_{j=1}^{i-1}\Big(\frac{m-\sum_{k=1}^{j-1}\delta_{(k)}(z)-\delta_{(j)}(z)\Delta_{(j)}}{m-\sum_{k=1}^{j}\delta_{(k)}(z)}\Big)$$

$$=\frac{1}{m}\cdot\delta_{(i)}(z)\Delta_{(i)}\prod_{j=1}^{i-1}\Big(1+\frac{\delta_{(j)}(z)(1-\Delta_{(j)})}{m\big(1-B_n(Y_{(j)}|z)\big)}\Big).$$

Thus, for the function $\psi:\mathbb{R}^2\times[0,1]^d\to\mathbb{R}$, we get

$$\int\psi(x,y,z)dH_n^{KM}(x|z)$$

$$=\frac{1}{m}\sum_{i=1}^{n}\psi(Y_{(i)},y,z)\Delta_{(i)}\delta_{(i)}(z)\prod_{j=1}^{i-1}\Big(1+\frac{\delta_{n(j)}(z)(1-\Delta_{(j)})}{m\big(1-B_n(Y_{(j)}|z)\big)}\Big)$$

$$=\frac{1}{m}\sum_{i=1}^{n}\psi(Y_i,y,z)\Delta_i\delta_i(z)\prod_{j=1}^{n}I(Y_j<Y_i)\Big(1+\frac{\delta_j(z)(1-\Delta_j)}{m\big(1-B_n(Y_j|z)\big)}\Big)$$

$$=\frac{1}{m}\sum_{i=1}^{n}\psi(Y_i,y,z)\Delta_i\delta_i(z)\exp\sum_{j=1}^{n}I(Y_j<Y_i)\ln\Big(1+\frac{\delta_j(z)(1-\Delta_j)}{m\big(1-B_n(Y_j|z)\big)}\Big)$$

$$=\frac{1}{m}\sum_{i=1}^{n}\psi(Y_i,y,z)\Delta_i\delta_i(z)\exp\sum_{j=1}^{n}I(Y_j<Y_i)(1-\Delta_j)\delta_j(z)\ln\Big(1+\frac{1}{m\big(1-B_n(Y_j|z)\big)}\Big)$$

$$=\frac{1}{m}\sum_{i=1}^{n}\psi(Y_i,y,z)\Delta_i\delta_i(z)\exp\left(m\int_{-\infty}^{Y_i-}\ln\Big(1+\frac{1}{m\big(1-B_n(x|z)\big)}\Big)dB_n^0(x|z)\right).$$

Analogously to Stute (1995), for $i\in\{1,...,n\}$, define the functions $A_{in},B_{in},C_{in}:[0,1]^d\to\mathbb{R}$ with

$$A_{in}(z):=m\int_{-\infty}^{Y_i-}\ln\Big(1+\frac{1}{m\big(1-B_n(x|z)\big)}\Big)dB_n^0(x|z)-\int_{-\infty}^{Y_i-}\frac{1}{1-B(x|z)}dB^0(x|z),$$

$$B_{in}(z):=m\int_{-\infty}^{Y_i-}\ln\Big(1+\frac{1}{m\big(1-B_n(x|z)\big)}\Big)dB_n^0(x)-\int_{-\infty}^{Y_i-}\frac{1}{1-B_n(x|z)}dB_n^0(x|z),$$

$$C_{in}(z) := \int_{-\infty}^{Y_i-} \frac{1}{1 - B_n(x|z)} dB_n^0(x|z) - \int_{-\infty}^{Y_i-} \frac{1}{1 - B(x|z)} dB^0(x|z).$$

Note that

$$A_{in}(z) = B_{in}(z) + C_{in}(z).$$

By a Taylor expansion, for $a, b \in \mathbb{R}$ there exists a ξ lying between a and b such that

$$\exp(a) = \exp(b) + \exp(b)(a - b) + \frac{1}{2}\exp(\xi)(a - b)^2$$

Thus, by setting

$$a = m \int_{-\infty}^{Y_i-} \ln\Big(1 + \frac{1}{m\big(1 - B_n(x|z)\big)}\Big) dB_n^0(x|z)$$

and

$$b = \int_{-\infty}^{Y_i-} \frac{1}{1 - B(x|z)} dB^0(x|z),$$

we get

$$\begin{aligned}
&\exp\left(m \int_{-\infty}^{Y_i-} \ln\Big(1 + \frac{1}{m\big(1 - B_n(x|z)\big)}\Big) dB_n^0(x|z)\right) \\
=&\gamma(Y_i|z) + \gamma(Y_i|z)A_{in}(z) + \frac{1}{2}e^{\xi_i(z)} \cdot A_{in}^2(z) \\
=&\gamma(Y_i|z)\big(1 + B_{in}(z) + C_{in}(z)\big) + \frac{1}{2}e^{\xi_i(z)}\big(B_{in}(z) + C_{in}(z)\big)^2,
\end{aligned}$$

where $\xi_i(z)$ lies between and

$$m \int_{-\infty}^{Y_i-} \ln\Big(1 + \frac{1}{m\big(1 - B_n(x|z)\big)}\Big) dB_n^0(x|z)$$

and

$$\int_{-\infty}^{Y_i-} \frac{1}{1 - B(x|z)} dB^0(x|z).$$

Therefore, we get

$$\begin{aligned}
\int \psi(x, y, z) dH_n^{KM}(x|z) =& \frac{1}{m}\sum_{i=1}^{n} \psi(Y_i, y, z)\Delta_i\delta_i(z)\gamma(Y_i|z)\big(1 + B_{in}(z) + C_{in}(z)\big) \\
&+ \frac{1}{m}\sum_{i=1}^{n} \psi(Y_i, y, z)\Delta_i\delta_i(z)\frac{1}{2}e^{\xi_i(z)}\big(B_{in}(z) + C_{in}(z)\big)^2.
\end{aligned}$$

Consequently, for the function $\psi : \mathbb{R}^2 \times [0, 1]^d \to \mathbb{R}$, we get

$$\sqrt{n} \cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \int\int \psi(x, y, z_k)\, dH_n^{KM}(x|z_k) dH_n^{KM}(y|z_k)$$

$$
\begin{aligned}
=&\sqrt{n}\cdot\frac{1}{n_0m^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n_0}\Big(\psi(Y_i,Y_j,z_k)\gamma(Y_i|z_k)\gamma(Y_j|z_k)\Delta_i\Delta_j\delta_i(z_k)\delta_j(z_k)\\
&\qquad\times\big(1+B_{in}(z_k)+C_{in}(z_k)\big)\cdot\big(1+B_{jn}(z_k)+C_{jn}(z_k)\big)\Big)\\
&+\sqrt{n}\cdot\frac{1}{n_0m^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n_0}\frac{1}{2}\Big(\psi(Y_i,Y_j,z_k)\gamma(Y_i|z_k)\Delta_i\Delta_j\delta_i(z_k)\delta_j(z_k)e^{\xi_j(z_k)}\\
&\qquad\times\big(1+B_{in}(z_k)+C_{in}(z_k)\big)\cdot\big(B_{jn}(z_k)+C_{jn}(z_k)\big)^2\Big)\\
&+\sqrt{n}\cdot\frac{1}{n_0m^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n_0}\frac{1}{2}\Big(\psi(Y_i,Y_j,z_k)\gamma(Y_j|z_k)\Delta_i\Delta_j\delta_i(z_k)\delta_j(z_k)e^{\xi_i(z_k)}\\
&\qquad\times\big(B_{in}(z_k)+C_{in}(z_k)\big)^2\cdot\big(1+B_{jn}(z_k)+C_{jn}(z_k)\big)\Big)\\
&+\sqrt{n}\cdot\frac{1}{n_0m^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n_0}\frac{1}{4}\Big(\psi(Y_i,Y_j,z_k)\Delta_i\Delta_j\delta_i(z_k)\delta_j(z_k)e^{\xi_i(z_k)}e^{\xi_j(z_k)}\\
&\qquad\times\big(B_{in}(z_k)+C_{in}(z_k)\big)^2\cdot\big(B_{jn}(z_k)+C_{jn}(z_k)\big)^2\Big).\qquad\text{(A.3.1)}
\end{aligned}
$$

In the first term of (A.3.1), by Lemma A.1.1 with $k=2$, $X_{1j}=X_{2j}=\Delta_j\cdot Y_j$ for $j\in\{1,...,n\}$ and $\psi(x,y,z)=\psi(x,y,z)\gamma(x|z)\gamma(y|z)$, we can write

$$
\begin{aligned}
&\sqrt{n}\cdot\frac{1}{n_0m^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n_0}\psi(Y_i,Y_j,z_k)\gamma(Y_i|z_k)\gamma(Y_j|z_k)\Delta_i\Delta_j\delta_i(z_k)\delta_j(z_k)\\
=&\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\int\int\psi(x,y,z_k)\gamma(x|z_k)\gamma(y|z_k)dB_n^1(x|z_k)dB_n^1(y|z_k)\\
=&\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\int\int\psi(x,y,z_k)\gamma(x|z_k)\gamma(y|z_k)\big[dB_n^1(x|z_k)dB^1(y|z_k)\\
&\qquad+dB^1(x|z_k)dB_n^1(y|z_k)-dB^1(x|z_k)dB^1(y|z_k)\big]+o_p(1).\qquad\text{(A.3.2)}
\end{aligned}
$$

In the sequel, we show that

$$
\sqrt{n}\cdot\frac{1}{n_0m^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n_0}\psi(Y_i,Y_j,z_k)\gamma(Y_i|z_k)\gamma(Y_j|z_k)\Delta_i\Delta_j\delta_i(z_k)\delta_j(z_k)B_{in}(z_k)=o_p(1).\qquad\text{(A.3.3)}
$$

Since for $a\geq 0$,

$$
-\frac{a^2}{2}\leq\ln(1+a)-a\leq 0.
$$

By setting

$$
a=\frac{1}{m\big(1-B_n(Y_j|z)\big)},
$$

we get

$$-\frac{1}{2m^2\big(1-B_n(Y_j|z)\big)^2} \le \ln\Big(1+\frac{1}{m\big(1-B_n(Y_j|z)\big)}\Big) - \frac{1}{m\big(1-B_n(Y_j|z)\big)} \le 0.$$

Thus,

$$-\frac{1}{2m}\int_{-\infty}^{Y_i-} \frac{dB_n^0(x|z)}{\big(1-B_n(x|z)\big)^2} \le B_{in}(z) \le 0.$$

Hence, we obtain for any $k \in \{1, ..., n_0\}$ and $i \in \{1, ..., n\}$ with $Y_i \le \tau$,

$$|B_{in}(z_k)| \le \frac{1}{2m}\int_{-\infty}^{Y_i-} \frac{dB_n^0(x|z_k)}{\min_{k\in\{1,...,n_0\}}\big(1-B_n(\tau|z_k)\big)^2}. \tag{A.3.4}$$

Note that

$$\min_{k\in\{1,...,n_0\}}\big(1-B_n(\tau|z_k)\big) \ge \min_{k\in\{1,...,n_0\}}\big(1-B(\tau|z_k)\big) + \min_{k\in\{1,...,n_0\}}\big(B(\tau|z_k)-B_n(\tau|z_k)\big)$$

Since $\tau < \tau_B$, thus there exists a constant $\delta > 0$ such that

$$\min_{k\in\{1,...,n_0\}}\big(1-B(\tau|z_k)\big) > \delta.$$

For $k \in \{1, ..., n_0\}$, by law of large numbers, $\big|B(\tau|z_k) - B_n(\tau|z_k)\big| \xrightarrow{a.s.} 0$. Since n_0 is a constant number, we have

$$\min_{k\in\{1,...,n_0\}}\big(B(\tau|z_k)-B_n(\tau|z_k)\big) \xrightarrow{a.s.} 0.$$

Therefore, for n large enough, we have

$$\min_{k\in\{1,...,n_0\}}\big(1-B_n(\tau|z_k)\big) \ge \delta \quad \text{a.s.} \tag{A.3.5}$$

Consequently, by (A.3.4), there exists a constant $C > 0$ such that for any $k \in \{1, ..., n_0\}$ and $i \in \{1, ..., n\}$ with $Y_i \le \tau$,

$$|B_{in}(z_k)| \le C \cdot m^{-1}, \quad \text{in probability}$$

Hence, by the assumption $\psi(x, y, z) = 0$, if $\max(x, y) > \tau$, we get

$$\Big|\sqrt{n}\cdot\frac{1}{n_0 m^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n_0}\psi(Y_i,Y_j,z_k)\gamma(Y_i|z_k)\gamma(Y_j|z_k)\Delta_i\Delta_j\delta_i(z_k)\delta_j(z_k)B_{in}(z_k)\Big|$$

$$\le C\sqrt{n}\cdot\frac{1}{n_0 m^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n_0}\big|\psi(Y_i,Y_j,z_k)\big|\gamma(Y_i|z_k)\gamma(Y_j|z_k)\Delta_i\Delta_j\delta_i(z_k)\delta_j(z_k)$$

$$=Cn^{-1/2}\sum_{k=1}^{n_0}\int\int\big|\psi(x,y,z_k)\big|\gamma(x|z_k)\gamma(y|z_k)dB_n^1(x|z_k)dB_n^1(y|z_k)=o_p(1)$$

where the last term follows from Corolary A.1.2 with $k=2$, $X_{1j}=X_{2j}=\Delta_j{\cdot}Y_j$ for $j\in\{1,...,n\}$ and $|\psi(x,y,z)|=\psi(x,y,z)\gamma(x|z)\gamma(y|z)$. Therefore, Equation (A.3.3) holds.

Next we show that

$$\begin{aligned}&\sqrt{n}\cdot\frac{1}{n_0m^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n_0}\psi(Y_i,Y_j,z_k)\gamma(Y_i|z_k)\gamma(Y_j|z_k)\Delta_i\Delta_j\delta_i(z_k)\delta_j(z_k)C_{in}(z_k)\\
=&\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\int\int\psi(x,y,z_k)\,dH_n^{KM}(x|z_k)dH(y|z_k)\\
&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\int\int\psi(x,y,z_k)\cdot\gamma(x|z_k)\,dB_n^1(x|z_k)dH(y|z_k)+o_p(1).\quad\text{(A.3.6)}\end{aligned}$$

Note that we can write

$$\begin{aligned}C_{in}(z)=&-\int_{-\infty}^{Y_i-}\int\frac{I(x<u)}{\big(1-B(x|z)\big)^2}dB_n(u|z)dB_n^0(x|z)\\
&+\int_{-\infty}^{Y_i-}\frac{2}{1-B(x|z)}dB_n^0(x|z)-\int_{-\infty}^{Y_i-}\frac{1}{1-B(x|z)}dB^0(x|z)\\
&+\int_{-\infty}^{Y_i-}\frac{\big(B_n(x|z)-B(x|z)\big)^2}{\big(1-B(x|z)\big)^2\big(1-B_n(x|z)\big)}dB_n^0(x|z).\end{aligned}$$

Denote the function $\tilde{\psi}:\mathbb{R}^3\times[0,1]^d\to\mathbb{R}$ with

$$\tilde{\psi}(x,y,u,z):=\psi(x,y,z)I(u<x)\gamma(x|z)\gamma(y|z)/\big(1-B(u|z)\big),$$

we get then

$$\begin{aligned}&\sqrt{n}\cdot\frac{1}{n_0m^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n_0}\psi(Y_i,Y_j,z_k)\gamma(Y_i|z_k)\gamma(Y_j|z_k)\Delta_i\Delta_j\delta_i(z_k)\delta_j(z_k)C_{in}(z_k)\\
=&\sqrt{n}\cdot\frac{1}{n_0m}\sum_{i=1}^{n}\sum_{k=1}^{n_0}\int\psi(Y_i,y,z_k)\gamma(y|z_k)dB_n^1(y|z_k)\cdot\gamma(Y_i|z_k)\Delta_i\delta_i(z_k)C_{in}(z_k)\\
=&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\iiiint\frac{\tilde{\psi}(x,y,u,z_k)I(u<t)}{1-B(u|z_k)}dB_n(t|z_k)dB_n^0(u|z_k)dB_n^1(x|z_k)dB_n^1(y|z_k)\\
&+2\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\iiint\tilde{\psi}(x,y,u,z_k)dB_n^0(u|z_k)dB_n^1(x|z_k)dB_n^1(y|z_k)\\
&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\iiint\tilde{\psi}(x,y,u,z_k)dB^0(u|z_k)dB_n^1(x|z_k)dB_n^1(y|z_k)\end{aligned}$$

$$+\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\iiint\frac{\tilde{\psi}(x,y,u,z_k)\big(B_n(u|z_k)-B(u|z_k)\big)^2}{\big(1-B(u|z_k)\big)\big(1-B_n(u|z_k)\big)}dB_n^0(u|z_k)dB_n^1(x|z_k)dB_n^1(y|z_k)$$

$$=:C_{n1}+C_{n2}+C_{n3}+C_{n4}.$$

By Lemma A.1.1 with $k=4$, $X_{1j}=Y_j$, $X_{2j}=(1-\Delta_j)\cdot Y_j$, $X_{3j}=H_{4j}=\Delta_j\cdot Y_j$ for $j\in\{1,...,n\}$ and

$$\psi(x,y,u,t,z)=\frac{\tilde{\psi}(x,y,u,z)I(u<t)}{1-B(u|z)},$$

we can write

$$\begin{aligned}C_{n1}=&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\iiiint\frac{\tilde{\psi}(x,y,u,z_k)I(u<t)}{1-B(u|z_k)}\Big[dB_n(t|z_k)dB^0(u|z_k)dB^1(x|z_k)dB^1(y|z_k)\\&+dB(t|z_k)dB_n^0(u|z_k)dB^1(x|z_k)dB^1(y|z_k)+dB(t|z_k)dB^0(u|z_k)dB_n^1(x|z_k)dB^1(y|z_k)\\&+dB(t|z_k)dB^0(u|z_k)dB^1(x|z_k)dB_n^1(y|z_k)\\&-3dB(t|z_k)dB^0(u|z_k)dB^1(x|z_k)dB^1(y|z_k)\Big]+o_p(1)\\=&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\iiiint\frac{\tilde{\psi}(x,y,u,z_k)I(u<t)}{1-B(u|z_k)}dB_n(t|z_k)dB^0(u|z_k)dB^1(x|z_k)dB^1(y|z_k)\\&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\iiint\tilde{\psi}(x,y,u,z_k)\Big[dB_n^0(u|z_k)dB^1(x|z_k)dB^1(y|z_k)\\&+dB^0(u|z_k)dB_n^1(x|z_k)dB^1(y|z_k)+dB^0(u|z_k)dB^1(x|z_k)dB_n^1(y|z_k)\\&-3dB^0(u|z_k)dB^1(x|z_k)dB^1(y|z_k)\Big]+o_p(1).\end{aligned}$$

With similar arguments, it can be shown that

$$\begin{aligned}C_{n2}=&2\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\iiint\tilde{\psi}(x,y,u,z_k)\Big[dB_n^0(u|z_k)dB^1(x|z_k)dB^1(y|z_k)\\&+dB^0(u|z_k)dB_n^1(x|z_k)dB^1(y|z_k)+dB^0(u|z_k)dB^1(x|z_k)dB^1(y|z_k)\\&-2dB^0(u|z_k)dB^1(x|z_k)dB^1(y|z_k)\Big]+o_p(1).\end{aligned}$$

and

$$\begin{aligned}C_{n3}=&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\iiint\tilde{\psi}(x,y,u,z_k)\Big[dB^0(u|z_k)dB_n^1(x|z_k)dB^1(y|z_k)\\&+B^0(u|z_k)dB^1(x|z_k)dB_n^1(y|z_k)-B^0(u|z_k)dB^1(x|z_k)dB^1(y|z_k)\Big]+o_p(1).\end{aligned}$$

By (A.3.5), there exists a $C > 0$ such that C_{n4} is bounded by

$$\sqrt{n} \cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \iiint \frac{|\tilde{\psi}(x, y, u, z_k)| \big(B_n(u|z_k) - B(u|z_k)\big)^2}{\big(1 - B(u|z_k)\big)\big(1 - B_n(\tau|z_k)\big)} dB_n^0(u|z_k) dB_n^1(x|z_k) dB_n^1(y|z_k)$$

$$\leq C\sqrt{n} \cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \iiint \frac{|\tilde{\psi}(x, y, u, z_k)| \big(B_n(u|z_k) - B(u|z_k)\big)^2}{1 - B(u|z_k)} dB_n^0(u|z_k) dB_n^1(x|z_k) dB_n^1(y|z_k),$$

which is equal to $o_p(1)$ by using the same arguments for T_{1n} in Theorem 2.2.4. Thus, $C_{n4} = o_p(1)$. Therefore, we get

$$\sqrt{n} \cdot \frac{1}{n_0 m^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n_0} \psi(Y_i, Y_j, z_k) \gamma(Y_i|z_k) \gamma(Y_j|z_k) \Delta_i \Delta_j \delta_i(z_k) \delta_j(z_k) C_{in}(z_k)$$

$$= -\sqrt{n} \cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \iiiint \frac{\tilde{\psi}(x, y, u, z_k) I(u < t)}{1 - B(u|z_k)} dB_n(t|z_k) dB^0(u|z_k) dB^1(x|z_k) dB^1(y|z_k)$$

$$+ \sqrt{n} \cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \iiint \tilde{\psi}(x, y, u, z_k) dB_n^0(u|z_k) dB^1(x|z_k) dB^1(y|z_k) + o_p(1)$$

$$= -\sqrt{n} \cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \iint \Big(\int \psi(x, y, z_k) dH(y|z_k) \Big) \cdot C(u \wedge x|z_k)\, dH(x|z_k) dB_n(u|z_k)$$

$$+ \sqrt{n} \cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \iint \frac{I(u < x) \int \psi(x, y, z_k) dH(y|z_k)}{1 - B(u|z_k)}\, dH(x|z_k) dB_n^0(u|z_k) + o_p(1)$$

$$= \sqrt{n} \cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \iint \psi(x, y, z_k)\, dH(y|z_k) dH_n^{KM}(x|z_k)$$

$$- \sqrt{n} \cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \iint \psi(x, y, z_k)\, dH(y|z_k) \cdot \gamma(x|z_k) dB_n^1(x|z_k) + o_p(1),$$

where the last step follows from Lemma A.3.1.

By symmetry arguments, we get further

$$\sqrt{n} \cdot \frac{1}{n_0 m^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n_0} \psi(Y_i, Y_j, z_k) \gamma(Y_i|z_k) \gamma(Y_j|z_k) \Delta_i \Delta_j \delta_i(z_k) \delta_j(z_k) C_{jn}(z_k)$$

$$= \sqrt{n} \cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \iint \psi(x, y, z_k)\, dH(x|z_k) dH_n^{KM}(y|z_k)$$

$$- \sqrt{n} \cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \iint \psi(x, y, z_k)\, dH(x|z_k) \cdot \gamma(y|z_k) dB_n^1(y|z_k) + o_p(1). \quad \text{(A.3.7)}$$

Analogously to (A.3.3) and (A.3.6), it can be shown that

$$\sqrt{n} \cdot \frac{1}{n_0 m^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n_0} \Big(\psi(Y_i, Y_j, z_k) \gamma(Y_i|z_k) \gamma(Y_j|z_k) \Delta_i \Delta_j \delta_i(z_k) \delta_j(z_k)$$

$$\times \big(B_{jn}(z_k) + B_{in}(z_k)B_{jn}(z_k) + B_{in}(z_k)C_{jn}(z_k) \\ + C_{in}(z_k)B_{jn}(z_k) + C_{in}(z_k)C_{jn}(z_k)\big)\Big) = o_p(1). \tag{A.3.8}$$

Therefore, it follows from (A.3.2), (A.3.3), (A.3.6), (A.3.7) and (A.3.8) that the first term on the right-hand side of (A.3.1) equals

$$\begin{aligned}
&\sqrt{n} \cdot \frac{1}{n_0 m^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n_0} \Big(\psi(Y_i, Y_j, z_k)\gamma(Y_i|z_k)\gamma(Y_j|z_k)\Delta_i\Delta_j\delta_i(z_k)\delta_j(z_k) \\
&\qquad \times \big(1 + C_{in}(z_k) + C_{jn}(z_k)\big)\Big) + o_p(1) \\
=&\sqrt{n} \cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \int\int \psi(x, y, z_k) \big[dH_n^{KM}(x|z_k)dH(y|z_k) \\
&\qquad + dH(x|z_k)dH_n^{KM}(y|z_k) - dH(x|z_k)dH(y|z_k)\big] + o_p(1).
\end{aligned}$$

By similar arguments, we can show all the last three terms on the right-hand side of (A.3.1) converge to zero in probability. Hence, the assertion holds. □

Lemma A.3.3. *Let $\psi : \mathbb{R}^3 \times [0,1]^d \to \mathbb{R}$ be a bounded function with $\psi(x_1, x_2, x_3, z) = 0$, if $\max(x_1, x_2, x_3) > \tau$. Then we have*

$$\begin{aligned}
\sqrt{n}\cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \iiint \psi(x_1, x_2, x_3, z_k) \Big[&dH_n^{KM}(x_1|z_k)dH_n^{KM}(x_2|z_k)dH_n^{KM}(x_3|z_k) \\
&-dH_n^{KM}(x_1|z_k)dH(x_2|z_k)dH(x_3|z_k) - dH(x_1|z_k)dH_n^{KM}(x_2|z_k)dH(x_3|z_k) \\
&-dH(x_1|z_k)dH(x_2|z_k)dH_n^{KM}(x_3|z_k) + 2dH(x_1|z_k)dH(x_2|z_k)dH(x_3|z_k)\Big] = o_p(1).
\end{aligned}$$

Proof. It can be shown similarly as Lemma A.3.2. □

A.4 Appendix of Section 3.3 ($n_0 \to \infty, m$ Fixed)

Lemma A.4.1. *There exists a constant $\delta > 0$ such that for all $z \in [0,1]^d$ and eventually all $n \in m \cdot \mathbb{N}$,*

$$1 - E[\hat{B}_n(\tau|z)] > \delta \quad \textit{and} \quad 1 - \hat{B}_n(\tau|z) > \delta \quad \textit{in probability}$$

Proof. By Assumption (i) and the independence of X_i and C_i, the function B has bounded derivative and Hessian matrix with respect to z. Hence, analogously to Lemma A.2.6, we can show there exists a constant $C > 0$, such that for any $z \in [0,1]^d$,

$$\big|E[\hat{B}_n(\tau|z)] - B(\tau|z)\big| \le Ch.$$

By definition of τ, there exists a constant $\delta > 0$, such that for any $z \in [0,1]^d$

$$1 - B(\tau|z) > \delta.$$

Hence, for any $z \in [0,1]^d$ and n large enough, we have

$$1 - E[\hat{B}_n(\tau|z)] = 1 - B(\tau|z) + B(\tau|z) - E[\hat{B}_n(\tau|z)] > \delta. \tag{A.4.1}$$

For the second part of assertion we show first

$$\max_{z\in[0,1]^d} \left|\hat{B}_n(\tau|z) - E[\hat{B}_n(\tau|z)]\right| = O_p(n_0^{-1/2} h^{-d/2} (\log n_0)^{1/2}).$$

By the compactness of $[0,1]^d$, there exist l_{n_0} subsets $S_1, ..., S_{l_{n_0}} \subset [0,1]^d$, such that

$$\max_i \max_{z,\tilde{z}\in S_i} \|z - \tilde{z}\| \le n_0^{-1/2} h^{1-d/2} (\log n_0)^{1/2} \quad \text{and} \quad [0,1]^d \subset \bigcup_{i=1}^{l_{n_0}} S_i.$$

For $i \in \{1, ..., l_{n_0}\}$, let $\tilde{z}_i$ be a vector in S_i, then we have

$$\begin{aligned}
&\max_{z\in[0,1]^d} \left|\hat{B}_n(\tau|z) - E[\hat{B}_n(\tau|z)]\right| \\
=&\max_i \max_{z\in S_i} \left|\hat{B}_n(\tau|z) - E[\hat{B}_n(\tau|z)]\right| \\
\le&\max_i \max_{z\in S_i} \left|\hat{B}_n(\tau|z) - \hat{B}_n(\tau|\tilde{z}_i)\right| + \max_i \left|\hat{B}_n(\tau|\tilde{z}_i) - E[\hat{B}_n(\tau|\tilde{z}_i)]\right| \\
&+ \max_i \max_{z\in S_i} \left|E[\hat{B}_n(\tau|\tilde{z}_i)] - E[\hat{B}_n(\tau|z)]\right|. \qquad \text{(A.4.2)}
\end{aligned}$$

Note that if $z \in S_i$,

$$\begin{aligned}
&\left|\hat{B}_n(\tau|z) - \hat{B}_n(\tau|\tilde{z}_i)\right| \\
=&\left|\sum_{j=1}^n w_{nj}(z,h) I(Y_j \le \tau) - \sum_{j=1}^n w_{nj}(\tilde{z}_i,h) I(Y_j \le \tau)\right| \\
=&\left|\frac{\sum_{j=1}^n K(\frac{z_j - z}{h}) I(Y_j \le \tau)}{\sum_{k=1}^n K(\frac{z_k - z}{h})} - \frac{\sum_{j=1}^n K(\frac{z_j - \tilde{z}_i}{h}) I(Y_j \le \tau)}{\sum_{k=1}^n K(\frac{z_k - \tilde{z}_i}{h})}\right| \\
\le&\left|\frac{\sum_{j=1}^n K(\frac{z_j - z}{h}) I(Y_j \le \tau)}{\sum_{k=1}^n K(\frac{z_k - z}{h})} - \frac{\sum_{j=1}^n K(\frac{z_j - z}{h}) I(Y_j \le \tau)}{\sum_{k=1}^n K(\frac{z_k - \tilde{z}_i}{h})}\right| \\
&+ \left|\frac{\sum_{j=1}^n K(\frac{z_j - z}{h}) I(Y_j \le \tau)}{\sum_{k=1}^n K(\frac{z_k - \tilde{z}_i}{h})} - \frac{\sum_{j=1}^n K(\frac{z_j - \tilde{z}_i}{h}) I(Y_j \le \tau)}{\sum_{k=1}^n K(\frac{z_k - \tilde{z}_i}{h})}\right|. \qquad \text{(A.4.3)}
\end{aligned}$$

In the following, we show that there exists a constant δ such that for all $z \in [0,1]^d$,

$$\frac{1}{nh^d}\sum_{k=1}^{n} K(\frac{z_k - z}{h}) > \delta. \tag{A.4.4}$$

By (A.2.1), if $z \in \{z_1, ...z_{n_0}\}$,

$$\frac{1}{nh^d}\sum_{k=1}^{n} K(\frac{z_k - z}{h}) \geq \frac{1}{n_0 h^d}\sum_{z'_k \in I'_0} K(z'_k).$$

where I'_0 is defined in Section A.2. Further, we have

$$\frac{1}{n_0 h^d}\sum_{z'_k \in I'_0} K(z'_k) \to \int_{[0,1]^d} K(x)dx.$$

Hence, for n large enough there exists a constant $\delta > 0$ such that for all $z \in \{z_1, ...z_{n_0}\}$,

$$\frac{1}{nh^d}\sum_{k=1}^{n} K(\frac{z_k - z}{h}) > \delta. \tag{A.4.5}$$

For any $z \in [0,1]^d$, by the definition of $\{z_1, ...z_{n_0}\}$, there exists a $z' \in \{z_1, ...z_{n_0}\}$ with $\|z - z'\| \leq 1/\bar{n}_0$ and

$$\begin{aligned}
&\Big|\frac{1}{nh^d}\sum_{k=1}^{n} K(\frac{z_k - z}{h}) - \frac{1}{nh^d}\sum_{k=1}^{n} K(\frac{z_k - z'}{h})\Big| \\
\leq &\frac{1}{n_0 h^d} \cdot \sum_{k=1}^{n_0} \Big|K(\frac{z_k - z}{h}) - K(\frac{z_k - z'}{h})\Big|
\end{aligned}$$

By Assumption (iii) $K(x) = 0$ if $\|x\| > 1$, hence, maximal $2 \cdot (2\lceil \bar{n}_0 h \rceil + 1)^d$ summands in the last term are nonzero. Thus, by the Lipschitz continuity of the function K, there exists a constant $C > 0$, such that

$$\begin{aligned}
&\Big|\frac{1}{nh^d}\sum_{k=1}^{n} K(\frac{z_k - z}{h}) - \frac{1}{nh^d}\sum_{k=1}^{n} K(\frac{z_k - z'}{h})\Big| \\
\leq &\frac{1}{n_0 h^d} \cdot (2\lceil \bar{n}_0 h \rceil + 1)^d \cdot C \cdot \frac{\|z - z'\|}{h} \\
\leq &\frac{1}{n_0 h^d} \cdot (2\lceil \bar{n}_0 h \rceil + 1)^d \cdot C \cdot \frac{1}{\bar{n}_0 h} = o(1).
\end{aligned} \tag{A.4.6}$$

Thus, it follows from (A.4.5) that for any $z \in [0,1]^d$, (A.4.4) holds.

The first term on the right-hand side of (A.4.3)

$$
\begin{aligned}
&\Big|\frac{\sum_{j=1}^n K(\frac{z_j-z}{h})I(Y_j\le\tau)}{\sum_{k=1}^n K(\frac{z_k-z}{h})}-\frac{\sum_{j=1}^n K(\frac{z_j-z}{h})I(Y_j\le\tau)}{\sum_{k=1}^n K(\frac{z_k-\tilde z_i}{h})}\Big|\\
=&\Big(\sum_{j=1}^n K(\frac{z_j-z}{h})I(Y_j\le\tau)\Big)\Big|\frac{\sum_{k=1}^n K(\frac{z_k-z}{h})-\sum_{k=1}^n K(\frac{z_k-\tilde z_i}{h})}{\sum_{k=1}^n K(\frac{z_k-z}{h})\sum_{k=1}^n K(\frac{z_k-\tilde z_i}{h})}\Big|\\
\le&\Big(\sum_{j=1}^n K(\frac{z_j-z}{h})\Big)\Big|\frac{\sum_{k=1}^n K(\frac{z_k-z}{h})-\sum_{k=1}^n K(\frac{z_k-\tilde z_i}{h})}{\sum_{k=1}^n K(\frac{z_k-z}{h})\sum_{k=1}^n K(\frac{z_k-\tilde z_i}{h})}\Big|\\
=&\Big|\frac{\sum_{k=1}^n K(\frac{z_k-z}{h})-\sum_{k=1}^n K(\frac{z_k-\tilde z_i}{h})}{\sum_{k=1}^n K(\frac{z_k-\tilde z_i}{h})}\Big|\\
=&\Big|\frac{\frac{1}{nh^d}\sum_{k=1}^n K(\frac{z_k-z}{h})-\frac{1}{nh^d}\sum_{k=1}^n K(\frac{z_k-\tilde z_i}{h})}{\frac{1}{nh^d}\sum_{k=1}^n K(\frac{z_k-\tilde z_i}{h})}\Big|=O(n_0^{-1/2}h^{-d/2}(\log n_0)^{1/2})
\end{aligned}
$$

where the last term can be shown analogously to (A.4.6) by (A.4.4). The same results can be shown for the second term on the right-hand side of (A.4.3). Therefore, we get

$$\max_i\max_{z\in S_i}\big|\hat B_n(\tau|z)-\hat B_n(\tau|\tilde z_i)\big|=O_p(n_0^{-1/2}h^{-d/2}(\log n_0)^{1/2})\ \text{ a.s.}$$

With the same arguments, it can be shown that

$$\max_i\max_{z\in S_i}\big|E[\hat B_n(\tau|\tilde z_i)]-E[\hat B_n(\tau|z)]\big|=O(n_0^{-1/2}h^{-d/2}(\log n_0)^{1/2})$$

as well.

In the following we show that

$$\max_i\big|\hat B_n(\tau|\tilde z_i)-E[\hat B_n(\tau|\tilde z_i)]=O_p(n_0^{-1/2}h^{-d/2}(\log n_0)^{1/2}). \tag{A.4.7}$$

Denote for each $j\in\{1,...,n\}$,

$$U_j:=w_{nj}(\tilde z_i,h)\big(I(Y_j\le\tau)-B(\tau|z_j)\big),$$

then we get $U_1,...,U_n$ are independent,

$$\hat B_n(\tau|\tilde z_i)-E[\hat B_n(\tau|\tilde z_i)]=\sum_{j=1}^n w_{nj}(\tilde z_i,h)\big(I(Y_j\le\tau)-B(\tau|z_j)\big)=\sum_{j=1}^n U_j$$

and $E[U_j]=0$ for $j\in\{1,...,n\}$. Note further by (A.2.12) there exists a constant $C>0$ such that for any $j\in\{1,...,n\}$,

$$|U_j|\le w_{nj}(\tilde z_i,h)\le\frac{C}{n_0h^d}$$

and

$$\sum_{j=1}^{n} Var[U_j] = Var[\hat{B}_n(\tau|\tilde{z}_i)] = \sum_{j=1}^{n} w_{nj}^2(\tilde{z}_i, h) \cdot \big(B(\tau|\tilde{z}_i) - B^2(\tau|\tilde{z}_i)\big)$$
$$\leq \sum_{j=1}^{n} w_{nj}^2(\tilde{z}_i, h) \leq \frac{C}{n_0 h^d} \sum_{j=1}^{n} w_{nj}(\tilde{z}_i, h) \leq \frac{C}{n_0 h^d}.$$

Thus, Corollary A.9 of Ferraty and Vieu (2006) implies that there exists a constant $C > 0$ such that for any $\epsilon > 0$,

$$P\big(\big|\hat{B}_n(\tau|\tilde{z}_i) - E[\hat{B}_n(\tau|\tilde{z}_i)]\big| > \epsilon\big) \leq 2 \cdot \exp(-C\epsilon^2 n_0 h^d).$$

Therefore,

$$P(\max_i \big|\hat{B}_n(\tau|\tilde{z}_i) - E[\hat{B}_n(\tau|\tilde{z}_i)]\big| > \epsilon)$$
$$\leq l_{n_0} \cdot \max_i P(\big|\hat{B}_n(\tau|\tilde{z}_i) - E[\hat{B}_n(\tau|\tilde{z}_i)]\big| > \epsilon) \leq 2 \cdot l_{n_0} \cdot \exp(-C\epsilon^2 n_0 h^d).$$

By setting

$$\epsilon = C' n_0^{-1/2} h^{-d/2} (\log n_0)^{1/2}$$

for a constant $C' > 0$ with $\sum_{n_0=1}^{\infty} l_{n_0} \cdot n_0^{-CC'^2} = o(1)$, we get

$$P\big(\max_i \big|\hat{B}_n(\tau|\tilde{z}_i) - E[\hat{B}_n(\tau|\tilde{z}_i)]\big| > C' n_0^{-1/2} h^{-d/2} (\log n_0)^{1/2}\big) \leq 2 \cdot l_{n_0} \cdot n_0^{-CC'^2}.$$

Hence, Equality (A.4.7) follows from an application of Borel-Cantelli Lemma. By (A.4.2), we get then

$$\max_{z \in [0,1]^d} \big|\hat{B}_n(\tau|z) - E[\hat{B}_n(\tau|z)]\big| = O_p(n_0^{-1/2} h^{-d/2} (\log n_0)^{1/2}).$$

Therefore, by (A.4.1) and Assumption (ii) that $n_0^{-1/2} h^{-d/2}(\log n_0) \to 0$, for any $z \in [0,1]^d$ and n large enough, there exists a $\delta > 0$, such that

$$1 - \hat{B}_n(\tau|z) = 1 - E[\hat{B}_n(\tau|z)] + E[\hat{B}_n(\tau|z)] - \hat{B}_n(\tau|z) > \delta.$$

in probability. □

Lemma A.4.2. *Let $(\psi_n)_{n \in m\cdot\mathbb{N}} : \mathbb{R} \times [0,1]^d$ be a sequence of uniformly bounded functions, with $\psi_n(x,z) = 0$ if $x > \tau$ and $z \in [0,1]^d$. Then we have*

$$\sqrt{n} \cdot \frac{1}{n_0} \sum_{j=1}^{n_0} \int \psi_n(x, z_j)\, d\hat{H}_n^{KM}(x|z_j)$$

$$
\begin{aligned}
=&\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\int\psi_n(x,z_j)\gamma_n(x|z_j)\,d\hat{B}_n^1(x|z_j)\\
&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iint\psi_n(u,z_j)C_n(x\wedge u|z_j)\gamma_n(u|z_j)dE[\hat{B}_n^1(u|z_j)]d\hat{B}_n(x|z_j)\\
&+\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iint\frac{I(x<u)\psi_n(u,z_j)\gamma_n(u|z_j)}{1-E[\hat{B}_n(x|z_j)]}\,dE[\hat{B}_n^1(u|z_j)]d\hat{B}_n^0(x|z_j)+o_p(1).
\end{aligned}
$$

Proof. By definition of Beran's estimate, for each z, it is a step function with possible jumps at $Y_1,...,Y_n$. The mass attached to ith order statistic $Y_{(i)}$ is equal to

$$
\begin{aligned}
&1-\prod_{j=1}^{i}\big(1-\frac{w_{n(j)}(z,h)\Delta_{(j)}}{1-\sum_{k=1}^{j-1}w_{n(k)}(z,h)}\big)-\Big(1-\prod_{j=1}^{i-1}\big(1-\frac{w_{n(j)}(z,h)\Delta_{(j)}}{1-\sum_{k=1}^{j-1}w_{n(k)}(z,h)}\big)\Big)\\
=&\prod_{j=1}^{i-1}\big(1-\frac{w_{n(j)}(z,h)\Delta_{(j)}}{1-\sum_{k=1}^{j-1}w_{n(k)}(z,h)}\big)-\prod_{j=1}^{i}\big(1-\frac{w_{n(j)}(z,h)\Delta_{(j)}}{1-\sum_{k=1}^{j-1}w_{n(k)}(z,h)}\big)\\
=&\Big(1-\big(1-\frac{w_{n(i)}(z,h)\Delta_{(i)}}{1-\sum_{k=1}^{i-1}w_{n(k)}(z,h)}\big)\Big)\cdot\prod_{j=1}^{i-1}\big(1-\frac{w_{n(j)}(z,h)\Delta_{(j)}}{1-\sum_{k=1}^{j-1}w_{n(k)}(z,h)}\big)\\
=&\frac{w_{n(i)}(z,h)\Delta_{(i)}}{1-\sum_{k=1}^{i-1}w_{n(k)}(z,h)}\cdot\prod_{j=1}^{i-1}\big(1-\frac{w_{n(j)}(z,h)\cdot\Delta_{(j)}}{1-\sum_{k=1}^{j-1}w_{n(k)}(z,h)}\big)\\
=&\frac{w_{n(i)}(z,h)\Delta_{(i)}}{1-\sum_{k=1}^{i-1}w_{n(k)}(z,h)}\cdot\prod_{j=1}^{i-1}\Big(\frac{1-\sum_{k=1}^{j-1}w_{n(k)}(z,h)-w_{n(j)}(z,h)\Delta_{(j)}}{1-\sum_{k=1}^{j-1}w_{n(k)}(z,h)}\Big)\\
=&w_{n(i)}(z,h)\Delta_{(i)}\prod_{j=1}^{i-1}\Big(\frac{1-\sum_{k=1}^{j-1}w_{n(k)}(z,h)-w_{n(j)}(z,h)\Delta_{(j)}}{1-\sum_{k=1}^{j}w_{n(k)}(z,h)}\Big)\\
=&w_{n(i)}(z,h)\Delta_{(i)}\prod_{j=1}^{i-1}\Big(1+\frac{w_{n(j)}(z,h)(1-\Delta_{(j)})}{1-\hat{B}_n(Y_{(j)}|z)}\Big).
\end{aligned}
$$

Thus, for $z\in[0,1]^d$, we have

$$
\begin{aligned}
&\int\psi_n(x,z)d\hat{H}_n^{KM}(x|z)\\
=&\sum_{i=1}^{n}\psi_n(Y_{(i)},z)\Delta_{(i)}w_{n(i)}(z,h)\prod_{j=1}^{i-1}\Big(1+\frac{w_{n(j)}(z,h)(1-\Delta_{(j)})}{1-\hat{B}_n(Y_{(j)}|z)}\Big)\\
=&\sum_{i=1}^{n}\psi_n(Y_i,z)\Delta_i w_{ni}(z,h)\prod_{j=1}^{n}I(Y_j<Y_i)\Big(1+\frac{w_{nj}(z,h)(1-\Delta_j)}{1-\hat{B}_n(Y_j|z)}\Big)\\
=&\sum_{i=1}^{n}\psi_n(Y_i,z)\Delta_i w_{ni}(z,h)\exp\sum_{j=1}^{n}I(Y_j<Y_i)\ln\Big(1+\frac{w_{nj}(z,h)(1-\Delta_j)}{1-\hat{B}_n(Y_j|z)}\Big)
\end{aligned}
$$

$$=\sum_{i=1}^{n}\psi_n(Y_i,z)\Delta_i w_{ni}(z,h)\exp\sum_{j=1}^{n}I(Y_j<Y_i)(1-\Delta_j)\ln\Big(1+\frac{w_{nj}(z,h)}{1-\hat{B}_n(Y_j|z)}\Big).$$

For $i\in\{1,...,n\}$, define the functions $A_{in},B_{in},C_{in}:[0,1]^d\to\mathbb{R}$ with

$$\begin{aligned}
A_{in}(z):=&\sum_{j=1}^{n}I(Y_j<Y_i)(1-\Delta_j)\ln\Big(1+\frac{w_{nj}(z,h)}{1-\hat{B}_n(Y_j|z)}\Big)\\
&-\int_{-\infty}^{Y_i-}\frac{1}{1-E\big[\hat{B}_n(x|z)\big]}dE\big[\hat{B}_n^0(x|z)\big],\\
B_{in}(z):=&\sum_{j=1}^{n}I(Y_j<Y_i)(1-\Delta_j)\ln\Big(1+\frac{w_{nj}(z,h)}{1-\hat{B}_n(Y_j|z)}\Big)\\
&-\int_{-\infty}^{Y_i-}\frac{1}{1-\hat{B}_n(x|z)}d\hat{B}_n^0(x|z),\\
C_{in}(z):=&\int_{-\infty}^{Y_i-}\frac{1}{1-\hat{B}_n(x|z)}d\hat{B}_n^0(x|z)-\int_{-\infty}^{Y_i-}\frac{1}{1-E\big[\hat{B}_n(x|z)\big]}dE\big[\hat{B}_n^0(x|z)\big].
\end{aligned}$$

By a Taylor expansion, for $a,b\in\mathbb{R}$ there exists a constant ξ lying between a and b such that

$$\exp(a)=\exp(b)+\exp(b)(a-b)+\frac{1}{2}\exp(\xi)(a-b)^2.$$

Thus, by setting

$$a=\sum_{j=1}^{n}I(Y_j<Y_i)(1-\Delta_j)\ln\Big(1+\frac{w_{nj}(z,h)}{1-\hat{B}_n(Y_j|z)}\Big)$$

and

$$b=\int_{-\infty}^{Y_i-}\frac{1}{1-E\big[\hat{B}_n(x|z)\big]}dE\big[\hat{B}_n^0(x|z)\big]$$

we get

$$\begin{aligned}
&\exp\sum_{j=1}^{n}I(Y_j<Y_i)(1-\Delta_j)\ln\Big(1+\frac{w_{nj}(z,h)}{1-\hat{B}_n(Y_j|z)}\Big)\\
=&\gamma_n(Y_i|z)+\gamma_n(Y_i|z)A_{in}(z)+\frac{1}{2}\exp\big(\xi_i(z)\big)A_{in}^2(z)\\
=&\gamma_n(Y_i|z)\big(1+B_{in}(z)+C_{in}(z)\big)+\frac{1}{2}\exp\big(\xi_i(z)\big)\big(B_{in}(z)+C_{in}(z)\big)^2
\end{aligned}$$

where $\xi_i(z)$ lies between

$$\sum_{j=1}^{n}I(Y_j<Y_i)(1-\Delta_j)\ln\Big(1+\frac{w_{nj}(z,h)}{1-\hat{B}_n(Y_j|z)}\Big)$$

and

$$\int_{-\infty}^{Y_i-} \frac{1}{1-E\big[\hat{B}_n(x|z)\big]} dE\big[\hat{B}_n^0(x|z)\big].$$

Thus, we obtain

$$\begin{aligned}
&\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\int\psi_n(x,z_j)d\hat{H}_n^{KM}(x|z_j)\\
=&\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n}\sum_{j=1}^{n_0}\psi_n(Y_i,z_j)\Delta_i w_{ni}(z_j,h)\gamma_n(Y_i|z_j)\big(1+B_{in}(z_j)+C_{in}(z_j)\big)\\
&+\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n}\sum_{j=1}^{n_0}\frac{1}{2}\psi_n(Y_i,z_j)\Delta_i w_{ni}(z_j,h)e^{\xi_i(z_j)}\big(B_{in}(z_j)+C_{in}(z_j)\big)^2. \quad \text{(A.4.8)}
\end{aligned}$$

First we have

$$\begin{aligned}
&\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n}\sum_{j=1}^{n_0}\psi_n(Y_i,z)\Delta_i w_{ni}(z_j,h)\gamma_n(Y_i|z_j)\\
=&\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\int\psi_n(x,z_j)\gamma_n(x|z_j)\,d\hat{B}_n^1(x|z_j). \qquad \text{(A.4.9)}
\end{aligned}$$

In the sequel, we show that

$$\begin{aligned}
&\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n}\sum_{j=1}^{n_0}\psi_n(Y_i,z)\Delta_i w_{ni}(z_j,h)\gamma_n(Y_i|z_j)C_{in}(z_j)\\
=&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iint\psi_n(u,z_j)C_n(x\wedge u|z_j)\gamma_n(u|z_j)dE[\hat{B}_n^1(u|z_j)]d\hat{B}_n(x|z_j)\\
&+\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iint\frac{I(x<u)\psi_n(u,z_j)\gamma_n(u|z_j)}{1-E[\hat{B}_n(x|z_j)]}dE[\hat{B}_n^1(u|z_j)]d\hat{B}_n^0(x|z_j)+o_p(1).
\end{aligned}$$

(A.4.10)

Note that we can write

$$\begin{aligned}
C_{in}(z)=&-\int_{-\infty}^{Y_i-}\int\frac{I(x<u)}{\big(1-E\big[\hat{B}_n(x|z)\big]\big)^2}d\hat{B}_n(u|z)d\hat{B}_n^0(x|z)\\
&+\int_{-\infty}^{Y_i-}\frac{2}{1-E\big[\hat{B}_n(x|z)\big]}d\hat{B}_n^0(x|z)-\int_{-\infty}^{Y_i-}\frac{1}{1-E\big[\hat{B}_n(x|z)\big]}dE\big[\hat{B}_n^0(x|z)\big]\\
&+\int_{-\infty}^{Y_i-}\frac{\big(\hat{B}_n(x|z)-E\big[\hat{B}_n(x|z)\big]\big)^2}{\big(1-E\big[\hat{B}_n(x|z)\big]\big)^2\big(1-\hat{B}_n(x|z)\big)}d\hat{B}_n^0(x|z).
\end{aligned}$$

For simplicity of notation we denote the functions $(\tilde{\psi}_n)_{n\in m\cdot\mathbb{N}}:\mathbb{R}^2\times[0,1]^d\to\mathbb{R}$ with

$$\tilde{\psi}_n(x,u,z):=I(u<x)\psi_n(x,z)\gamma_n(x|z)/\big(1-E\big[\hat{B}_n(u|z)\big]\big),$$

then we get

$$
\begin{aligned}
&\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n}\sum_{j=1}^{n_0}\psi_n(Y_i,z_j)\Delta_i w_{ni}(z,h)\gamma_n(Y_i|z_j)C_{in}(z_j)\\
=&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iiint\frac{I(u<t)\tilde{\psi}_n(x,u,z_j)}{1-E\big[\hat{B}_n(u|z_j)\big]}d\hat{B}_n(t|z_j)d\hat{B}_n^0(u|z_j)d\hat{B}_n^1(x|z_j)\\
&+2\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iint\tilde{\psi}_n(x,u,z_j)d\hat{B}_n^0(u|z_j)d\hat{B}_n^1(x|z_j)\\
&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iint\tilde{\psi}_n(x,u,z_j)dE\big[\hat{B}_n^0(u|z_j)\big]d\hat{B}_n^1(x|z_j)\\
&+\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iint\frac{\tilde{\psi}_n(x,u,z_j)\big(\hat{B}_n(u|z_j)-E\big[\hat{B}_n(u|z_j)\big]\big)^2}{\big(1-E\big[\hat{B}_n(u|z_j)\big]\big)\big(1-\hat{B}_n(u|z_j)\big)}d\hat{B}_n^0(u|z_j)d\hat{B}_n^1(x|z_j)\\
=:&\bar{C}_{n1}+\bar{C}_{n2}+\bar{C}_{n3}+\bar{C}_{n4}.
\end{aligned}
$$

By Lemma A.2.9, with $k=3$, $X_{1j}=Y_j$, $X_{2j}=(1-\Delta_j)\cdot Y_j$, $H_{3j}=\Delta_j\cdot Y_j$ for each $j\in\{1,...,n\}$ and

$$
\psi_n(x,u,t,z)=\frac{I(u<t)\tilde{\psi}_n(x,u,z)}{1-E\big[\hat{B}_n(u|z)\big]}
$$

which is a bounded function, we get

$$
\begin{aligned}
\bar{C}_{n1}=&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iiint\frac{I(u<t)\tilde{\psi}_n(x,u,z_j)}{1-E\big[\hat{B}_n(u|z_j)\big]}\Big[d\hat{B}_n(t|z_j)dE\big[\hat{B}_n^0(u|z_j)\big]dE\big[\hat{B}_n^1(x|z_j)\big]\\
&+dE\big[\hat{B}_n(t|z_j)\big]d\hat{B}_n^0(u|z_j)dE\big[\hat{B}_n^1(x|z_j)\big]+dE\big[\hat{B}_n(t|z_j)\big]dE\big[\hat{B}_n^0(u|z_j)\big]d\hat{B}_n^1(x|z_j)\\
&-2dE\big[\hat{B}_n(t|z_j)\big]E\big[\hat{B}_n^0(u|z_j)\big]dE\big[\hat{B}_n^1(x|z_j)\big]\Big]+o_p(1)\\
=&-\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iiint\frac{I(u<t)\tilde{\psi}_n(x,u,z_j)}{1-E\big[\hat{B}_n(u|z_j)\big]}d\hat{B}_n(t|z_j)dE\big[\hat{B}_n^0(u|z_j)\big]dE\big[\hat{B}_n^1(x|z_j)\big]\\
&+\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iint\tilde{\psi}_n(x,u,z_j)\Big[d\hat{B}_n^0(u|z_j)dE\big[\hat{B}_n^1(x|z_j)\big]\\
&\quad+dE\big[\hat{B}_n^0(u|z_j)\big]d\hat{B}_n^1(x|z_j)-2dE\big[\hat{B}_n^0(u|z_j)\big]dE\big[\hat{B}_n^1(x|z_j)\big]\Big]+o_p(1).
\end{aligned}
$$

With similar arguments, it can be shown that

$$
\bar{C}_{n2}=2\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iint\tilde{\psi}_n(x,u,z_j)\Big[d\hat{B}_n^0(u|z_j)dE\big[\hat{B}_n^1(x|z_j)\big]
$$

$$+ dE\big[\hat{B}_n^0(u|z_j)\big]d\hat{B}_n^1(x|z_j) - dE\big[\hat{B}_n^0(u|z_j)\big]dE\big[\hat{B}_n^1(x|z_j)\big]\Big] + o_p(1).$$

By Lemma A.4.1 and the assumption $\psi_n(x,z) = 0$ if $x > \tau$ and $z \in [0,1]^d$, there exists a $C > 0$ such that $\bar{C}_{n4}$ is bounded by

$$\begin{aligned}
&\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iint \frac{\tilde{\psi}_n(x,u,z_j)\big(\hat{B}_n(u|z_j) - E\big[\hat{B}_n(u|z_j)\big]\big)^2}{\big(1 - E\big[\hat{B}_n(\tau|z_j)\big]\big)\big(1 - \hat{B}_n(u|z_j)\big)} d\hat{B}_n^0(u|z_j)d\hat{B}_n^1(x|z_j)\\
\le& C\sqrt{n}\cdot\frac{1}{n_0}\sum_{j=1}^{n_0}\iint \frac{\tilde{\psi}_n(x,u,z_j)\big(\hat{B}_n(u|z_j) - E\big[\hat{B}_n(u|z_j)\big]\big)^2}{\big(1 - E\big[\hat{B}_n(u|z_j)\big]\big)} d\hat{B}_n^0(u|z_j)d\hat{B}_n^1(x|z_j)
\end{aligned}$$

which is equal to $o_p(1)$ by Lemma A.2.9 as in the proof of lemma A.3.2. Thus,

$$\bar{C}_{n4} = o_p(1).$$

Consequently, the Equality (A.4.10) holds.

Next, we show that

$$\sqrt{n}\cdot\frac{1}{n_0}\sum_{i=1}^{n}\sum_{j=1}^{n_0}\psi_n(Y_i,z_j)\Delta_i w_{ni}(z_j,h)\gamma_n(Y_i|z_j)B_{in}(z_j) = o_p(1). \qquad \text{(A.4.11)}$$

Since for $a \ge 0$,

$$-\frac{a^2}{2} \le \ln(1+a) - a \le 0.$$

By setting

$$a = \frac{w_{nj}(x,h)}{1 - \hat{B}_n(Y_j|z)},$$

we get

$$-\frac{1}{2}\cdot\frac{w_{nj}^2(x,h)}{\big(1 - \hat{B}_n(Y_j|z)\big)^2} \le \ln\left(1 + \frac{w_{nj}(x,h)}{1 - \hat{B}_n(Y_j|z)}\right) - \frac{w_{nj}(x,h)}{1 - \hat{B}_n(Y_j|z)} \le 0.$$

Thus,

$$-\frac{1}{2}\sum_{j=1}^{n} I(Y_j < Y_i)(1 - \Delta_j)\cdot\frac{w_{nj}^2(z,h)}{\big(1 - \hat{B}_n(Y_j|z)\big)^2} \le B_{in}(z) \le 0.$$

By (A.2.12), there exists a constant $C > 0$, such that for n large enough, for all $z \in [0,1]$ and $i \in \{1,...,n\}$ with $Y_i \le \tau$

$$|B_{in}(z)| \le \frac{1}{2}\sum_{j=1}^{n} I(Y_j < Y_i)(1 - \Delta_j)\cdot\frac{w_{nj}^2(z,h)}{\big(1 - \hat{B}_n(Y_j|z)\big)^2}$$

$$
\begin{aligned}
\leq & C \cdot \frac{1}{n_0 h^d} \sum_{j=1}^{n} I(Y_j < Y_i)(1-\Delta_j) \cdot \frac{w_{nj}(z,h)}{\big(1-\hat{B}_n(Y_j|z)\big)^2} \\
= & C \cdot \frac{1}{n_0 h^d} \int_{-\infty}^{Y_i-} \frac{1}{\big(1-\hat{B}_n(x|z)\big)^2} d\hat{B}_n^0(x|z) \\
\leq & C \cdot \frac{1}{n_0 h^d} \int_{-\infty}^{Y_i-} \frac{1}{\big(1-\hat{B}_n(\tau|z)\big)^2} d\hat{B}_n^0(x|z).
\end{aligned}
$$

By Lemma A.4.1 there exists a constant $C > 0$, such that for n large enough,

$$
|B_{in}(z)| \leq C \cdot \frac{1}{n_0 h^d} \text{ in probability.}
$$

Therefore, we obtain

$$
\begin{aligned}
& \Big| \sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n} \sum_{j=1}^{n_0} \psi_n(Y_i, z_j) \Delta_i w_{ni}(z_j,h) \gamma_n(Y_i|z_j) B_{in}(z_j) \Big| \\
\leq & C \cdot \frac{\sqrt{m}}{\sqrt{n_0} h^d} \cdot \frac{1}{n_0} \sum_{i=1}^{n} \sum_{j=1}^{n_0} |\psi_n(Y_i, z_j)| \Delta_i w_{ni}(z_j,h) \gamma_n(Y_i|z_j) \\
= & C \cdot \frac{\sqrt{m}}{\sqrt{n_0} h^d} \int |\psi_n(x,z)| \gamma_n(x|z) d\hat{Q}_n^1(x,z) = o_p(1)
\end{aligned}
$$

where the last step follow from the assumption $n_0 h^{2d} \to \infty$ and the boundedness of the $|\psi_n|\gamma_n$. Hence, Equality (A.4.11) holds.

By similar arguments, we can show

$$
\sqrt{n} \cdot \frac{1}{n_0} \sum_{i=1}^{n} \sum_{j=1}^{n_0} \frac{1}{2} \psi_n(Y_i, z_j) \Delta_i w_{ni}(z_j,h) e^{\xi_i(z_j)} \big(B_{in}(z_j) + C_{in}(z_j)\big)^2 = o_p(1). \tag{A.4.12}
$$

Therefore, the assertion follows from Equalities (A.4.8)–(A.4.12). □

Lemma A.4.3. *Let* $(\psi_n)_{n \in m \cdot \mathbb{N}} : \mathbb{R} \times [0,1]^d$ *be a sequence of uniformly bounded functions with* $\psi_n(x,y,z) = 0$ *for any* $n \in m \cdot \mathbb{N}$*, if* $\max(x,y) > \tau$*. Then we have*

$$
\begin{aligned}
\sqrt{n} \cdot \frac{1}{n_0} \sum_{k=1}^{n_0} \iint \psi_n(x,y,z_k) \Big[& d\hat{H}_n^{KM}(x|z_k) d\hat{H}_n^{KM}(y|z_k) \\
& - \gamma_n(y|z_k) d\hat{H}_n^{KM}(x|z_k) dE[\hat{B}_n^1(y|z_k)] - \gamma_n(x|z_k) dE[\hat{B}_n^1(x|z_k)] d\hat{H}_n^{KM}(y|z_k) \\
& + \gamma_n(x|z_k) \gamma_n(y|z_k) dE[\hat{B}_n^1(x|z_k)] dE[\hat{B}_n^1(y|z_k)] \Big] = o_p(1).
\end{aligned}
$$

Proof. This lemma can be shown analogously as Lemma A.4.2. □

Lemma A.4.4. *Let* $(\psi_n)_{n\in m\cdot\mathbb{N}} : \mathbb{R}\times[0,1]^d$ *be a sequence of uniformly bounded functions with* $\psi_n(x_1,x_2,x_3,z) = 0$ *for any* $n \in m\cdot\mathbb{N}$, *if* $\max(x_1,x_2,x_3) > \tau$. *Then we have*

$$\begin{aligned}
\sqrt{n}\cdot\frac{1}{n_0}\sum_{k=1}^{n_0}\iiint \psi_n(x_1,x_2,x_3,z_k)\Big[&d\hat{H}_n^{KM}(x_1|z_k)d\hat{H}_n^{KM}(x_2|z_k)d\hat{H}_n^{KM}(x_3|z_k)\\
&-\gamma_n(x_2|z_k)\gamma_n(x_3|z_k)d\hat{H}_n^{KM}(x_1|z_k)dE[\hat{B}_n^1(x_2|z_k)]dE[\hat{B}_n^1(x_3|z_k)]\\
&-\gamma_n(x_1|z_k)\gamma_n(x_3|z_k)dE[\hat{B}_n^1(x_1|z_k)]d\hat{H}_n^{KM}(x_2|z_k)dE[\hat{B}_n^1(x_3|z_k)]\\
&-\gamma_n(x_1|z_k)\gamma_n(x_2|z_k)dE[\hat{B}_n^1(x_1|z_k)]dE[\hat{B}_n^1(x_2|z_k)]d\hat{H}_n^{KM}(x_3|z_k)\\
&-\gamma_n(x_1|z_k)\gamma_n(x_2|z_k)\gamma_n(x_3|z_k)dE[\hat{B}_n^1(x_1|z_k)]dE[\hat{B}_n^1(x_2|z_k)]dE[\hat{B}_n^1(x_3|z_k)]\Big] = o_p(1).
\end{aligned}$$

Proof. This lemma can be shown analogously as Lemma A.4.2. □

Lemma A.4.5. *There exists a constant* $C > 0$, *such that for all* $(x,z) \in (-\infty,\tau]\times[0,1]^d$ *and eventually all* $n \in m\cdot\mathbb{N}$,

$$\big|\gamma_n(x|z) - \gamma(x|z)\big| \le Ch$$

and for all $(x,z_i) \in (-\infty,\tau]\times S_h$ *and eventually all* $n \in m\cdot\mathbb{N}$,

$$\big|\gamma_n(x|z_i) - \gamma(x|z_i)\big| \le Ch^2.$$

Proof. By a Taylor expansion, there exists a function $\tilde{\gamma}$ lying between γ_n and γ such that

$$\gamma_n(x|z) - \gamma(x|z) = \tilde{\gamma}(x|z)\big(\ln\gamma_n(x|z) - \ln\gamma(x|z)\big).$$

By Lemma A.4.1, it can be seen that the functions γ_n and γ are bounded on $(-\infty,\tau]\times[0,1]^d$. Thus, the function $\tilde{\gamma}$ is also bounded on $(-\infty,\tau]\times[0,1]^d$. Further, for any $(x,z) \in (-\infty,\tau]\times[0,1]^d$,

$$\begin{aligned}
&\big|\ln\gamma_n(x|z) - \ln\gamma(x|z)\big|\\
\le&\Big|\int_{-\infty}^{x-}\frac{1}{1-E\big[\hat{B}_n(u|z)\big]} - \frac{1}{1-B(u|z)}dE\big[\hat{B}_n^0(u|z)\big]\Big|\\
&+\Big|\int_{-\infty}^{x-}\frac{1}{1-B(u|z)}d\big(E[\hat{B}_n^0(u|z)] - B^0(u|z)\big)\Big|\\
\le&\int_{-\infty}^{x-}\frac{\big|E[\hat{B}_n(u|z)] - B(u|z)\big|}{\big(1-E[\hat{B}_n(u|z)]\big)\big(1-B(u|z)\big)}dE\big[\hat{B}_n^0(u|z)\big]
\end{aligned}$$

$$+\Big|\frac{\big(E[\hat{B}_n^0(x|z)]-B^0(x|z)\big)}{1-B(x|z)}-\int_{-\infty}^{x-}\frac{\big(E[\hat{B}_n^0(u|z)]-B^0(u|z)\big)}{\big(1-B(u|z)\big)^2}dB(u|z)\Big|$$

$$\leq\int_{-\infty}^{x-}\frac{\big|E[\hat{B}_n(u|z)]-B(u|z)\big|}{\big(1-E[\hat{B}_n(\tau|z)]\big)\big(1-B(\tau|z)\big)}dE\big[\hat{B}_n^0(u|z)\big]$$

$$+\frac{\big|E[\hat{B}_n^0(x|z)]-B^0(x|z)\big|}{1-B(\tau|z)}+\int_{-\infty}^{x-}\frac{\big|E[\hat{B}_n^0(u|z)]-B^0(u|z)\big|}{\big(1-B(\tau|z)\big)^2}dB(u|z).$$

Thus, by Lemma A.4.1 there exists a constant $C>0$ such that $\big|\ln\gamma_n(x|z)-\ln\gamma(x|z)\big|$ is bounded by

$$C\cdot\int_{-\infty}^{x-}\big|E[\hat{B}_n(u|z)]-B(u|z)\big|dE\big[\hat{B}_n^0(u|z)\big]$$

$$+C\cdot\big|E[\hat{B}_n^0(x|z)]-B^0(x|z)\big|+C\cdot\int_{-\infty}^{x-}\big|E[\hat{B}_n^0(u|z)]-B^0(u|z)\big|dB(u|z)$$

$$\leq 3C\cdot\max_{x\in(-\infty,\tau]}\big|E[\hat{B}_n^0(x|z)]-B^0(x|z)\big|$$

Hence, the assertion follows by similar arguments as used in Lemma A.2.6. □

Lemma A.4.6. *There exists a constant $C>0$, such that for all $(x,z)\in(-\infty,\tau]\times[0,1]^d$ and eventually all $n\in m\cdot\mathbb{N}$,*

$$\Big|\int_{-\infty}^{x}\gamma_n(u|z)dE[\hat{B}_n^1(u|z)]-H(x|z)\Big|\leq Ch$$

and for all $(x,z_i)\in(-\infty,\tau]\times S_h$ and eventually all $n\in m\cdot\mathbb{N}$,

$$\Big|\int_{-\infty}^{x}\gamma_n(u|z_i)dE[\hat{B}_n^1(u|z_i)]-H(x|z_i)\Big|\leq Ch^2.$$

Proof. We write first

$$\Big|\int_{-\infty}^{x}\gamma_n(u|z)dE[\hat{B}_n^1(u|z)]-H(x|z)\Big|$$

$$=\Big|\int_{-\infty}^{x}\gamma_n(u|z)dE[\hat{B}_n^1(u|z)]-\int_{-\infty}^{x}\gamma(u|z)dH(u|z)\Big|$$

$$\leq\Big|\int_{-\infty}^{x}\big(\gamma_n(u|z)-\gamma(u|z)\big)dE[\hat{B}_n^1(u|z)]\Big|+\Big|\int_{-\infty}^{x}\gamma(u|z)d\big(E[\hat{B}_n^1(u|z)]-B^1(u|z)\big)\Big|$$

$$\leq\int_{-\infty}^{x}\big|\gamma_n(u|z)-\gamma(u|z)\big|dE[\hat{B}_n^1(u|z)]+\big|E[\hat{B}_n^1(x|z)]-B^1(x|z)\big|\cdot\gamma(x|z)$$

$$+\int_{-\infty}^{x}\frac{\big|E[\hat{B}_n^1(u|z)]-B^1(u|z)\big|}{\big(1-J(u|z)\big)^2}dJ(u|z).$$

By similar arguments as used in Lemma A.2.6, the assertion follows from Lemma A.4.5. □

Lemma A.4.7. *Let $\psi : \mathbb{R} \times [0,1]^d$ be a bounded functions, with $\psi(x,z) = 0$ if $x > \tau$ and $z \in [0,1]^d$. If $\partial\psi/\partial z$ is dominated by a B^1 integrable function, then*

$$\frac{1}{n_0}\sum_{j=1}^{n_0}\int \psi(x,z_j)\,d\hat{H}_n^{KM}(x|z_j) = \int \psi(x,z)\,dQ(x,z) + o_p(1).$$

Proof. By Lemma A.4.2 we get

$$\begin{aligned}
&\frac{1}{n_0}\sum_{j=1}^{n_0}\int \psi(x,z_j)\,d\hat{H}_n^{KM}(x|z_j)\\
=&\frac{1}{n_0}\sum_{j=1}^{n_0}\int \psi(x,z_j)\gamma_n(x|z_j)\,d\hat{B}_n^1(x|z_j)\\
&-\frac{1}{n_0}\sum_{j=1}^{n_0}\iint \psi(u,z_j)C_n(x\wedge u|z_j)\gamma_n(u|z_j)dE[\hat{B}_n^1(u|z_j)]d\hat{B}_n(x|z_j)\\
&+\frac{1}{n_0}\sum_{j=1}^{n_0}\iint \frac{I(x<u)\psi(u,z_j)\gamma_n(u|z_j)}{1-E[\hat{B}_n(x|z_j)]}\,dE[\hat{B}_n^1(u|z_j)]d\hat{B}_n^0(x|z_j) + o_p(1).
\end{aligned}$$

Note that the functions

$$\psi(x,z),\quad \int \psi(u,z)C_n(x\wedge u|z)\gamma_n(u|z)dE[\hat{B}_n^1(u|z)],$$

and

$$\int \frac{I(x<u)\psi(u,z)\gamma_n(u|z)}{1-E[\hat{B}_n(x|z)]}\,dE[\hat{B}_n^1(u|z)]$$

are uniformly bounded for all $n \in m\cdot\mathbb{N}$ and $(x,z) \in \mathbb{R}\times[0,1]^d$. Thus, analogously to Lemma A.2.11, it can be shown that

$$\begin{aligned}
&\frac{1}{n_0}\sum_{j=1}^{n_0}\int \psi(x,z_j)\,d\hat{H}_n^{KM}(x|z_j)\\
=&\frac{1}{n_0}\sum_{j=1}^{n_0}\int \psi(x,z_j)\gamma_n(x|z_j)\,dE[\hat{B}_n^1(x|z_j)]\\
&-\frac{1}{n_0}\sum_{j=1}^{n_0}\iint \psi(u,z_j)C_n(x\wedge u|z_j)\gamma_n(u|z_j)dE[\hat{B}_n^1(u|z_j)]dE[\hat{B}_n(x|z_j)]\\
&+\frac{1}{n_0}\sum_{j=1}^{n_0}\iint \frac{I(x<u)\psi(u,z_j)\gamma_n(u|z_j)}{1-E[\hat{B}_n(x|z_j)]}\,dE[\hat{B}_n^1(u|z_j)]dE[\hat{B}_n^0(x|z_j)] + o_p(1)\\
=&\frac{1}{n_0}\sum_{j=1}^{n_0}\int \psi(x,z_j)\gamma_n(x|z_j)\,dE[\hat{B}_n^1(x|z_j)] + o_p(1)
\end{aligned}$$

where the last term follows from similar arguments as in (3.2.1). Further Lemma A.4.5 implies

$$\frac{1}{n_0}\sum_{j=1}^{n_0}\int \psi(x,z_j)\gamma_n(x|z_j)\,dE[\hat{B}_n^1(x|z_j)] = \frac{1}{n_0}\sum_{j=1}^{n_0}\int \psi(x,z_j)\gamma(x|z_j)\,dE[\hat{B}_n^1(x|z_j)]+o(1).$$

It follows from Assumption (i) that the function γ has a bounded derivative with respect to z. Thus, it can be shown analogously as Lemma A.2.12 that

$$\begin{aligned}
&\frac{1}{n_0}\sum_{j=1}^{n_0}\int \psi(x,z_j)\gamma(x|z_j)\,dE[\hat{B}_n^1(x|z_j)]\\
=&\frac{1}{n_0}\sum_{j=1}^{n_0}\int \psi(x,z_j)\gamma(x|z_j)\,dB^1(x|z_j)+o(1)\\
=&\int \psi(x,z)\,dQ(x,z)+o(1).
\end{aligned}$$

Hence, the assertion follows. □

Bibliography

[1] Aerts M. Janssen P. Veraverbeke N. Asymptotic theory for regression quantile estimators in the heteroscedastic regression model. *Asymptotic Statistics*, pages 151–161, 1994.

[2] Akaike H. Information theory and an extension of the maximum likelihood principle. *The second international symposium on information theory. Budapest*, pages 267–281, 1973.

[3] Akritas M. G. Nearest neighbor estimation of a bivariate distribution under random censoring. *Ann. Statist.*, 22:1299–1327, 1994.

[4] Akritas M.G. The central limit theorem under censoring. *Bernoulli*, 6:1109–1120, 2000.

[5] Andrews D.W.K. A conditional Kolmogorov test. *Econometrica*, 65:1097–1128, 1997.

[6] Basseville M. Divergence measures for statistical data processing. *Technical Report IRISA*, 2010.

[7] Beran R. Nonparametric regression with randomly censored survival data. *Tech. Rep. University of California, Berkeley.*, 1981.

[8] Bobrowski S. Döring M. Jensen U. Schinköthe W. Reliability prediction using the Cox proportional hazards model. *The 56th international scientific colloquium, Ilmenau*, 2011.

[9] Bobrowski S. Chen H. Döring M. Jensen U. Schinköthe W. Estimation of the lifetime distribution of mechatronic systems in the presence of a covariate: A comparison among parametric, semiparametric and nonparametric models. *Reliab. Eng. Syst. Safety*, 139:105–112, 2015.

[10] Chen X. Hong H. Shum M. Nonparametric likelihood ratio model selection tests between parametric likelihood and moment condition models. *J. Econometrics*, 141:109–140, 2007.

[11] Chen H. Döring M. Jensen U. Model selection using Cramér-von Mises distance. Stochastic Models, Statistics and Their Applications. . *Springer, Proceedings in Mathematics & Statistics*, 122:39–47, 2015.

[12] Claeskens G. Hjort N.L. *Model selection and model averaging*. Cambridge University Press, Cambridge, 2008.

[13] Dabrowska D. M. Non-parametric regression with censored survival time data. *Scand. J. Statist.*, 14:181–197, 1987.

[14] Dabrowska D. M. Uniform consistency of the kernel conditional Kaplan-Meier estimate. *Ann. Statist.*, 17:1157–1167, 1989.

[15] Du Y. Akritas M. G. Uniform strong representation of the conditional Kaplan-Meier process. *Math. Methods Statist.*, 11(2):152–182, 2002.

[16] Ducharme G. Ferrigno S. An omnibus test of goodness-of-fit for conditional distribution functions with applications to regression models. *J. Statist. Plann. Inference*, 142:2748–2761, 2012.

[17] Durbin J. Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.*, 1(2):279–290, 1973.

[18] Durbin J. Kolmogorov-Smirnov tests when parameters are estimated with applications to tests of exponentiality and tests on spacings. *Biometrika*, 62(1):5–22, 1975.

[19] Durbin J. The first-passage density of a continuous Gaussian process to a general boundary. *J. Appl. Probab.*, 22(1):99–122, 1985.

[20] Ferraty F. Vieu P. *Nonparametric functional data analysis theory and practice*. Springer, Berlin, 2006.

[21] Földes A. Rejto L. Strong uniform consistency for nonparametric survival curve estimators from randomly censored data. *Ann. Statist.*, 9(1):122–129, 1981.

[22] Gasser T. Müller H.G. Estimating regression functions and their derivatives by the kernel method. *Scand. J. Stat.*, 1:171–185, 1984.

[23] Gill R. D. Large sample behavior of the product-limit estimator on the whole line. *Ann. Statist.*, 11:49–58, 1983.

[24] González M. W. Cadarso S. C. Asymptotic properties of a generalized Kaplan-Meier estimator with some applications. *J. Nonparametric Statistics*, 4:65–78, 1994.

[25] Györfi L. Kohler M. Krzyżak A. Walk H. *A distribution-free theory of nonparametric regression.* Springer, 2002.

[26] Inoue A. Kilian L. On the selection of forecasting models. *J. Econometrics*, 130(2):273–306, 2006.

[27] Kaplan E.L. Meier P. Nonparametric estimation from incomplete observations. *J. Am. Stat. Assoc.*, 53:457–481, 1958.

[28] Khmaladze E. Martingale approach in the theory of goodness-of-fit tests. *Theory Probab. Appl.*, 26(2):240–257, 1981.

[29] Kitamura Y. Comparing misspecified dynamic econometric models using nonparametric likelihood. *Unpublished Manuscript University of Wisconsin*, 2001.

[30] Kosorok M.R. *Introduction to empirical processes and semiparametric inference.* Springer, New York, 2008.

[31] Li F. Tkacz G. A consistent test for multivariate conditional distributions. *Econometric reviews*, 30(3):251–273, 2011.

[32] Li Q. Racine J.S. *Nonparametric econometrics: theory and practice.* Princeton University Press, 2007.

[33] Li Q. Lin J. Racine J.S. Optimal bandwidth selection for nonparametric conditional distribution and quantile functions. *J. Bus. Econom. Statist.*, 31(1):57–65, 2013.

[34] Liebscher E. Approximation of distributions by using the Anderson-Darling statistic. *Commun. Stat. Theory Methods*, Published online: 25 Jan 2016.

[35] Lin D. Y. Wei L. J. The robust inference for the Cox proportional hazards model. *Am. Stat. Assoc.*, 84:1074–1078, 1989.

[36] Lo S. Singh K. The product limit estimator and the bootstrap: some asymptotic representations. *Probab. Theory Related Fields*, 71:455–466, 1986.

[37] McKeague I. W. Utikal K. J. Inference for a nonlinear counting process regression model. *Ann. Statist.*, 18:1172–1187, 1990.

[38] Nadaraya E.C. On estimating regression. *Theor. Probab. Appl.*, 9:141–142, 1964.

[39] Ng C. T. Joe H. Comparison of non-nested models under a general measure of distance. *J. Statist. Plann. Inference*, 170:166–185, 2016.

[40] Nikabadze A. Stute W. Model checks under random censorship. *Statist. and Prob. Letters*, 32:249–259, 1997.

[41] Nishii R. Maximum likelihood principle and model selection when the true model is unspecified. *J. Multivariate Anal.*, 27(2):392–403, 1988.

[42] Rothe C. Wied D. Misspecification testing in a class of conditional distributional models. *J. Am. Stat. Assoc.*, 108:314–324, 2013.

[43] Schick A. Susarla V. Koul H. Efficient estimation of functionals with censored data. *Statist. Decisions*, 6:349–360, 1988.

[44] Schwartz G. Estimating the dimension of a model. *Ann. Statist.*, 6:461–464, 1978.

[45] Shi X. Model selection test for moment inequality models. *J. Econometrics*, 187:1–17, 2015a.

[46] Shi X. A nondegenerate Vuong test. *Quant. Econs.*, 6(1):85–121, 2015b.

[47] Sin C.Y. White H. Information criteria for selecting possibly misspecified parametric models. *J. Econometrics*, 71(1-2):207–225, 1996.

[48] Struthers C. A. Kalbfleisch J. D. Misspecified proportional hazard models. *Biometrika*, 73:363–369, 1986.

[49] Stute W. The central limit theorem under random censorship. *Ann. Statist.*, 23:422–439, 1995.

[50] Stute W. Wang J. L. A strong law under random censorship. *Ann. Statist.*, 21:1591–1607, 1993.

[51] Sun Y. Weak convergence of the generalized parametric empirical processes and goodness-of-fit tests for parametric models. *Comm. Statist. and Theory Methods*, 26(10):2393–2413, 1997.

[52] Van der Vaart A. W. *Asymptotic statistics*. Cambridge University Press, Cambridge, 1998.

[53] Van Keilegom I. Veraverbeke N. Uniform strong convergence results for the conditional Kaplan-Meier estimator and its quantiles. *Commun. Statist. Theory Meth.*, 25:2251–2265, 1996.

[54] Van Keilegom I. Veraverbeke N. Weak convergence of the bootstrapped conditional Kaplan-Meier process and its quantile process. *Commun. Statist. Theory Meth.*, 26:853–869, 1997a.

[55] Van Keilegom I. Veraverbeke N. Estimation and bootstrap with censored data in fixed design nonparametric regression. *Ann. Inst. Statist. Math.*, 49:467–491, 1997b.

[56] Vuong Q.H. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57:307–333, 1989.

[57] White H. Maximum likelihood estimation of misspecified models. *Econometrica*, 50:1–25, 1982.

[58] Wooldridge J. A unified approach to robust, regression-based specification tests. *Econom. Theory*, 6(1):17–23, 1990.

[59] Yang S. A central limit theorem for functionals of the Kaplan-Meier estimator. *Statist. Probab. Lett.*, 21:337–345, 1994.

[60] Zheng J.X. A consistent test of conditional parametric distributions. *Economet. Theor.*, 16:667–691, 2000.

[61] Zheng X. Testing parametric conditional distributions using the nonparametric smoothing method. *Metrika*, 75:455–469, 2012.