Testing under a Modified Cox Regression Model

Dissertation

zur Erlangung des Doktorgrades
der Fakultät für Mathematik, Informatik
und Naturwissenschaften
der Universität Hamburg

vorgelegt
im Department Mathematik
von

Michael Brendel
aus Garmisch-Partenkirchen

Hamburg
2006

Als Dissertation angenommen vom Department
Mathematik der Universität Hamburg

auf Grund der Gutachten von Prof. Dr. Georg Neuhaus
und Prof. Dr. Arnold Janssen

Hamburg, den 1. November 2006
Prof. Dr. Hans Joachim Oberle
Leiter des Departments Mathematik

To my wife Cornelia
for her infinite patience.

Contents

Preface

The determination of the influence of covariates on survival times is a common issue in biomedical research. In this dissertation tests for various hypotheses are rigorously developed on the basis of Cox-type models. The investigated models are obtained by modifying the frequently applied Cox Regression Model (CRM) [13, 14]. For this purpose the basic concept of the rank tests with estimated scores provided by Behnen and Neuhaus [7, 8] is combined with the CRM, and methods derived from LAN and counting process theory are employed.

The results of this dissertation are practical and applied, even thought the structure of the monograph is theoretically oriented. In order to facilitate the access to the presented approach its organisation is outlined in the following paragraphs.

In the first chapter an introduction to the objectives of the dissertation is given. Furthermore, the Modified Cox Regression Model (MCRM) is motivated by the two-sample problem with randomly right censored data, since this well understood problem can be used to link the ideas of Behnen and Neuhaus [7, 8] with the CRM. Moreover, localized, q-dimensional parametric sub-models of the MCRM, which form the basis for the further statistical analysis, are introduced. These models incorporate the crucial aspects of the CRM and the models considered by Behnen and Neuhaus [7, 8]. In contrast to these authors who state their models using $\mathbb{L}_2$-differentiable distribution families, the models in this dissertation are specified by hazard rates. From an application-oriented point of view this approach is preferable, since the resulting counting process models can be easily interpreted and are very comprehensible. However, this direction unfortunately holds several methodological difficulties. Additionally, some of the vast amount of literature on the CRM and its generalizations, as well as the relevant literature on rank test theory, is discussed.

One aim of this dissertation is the development of a comprehensive, asymptotic theory of localized, q-dimensional parametric sub-models of the MCRM. In order to achieve this objective, important theorems and concepts required in the following sections are arranged and discussed. In Section 2.1, the concept of weak convergence on Polish spaces is sketched. Among other things, Rebolledo's Central Limit Theorem and Lenglart's Inequality are stated. Jacod's Formula for the Density Process and the above-named results are the foundation for the proof of a general result on asymptotic normality for counting process models, see Section 2.2. This general result, which can be regarded as a counting process analogue to the Second Le Cam Lemma, is applied to sequences of localized, q-dimensional parametric sub-models of the MCRM in Section 2.3.

The MCRM is an semi-parametric model, *i.e.* the interesting parameter is finite-dimensional and further parameters that are regarded as nuisance are infinite-dimensional. One of these nuisance parameters is the baseline hazard which is an element of an infinite-dimensional function space. Localized, q-dimensional parametric sub-models of the MCRM are obtained by – among others things – restricting the baseline hazard to some at-most q-dimensional sub-space of the before mentioned infinite-dimensional function space. Specifying this sub-space is a problem, since there are no reasons why certain sub-spaces are preferable to others. A well-known way out of this dilemma is the study of hardest parametric sub-models, cf. *e.g.* Neuhaus [60] or Andersen *et al.* [4]. However, in Literature there still remain questions concerning the construction and the definition of hardest parametric sub-models. In Section 3.1 statistical considerations are made to shape and provide such a definition. In the following section of Chapter 3 the properties of sequences of hardest parametric sub-models are investigated. First of all, their existence is established.

The further analysis of sequences of hardest parametric sub-models gives that an important sequence of statistics, cf. Section 4.1 and Section 4.2, is asymptotically equivalent to a sequence of statistics that can be independently chosen of the underlying sequence of localized, parametric sub-models. This sequence of statistics plays a significant role as can be seen in Section 4.3. Additionally, it is proved that this sequence of statistics converges in distribution to some normal distribution that only depends on the interesting parameter and a matrix that can be consistently estimated.

In Section 4.1 and Section 4.2 multivariate one-sided and linear testing problems are examined under fairly general conditions. The models treated in these two sections contain both interesting and nuisance parameters. Moreover, it is assumed that sequences of the models in question are asymptotically normal, *i.e.* they converge weakly to some Gauss Shift Experiment. Based on the likelihood ratio test statistic of the limit Gauss Shift Experiment a test statistic for finite sample-sizes is derived. Finally, it is shown that the resulting sequence of tests keeps asymptotically the level, is even asymptotically unbiased and admissible.

The findings of the two previous sections are applied to sequences of hardest parametric sub-models of the MCRM in Section 4.3. Using the results of Section 3.2, it is shown that the resulting sequence of tests is independent of the special choice of the sequence of localized, parametric sub-models. Therefore, it can be regarded as a sequence of tests for the MCRM. At the end of this chapter, see Section 4.4, it is proven that the tests received in Section 4.3 are projective-type tests. This theorem helps to establish a connection to well known results and provides a descriptive illustration of the effectiveness of the constructed tests.

Chapter 5 is devoted to examples and applications. The existence of the MCRM and localized, q-dimensional parametric sub-models of the MCRM is discussed with elementary methods in Section 5.1. In particular, filtered probability spaces that satisfy the assumptions required to prove the Theorems of Section 2.2 are constructed. In Section 5.2 the assumptions stated in Section 2.3 and Chapter 3, which are exactly the assumptions used in Section 4.3, are examined in detail.

In Section 5.3, the applicability of the MCRM is eventually demonstrated. The two-sample problem with and without concomitant covariates is one major example. Among the further examples are tests for trend and k-sample tests. Finally, model check problems are briefly discussed.

A permutation method to determine the critical values is introduced in Chapter 6. Under additional premises the stated permutation tests keep the level on a subset of the hypothesis even for finite sample sizes. The basic concept is presented in Section 6.1. In the following section, it is proven that the sequence of tests derived in Section 4.3 and the corresponding sequence of permutation tests is asymptotically equivalent. Again, Rebolledo's Central Limit Theorem is a major tool in the proof. The assumptions of Section 6.1 and Section 6.2 are discussed for the important case of time-independent covariates in Section 6.3.

In the previous chapters some proofs were omitted for diverse reason. These proofs can be found in Appendix A. Mainly technical propositions applied in the preceding chapters are collected in Appendix B.

Hamburg, in November 2006 Michael Brendel

Acknowledgement

First, I want to thank my advisor Georg Neuhaus for his constant support and the valuable discussions which led to this comprehensive dissertation. I benefited a lot from his excellent lectures and seminars, where he always managed to present interesting and exciting mathematics and its applications.

A special thank goes to Arnold Janssen who initiated this research during a conference at Oberwolfach in September 2003. I also want to thank him for the possibility to present my work at the Heinrich-Heine-Universität in Düsseldorf and the fruitful exchange.

Moreover, I am indebted to Alexander Smith for proof-reading. His comments on my use of the English language improved the presentation of the dissertation notably.

Last, but not least, I thank my family for offering me the possibility to study and for the encouragement to pursue my interests. In particular I am grateful to my wife Cornelia for her backing and her confidence in me. She always listened to my complaints and frustrations.

M.B.

List of Abbreviations

cf. .... confer

CRM .... Cox Regression Model

e.g. .... *exempli gratia* (for example)

et al. .... *et alii* (and others)

i.e. .... *id est* (that is)

i.i.d. .... independent, identically distributed

MCRM .... Modified Cox Regression Model

SHPSM .... Sequence of Hardest Parametric Sub-Models

WLLN .... Weak Law of Large Numbers

List of Symbols

List of General Symbols

$\mathbb{N}$ .... natural numbers, postive integers

$\mathbb{Z}$ .... integers

$\mathbb{R}$ .... real numbers

$\mathbb{R}_+$ .... $[0, \infty)$

$\mathbb{Q}$ .... rational numbers

$\mathbb{Q}_+$ .... non-negative rational numbers

$\mathbb{R}^{m \times n}$ .... space of all real $(m \times n)$ matrices

$\mathrm{Per}(1, \ldots, n)$ .... set of all permutation of the numbers $1, \ldots, n$

exp .... exponential function

log .... natural logarithm

∇g .... Jacobian matrix of g

$\nabla^2 g$ .... Hessian matrix of g

$\frac{\partial g}{\partial \eta}$ .... partial derivative of g with respect to η

F^{-1} ..... inverse or pseudo-inverse of the cumulative distribution function F

$o(\cdot)$, $O(\cdot)$ .... Landau symbols

$\delta_{i,j}$ .... Kronecker symbol

$<\cdot, \cdot>$ .... inner product

$\|\cdot\|$ .... norm on a vector space

$\|\cdot\|_\infty$ .... sup-norm on $\mathbb{R}^k$

$(\mathcal{V}, <\cdot, \cdot>)$ .... Hilbert space

$\begin{pmatrix} \mathscr{A} & \mathscr{B} \end{pmatrix}$ .... matrix consisting of the sub-matrices $\mathscr{A}$ and $\mathscr{B}$

$\beta^{(u)}$ .... u-th component of the vector β

$\mathscr{A}^{(u,v)}$ .... entry in the u-th row and v-th column of the matrix $\mathscr{A}$

$\mathscr{A}^{-}$ .... generalized inverse of $\mathscr{A}$

$\mathscr{E}_k$ .... $(k \times k)$ unity matrix

$\operatorname{diag}(a_1, \ldots, a_k)$ .... $(k \times k)$ diagonal matrix

$^{\mathrm{T}}$ .... transposed of a vector or matrix

$\ker(A)$ .... kernel of the linear mapping $\mathscr{A}$

$\ker(\mathscr{A})^{\perp}$ .... $\ker(\mathscr{A})^{\perp} = \{x \mid x^{\mathrm{T}} a = 0 \text{ for all } a \in \ker(\mathscr{A})\}$

$\operatorname{Im}(\mathscr{A})$ .... image of the linear mapping $\mathscr{A}$

$\operatorname{rank}(\mathscr{A})$ .... rank of the matrix $\mathscr{A}$

$\operatorname{span}(v_1, \ldots, v_k)$ .... real vector space generated by $v_1, \ldots, v_k$

$\mathbb{B}$ .... Borel σ-algebra on $\mathbb{R}$

$\mathbb{B}_+$ .... Borel σ-algebra on $\mathbb{R}_+$

$\mathbb{B}^p$ .... Borel σ-algebra on $\mathbb{R}^p$

$\mathbb{B}^{m \times n}$ .... Borel σ-algebra on $\mathbb{R}^{m \times n}$

$\mathbb{B}A$ .... $\{B \cap A \mid B \in \mathbb{B}\}$

$\mathcal{P}A$ .... power set of A

$\emptyset$ .... empty set

$A \subset \Omega$ .... A is a subset of Ω

$A \cup B$ .... union of the sets $A, B \subset \Omega$

$A \cap B$ .... intersection of the sets $A, B \subset \Omega$

$A^{\complement}$ .... complement of the set $A \subset \Omega$

$A \backslash B$ .... $A \backslash B = A \cap B^{\complement}$, $A, B \subset \Omega$

$(\Omega, \mathcal{F})$ .... measurable space

$\mathcal{F}$ .... σ-algebra

$\mathcal{F} \vee \mathcal{G}$ .... σ-algebra generated by the σ-algebras $\mathcal{F}$ and $\mathcal{G}$

$\bigvee_{t \in T} \mathcal{F}_t$ .... σ-algebra generated by the σ-algebras $\mathcal{F}_t$, $t \in T$

$\sigma(X)$ .... σ-algebra generated by the mapping X

$\mathbb{F}$ .... filtration

$\mathbb{P}$ .... generic notation for a probability measure

$(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, $(\Omega, \mathcal{F}, \mathbb{F}, \mathfrak{P})$ .... filtered probability space

$\mathfrak{L}(X \mid \mathbb{P})$ .... Law of X under the probability measure $\mathbb{P}$

$\mathbb{E}(X)$ .... expectation of X

$\mathbb{E}[X \mid \mathcal{F}]$, $\mathbb{E}[X \mid Z]$ .... conditional expectation

$\mathbb{P}(X \mid Z)$ .... conditional probability

$\mathrm{Var}(X)$ .... Variance of X

$\mathrm{Cov}(X, Y)$ .... Covariance of X and Y

$\xrightarrow{\mathfrak{D}}_{\mathbb{P}_n}$ .... convergence in distribution

$\longrightarrow_{\mathbb{P}_n}$ .... convergence in probability

$\mathbb{W}$ .... standard Wiener / Brownian motion

$\nu \ll \mu$ .... ν is μ-continuous

$\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ .... μ density of ν

$\frac{\mathrm{d}P}{\mathrm{d}Q}$ .... likelihood ratio of P with respect to Q

$\mathcal{N}(a, \mathscr{S})$ .... normal distribution with mean a and covariance matrix $\mathscr{S}$

u_α .... $(1 - \alpha)$ quantile of a normal distribution with mean 0 and variance 1

χ^2_l .... central χ^2-distribution with l degrees of freedom

$\chi^2_l(\delta)$ .... non-central χ^2-distribution with l degrees of freedom

$\chi^2_{l,\alpha}$ .. $(1-\alpha)$-quantile of a central χ^2-distribution with l degrees of freedom

$\alpha(s)$ .... hazard rate

$\aleph$ .... set of all hazard rates on $\mathbb{R}_+$

$a \wedge b$ .... minimum of a and b

$f \circ g$ .... composition of the function f and g

$\dot{u}$, $\ddot{u}$ .... index function

$\mathbb{1}$ .... indicator function

$\lfloor r \rfloor$ .... integer part of r, $\lfloor r \rfloor = \sup\{z \in \mathbb{Z} \mid z \le r\}$

$|\mathcal{J}|$ .... cardinality of the set $\mathcal{J}$

List of Symbols Introduced in Chapter 1

List of Symbols Introduced in Chapter 2

List of Symbols Introduced in Chapter 3

List of Symbols Introduced in Chapter 4

List of Symbols Introduced in Chapter 5

List of Symbols Introduced in Chapter 6

1 Survival Times and Covariates

This chapter is devoted to the presentation of the fundamental notions and notations used in this dissertation. In Section 1.1 different types of covariates are discussed and the main statistical questions are stated. In Section 1.2 the Cox Regression Model (CRM) is defined and some of the extensive literature on the CRM is summarized and reviewed, before the Modified Cox Regression model is motivated and introduced in Section 1.3.

1.1 Introduction

In a case study, one is often interested in finding out if a new treatment is better than a standard method, or if a new cure has any effect at all. Normally, one forms two groups of subjects. The first group, the so-called control group, receives the standard treatment or no treatment at all, and the members of the second group, the so-called test group, obtain an alternative therapy. Then one observes, how the different subjects respond to their treatments. In medicine this response is typically the time between the start of the treatment and the death of the subject, *i.e.* the survival time. More generally, the response is the time between the start of the treatment and a point in time, when the subject experiences a defined event, some examples being death, a decrease or increase of the subjects constitution or that the drug under consideration stops to be effective. In the following text the time to event is generally called survival time. So, one aims to compare the distributions of the survival times of the two groups in order to discover differences between the standard and the alternative treatment.

From a statistical point of view, this situation is a two-sample problem. There are two groups of observations and one wants to test if the distribution of the survival times in both samples are equal or if the distribution of the test group is stochastically larger. An alternative testing problem would be to test the hypothesis that the distribution of the survival times is the same in both samples against the alternative that the distribution of the survival times is not the same in both samples.

In the previous example the two groups only differed in one characteristic, *i.e.* the type of therapy they received. For a case study, one would try to find subjects that are quite similar, so that a difference in the survival time distribution can be attributed to the difference in treatment.

However, in biology and medicine one often encounters the situation that individuals differ in various characteristics, and one wants to determine the influence of these characteristics on the times to event. These characteristics – the explanatory variables or risk factors – are called covariates. Examples for covariates are physical variables like constitution, blood pressure, age and gender or demographic quantities like education, income or the ethnic group an individual belongs to, and last but not least behaviour variables like smoking and drinking habits, cf. Klein and Moeschberger [43, pp. 243].

Covariates can be classified as time-independent and time-dependent covariates. Typical examples of the first are the kind of therapy, gender or social status. These variables are fixed at the start of the study or do not change during the study. Time-dependent covariates are given by air pollution, constitution, stress, pulse or blood pressure.

Additionally, one has to distinguish between external and internal covariates. External covariates are classified as fixed, defined and ancillary. Time-independent covariates are considered as fixed covariates. Defined covariates are time-dependent but their path is already known at the start of the study, for instance any factor that is controlled by some experimenter and that is not adjusted according to the course of the experiment. In a psychological experiment such a factor could be a stress factor. The last type of external covariates are ancillary covariates which are the realisation of a stochastic process. The marginal distribution of the process is independent of the underlying model for the survival time and the survival time itself, cf. Kalbfleisch and Prentice [41, pp. 123].

Internal covariates are also the realisation of stochastic processes, but the distribution of this process depends on the individual under study, since an internal covariate can only be observed as long as the subject is at risk. Examples for internal covariates are blood pressure or white blood count, cf. Kalbfleisch and Prentice [41, pp. 123]. Other examples are disease complications that cannot be predicted from the history of the process, cf. Andersen *et al.* [4, pp. 169].

As we will see later, our model comprises both internal and external covariates.

In this dissertation it is aimed to develop tests which conclude whether covariates have influence on the survival times. A special case of this undertaking is the well understood two-sample problem, which serves us as a motivation and an illustration. More precisely, the following statistical questions are going to be considered:

- Does a covariate have any influence on the survival time at all?
- Does a large value of a covariate correspond with longer survival times?
- Can differences in the survival times be explained only with some of the covariates?

In the next sections the basic notation is introduced and the Cox Regression Model (CRM), which links the covariates and the survival times is presented. Under the CRM the aforementioned statistical questions can be transformed into parametric testing problems. The first one turns out to be a multivariate one-sided testing problem, the second and third can be transformed into linear hypotheses, cf. Chapter 4 and Chapter 5.

1.2 A First Mathematical Model

In this section we present the basic mathematical notation and a few fundamental notions and terms from counting process theory. Moreover, the CRM is introduced and some of the abounding literature on the CRM is briefly discussed.

First, let us fix some terminology. Suppose that $(\Omega, \mathcal{F})$ is some measurable space and let $\mathbb{P}$ denote some probability measure on $\mathcal{F}$. The survival times $T_1, \ldots, T_n$ are modelled by the measurable mappings $T_i : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}_+, \mathbb{B}_+)$, which can be identified with the stochastic processes $N_T^{(i)} = \big\{ N_T^{(i)}(t) \mid t \in \mathbb{R}_+ \big\}$, where

$$N_T^{(i)}(t) = \mathbb{1}(T_i \leq t), \qquad t \in \mathbb{R}_+, \quad i = 1, \ldots, n.$$

Such a process equals 0 as long as the individual has not experienced the event under consideration – *i.e.* is alive – and jumps to 1, when the event occurs. This jump process is a special example of a so-called counting process. Counting processes are increasing processes with right continuous paths that only take the numbers $\{0, 1, 2, \ldots\}$. The index T indicates that the counting process depends only on the survival time.

A survival time can also be identified with the so-called at-risk process $Y_T^{(i)} = \big\{ Y_T^{(i)}(t) \mid t \in \mathbb{R}_+ \big\}$, $i = 1, \ldots, n$, where

$$Y_T^{(i)}(t) = \mathbb{1}(T_i \geq t), \qquad t \in \mathbb{R}_+.$$

This process equals 1 as long as the individual has not experience the event under consideration, *i.e.* the individual is at risk. The process jumps to 0 after the event in question has occurred. As the random variable $N_T^{(i)}(t)$ tells us, whether the event at the time t with respect to the i-th individual has already occurred or not, the information on that individual up to time t is contained in the σ-algebra

$$\mathcal{F}_T^{(i)}(t) = \sigma\big(N_T^{(i)}(s) \mid s \leq t \big).$$

The family of σ-algebras $\mathbb{F}_T^{(i)} = \big\{ \mathcal{F}_T^{(i)}(t) \mid t \in \mathbb{R}_+ \big\}$, is called filtration.

More precisely, a filtration $\mathbb{F} = \{\mathcal{F}_t \mid t \in \mathbb{R}_+\}$ is a right continuous, increasing family of sub-sigma-algebras of $\mathcal{F}$, *i.e.* it holds that

$$\begin{aligned} \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F} \quad & \text{for all } s \leq t, \\ \mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s \quad & \text{for all } t \in [0,\infty). \end{aligned}$$

It is said to be $\mathbb{P}$-complete, if $\mathcal{F}$ and $\mathcal{F}_0$ contain all subsets of $\mathbb{P}$-null sets.

The tuple $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is called a filtered probability space. In the situation that we consider several probability laws we write $(\Omega, \mathcal{F}, \mathbb{F}, \mathfrak{P})$ for the filtered space, where $\mathfrak{P}$ denotes a family of probability measures on $\mathcal{F}$. The filtered probability space is called complete or satisfying the "usual conditions", if the filtration is $\mathbb{P}$-complete for all $\mathbb{P} \in \mathfrak{P}$. It is always possible to "complete" a filtered space, cf. Jacod and Shiryaev [32, pp. 2].

Additionally, Proposition B.5.1 yields that the family of σ-algebras $\mathbb{F}_T^{(i)} = \{\mathcal{F}_T^{(i)}(t) \mid t \in \mathbb{R}_+\}$ is right continuous.

The process $N_T^{(i)}(t)$, $t \in \mathbb{R}_+$, is a $\mathbb{F}_T^{(i)}$-sub-martingale. By the Doob-Meyer decomposition we know the existence of a predictable, increasing process $A_T^{(i)} = \{A_T^{(i)}(t) \mid t \in \mathbb{R}_+\}$ with $A_T^{(i)}(0) = 0$, such that the process $N_T^{(i)} - A_T^{(i)}$ is a $\mathbb{F}_T^{(i)}$-martingale. The dual predictable projection $A_T^{(i)}$ is unique up to indistinguishability. Note that a process $X = \{X_t \mid t \in \mathbb{R}_+\}$ on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is called predictable, if it is adapted, *i.e.* X_t is $\mathcal{F}_t$-measurable, and if the mapping $(\omega, t) \mapsto X_t(\omega)$ is measurable with respect to the predictable σ-algebra – the σ-algebra generated by the adapted process with left continuous paths. A detailed explanation of the terms introduced in these paragraphs can be found, for example, in the books of Fleming and Harrington [19] or Jacod and Shiryaev [32].

The dual predictable projection $A_T^{(i)}$ is given by

$$A_T^{(i)}(t) = \int_{[0,t]} \frac{Y_T^{(i)}(s)}{\mathbb{P}(T_i \geq s)} \, \mathrm{d}\mathbb{P}^{T_i}(s), \qquad t \in \mathbb{R}_+,$$

and in the case that $\mathbb{P}^{T_i}$ is absolutely continuous by

$$A_T^{(i)}(t) = \int_{[0,t]} Y_T^{(i)}(s)\,\alpha(s)\,\mathrm{d}s, \qquad \alpha(s) = \lim_{h\downarrow 0} \frac{\mathbb{P}\big(s < T_i \leq s+h\big)}{\mathbb{P}\big(T_i > s\big)}, \tag{1.1}$$

cf. Fleming and Harrington [19, Section 1.3] or Andersen *et al.* [4, Example II.4.1]. We note that the distribution of the survival time T_i is reflected by the dual predictable projections of the counting process $N_T^{(i)}$. Later, different underlying probability measures are modelled by stating the dual predictable projections of the counting processes $N_T^{(i)}$, $i = 1, \ldots, n$. But first, let us consider the case that the survival data is right censored.

In clinical studies, it is often not possible to observe the survival time one is interested in. One only registers that the event in question has not happened up to some time t and must have occurred after t. In this case we say that the survival time was right censored. More precisely, one observes an event time and an indicator stating, whether the survival time in question or some censoring time was observed. There are different reasons for right censoring. Among others, subjects drop out of the study because they move away or die and the cause of the death is not related to the investigation, *e.g.* someone dies due to an accident.

This situation can be modelled as follows. Let $C_i : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}_+, \mathbb{B}_+)$, $i = 1, \ldots, n$, denote the censoring times. The survival time T_i is right censored, if $T_i > C_i$, so one merely observes

$$X_i = T_i \wedge C_i \quad \text{and} \quad \Delta_i = \mathbb{1}(T_i \leq C_i),$$

where X_i is the censored survival time and Δ_i the censoring indicator. $\Delta_i = 1$ (0) means that the i-th observation was non-censored (censored).

The random variables X_i and Δ_i can be used to define a new counting process, namely

$$N^{(i)} = \big\{N_t \mid t \in \mathbb{R}_+\big\}, \quad N_t^{(i)} = \Delta_i\, N_T^{(i)}(t) = \int_{[0,t]} Y_C^{(i)}(s)\,\mathrm{d}N_T^{(i)}(s),$$

where $Y_C^{(i)}(t) = \mathbb{1}(C_i \geq t)$, $t \in \mathbb{R}_+$, is the at-risk process of the censoring time. This counting process only jumps to 1, if one observes the event. At first glance, one might think that information is being wasted by only considering the non-censored observations, but we will see that the likelihood in the CRM and also in the MCRM primarily depends on the counting processes $N^{(i)}$. The censored observations are going to be used for estimation.

Setting $\widetilde{N}^{(i)}(t) = (1 - \Delta_i) \cdot \mathbb{1}(X_i \leq t)$, $t \in \mathbb{R}_+$, $i = 1, \ldots, n$, we can define the filtration

$$\mathbb{F}^{(i)} = \big\{\mathcal{F}_t^{(i)} \mid t \in \mathbb{R}_+\big\}, \qquad \mathcal{F}_t^{(i)} = \sigma\big(N^{(i)}(s),\, \widetilde{N}^{(i)}(s) \mid s \leq t\big).$$

Under the additional assumption that T_i and C_i are stochastically independent and that the distribution of T_i is absolutely continuous, the dual predictable projection of N_i is given by

$$A^{(i)}(t) = \int_{[0,t]} Y^{(i)}(s)\, \alpha(s)\, \mathrm{d}s, \quad t \in \mathbb{R}_+,$$

where $Y^{(i)}(t) = Y_T^{(i)}(t) \cdot Y_C^{(i)}(t)$ is the censored at-risk process and α is the hazard rate given in equation (1.1), cf. Fleming and Harrington [19, Theorem 1.3.1]. Note that the dual predictable projection does not formally depend on the distribution of the censoring time, which, in this context, is an infinite dimensional nuisance parameter.

In order to get a "complete" statistical model we still have to integrate the covariates. A straightforward way to model an influence on the distribution of the survival times is by linking the hazard rate of the survival time distribution with the covariates. The proceeding is illustrated by the next example – the two-sample problem – which is a leitmotif of the whole dissertation, since this problem is well understood and serves as the starting point for the modification of the CRM considered in this dissertation. Moreover, we will show that our results always contain the two-sample model and a lot of well known related results as special cases.

1.2.1 Example (Two-sample problem). Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and suppose that $T_i, C_i : \Omega \longrightarrow \mathbb{R}_+$, $i = 1, \dots, n$ and $Z_i : \Omega \longrightarrow \{0, 1\}$, $i = 1, \dots, n$, are measurable mappings denoting the survival and censoring time and the covariate. The covariate Z_i is interpreted as follows, $Z_i = 1$ (0) means that the i-th observation belongs to the first (second) sample. Under this model, it is also assumed that survival times and censoring times T_i and C_i, $i = 1, \dots, n$, are mutually stochastically independent and that we merely observe $X_i = T_i \wedge C_i$, $\Delta_i = \mathbb{1}(T_i \le C_i)$ and the covariate Z_i, $i = 1, \dots, n$.

If F_1 and F_2 denote the cumulative distribution functions

$$F_1(t) = \mathbb{P}\big(T_i \le t \mid Z_i = 1\big) \quad \text{and} \quad F_2(t) = \mathbb{P}\big(T_i \le t \mid Z_i = 0\big), \quad t \ge 0,$$

under the two-sample problem one is interested to test the hypotheses

$$\mathcal{H}: \; F_1 = F_2 \quad \text{versus} \quad \mathcal{K}: \; F_1 \ge F_2, \;\; F_1 \ne F_2,$$

i.e. no difference in survival times versus the distribution of the survival times in the second sample is stochastically larger. Or even more colloquial one wants to test no difference in methods applied to the first and second group versus the method applied to the second group is better. This is a classical non-parametric testing problem.

In the next step we intend to transform the testing problem into a parametric one, because there exists a lot of approved methods to develop reasonable testing procedures for parametric testing problems. Therefore, let us assume that $\mathfrak{P} = \{P_{\beta,\alpha} \mid \beta \in \mathbb{R}, \, \alpha \in \aleph\}$ is some family of probability distributions on $\mathcal{F}$ and $\mathbb{P} \in \mathfrak{P}$, where $\aleph$ denotes the set of all hazard rates on $\mathbb{R}_+$. Analog to the previous consideration we can identify the censored survival times X_i, $i = 1, \dots, n$, with the counting processes $N^{(i)}$, $i = 1, \dots, n$. Moreover, a suitable filtration is given by

$$\mathbb{F} = \{\mathcal{F}_t \mid t \in \mathbb{R}_+\}, \qquad \mathcal{F}_t = \bigvee\nolimits_{i=1,\dots,n} \mathcal{F}_t^{(i)}.$$

Assuming that the dual predictable projection of $N^{(i)}$ under $P_{\beta,\alpha} \in \mathfrak{P}$ is given by

$$A^{(i)}_{\beta,\alpha}(t) = \int_{[0,t]} Y^{(i)}(s)\, \exp(\beta\, Z_i)\, \alpha(s)\, \mathrm{d}s, \qquad t \in \mathbb{R}_+,$$

the testing problem $\mathcal{H}$ versus $\mathcal{K}$ transforms to

$$\mathcal{H}: \ \beta = 0 \quad \text{versus} \quad \mathcal{K}: \ \beta > 0,$$

where we use the fact that

$$\mathbb{P}\big(T_i > t \mid Z_i = z\big) = \exp\left(- \int_{[0,t]} \exp(\beta\, z)\, \alpha(s)\, \mathrm{d}s \right),$$

cf. Fleming and Harrington [19, Theorem 1.3.1]. Implicitly, it is also assumed that the distribution of the covariates does not depend on the underlying probability distribution. Moreover, the baseline hazard α is a nuisance parameter.

The previous example is a special case of the CRM and will be discussed later in greater detail. Note that the above procedure is a standard method to transform a non-parametric statistical question into a parametric testing problem.

1.2.2 Definition (Cox Regression Model). Consider the filtered space $(\Omega, \mathcal{F}, \mathbb{F}, \mathfrak{P})$, where $\mathfrak{P} = \big\{P_{\beta,\alpha} \mid \beta \in \mathbb{R}^p,\ \alpha \in \aleph\big\}$. The observations are given by the tuples (X_i, Δ_i, Z_i), $i = 1, \ldots, n$, where $X_i : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}_+, \mathbb{B}_+)$ denotes a censored survival time and $\Delta_i : (\Omega, \mathcal{F}) \longrightarrow \big(\{0,1\}, \mathcal{P}\{0,1\}\big)$ the corresponding censoring indicator. $\mathcal{P}\{0,1\}$ represents the power set of $\{0,1\}$. $Z_i = \big\{Z_i(t) \mid t \in \mathbb{R}_+\big\}$, $Z_i(t) : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}^p, \mathbb{B}^p)$, is a predictable, and therefore $\mathbb{F}$-adapted, stochastic process. The counting processes $N^{(i)} = \big\{N^{(i)}(t) \mid t \in \mathbb{R}_+\big\}$, $N^{(i)}(t) = \Delta_i \cdot \mathbb{1}(X_i \le t)$, $i = 1, \ldots, n$, are also supposed to be adapted to $\mathbb{F}$. Under the Cox Regression Model (CRM) it is assumed that the dual predictable projection of $N^{(i)}$ under $P_{\beta,\alpha} \in \mathfrak{P}$ is given by $A^{(i)}_{\beta,\alpha} = \big\{A^{(i)}_{\beta,\alpha}(t) \mid t \in \mathbb{R}_+\big\}$, where

$$A^{(i)}_{\beta,\alpha}(t) = \int_{[0,t]} Y^{(i)}(s)\, \exp\big(\beta^{\mathrm{T}} Z_i(s)\big)\, \alpha(s)\, \mathrm{d}s,$$

$Y^{(i)}(s) = \mathbb{1}(X_i \ge s)$, $i = 1, \ldots, n$.

Clearly, Example 1.2.1 is a special case of the CRM model introduced by D. R. Cox [13] in 1972. The analysis of the model – the estimation of the parameter β – by the so-called marginal likelihood caused a lot of attention and discussion. Therefore, Cox [14] formalised the idea of analysis by introducing the partial likelihood. It was shown that maximum likelihood estimates and tests derived from the partial likelihood have the usual large sample properties. Tsiatis [69] proved strong consistency and asymptotic normality of the estimates in the CRM. He also suggested estimates for the underlying baseline hazard and the survivor function and investigated their asymptotic properties. For a modern treatment of the CRM with martingale methods, see Andersen and Gill [5] or Andersen *et al.* [4, Chapter VII]. In this dissertation we also rely on this martingale approach introduced by Aalen [1]. A different method for estimating β on the basis of the method of local likelihoods was proposed by Crowley and Gentlemen [21].

In application the main interest is estimating β and testing hypothesis on β using either the Wald, likelihood ratio or score test, where the estimates and the statistics are derived from Cox partial likelihood cf. Klein and Moeschberger [43, chapter 8]. The concept of the partial likelihood has been a subject of discussion since its introduction. Although it is not a full likelihood, methods based on the partial likelihoods share many properties of methods based on likelihoods. For more detailed information on partial likelihoods consult *e.g.* Wong [73], Jacod [31], Slud [66], Greenwood and Wefelmeyer [25].

The advantage of the CRM model is its simplicity and easy manageability. Its drawback is the assumption that the influence of the covariates is constant in time. In particular, this means that the hazard rates of time-independent covariates are proportional. This drawback has been the starting point for many generalizations of the CRM, and in this dissertation it is also attempted to overcome this assumption, see also Example 1.3.1. A possible way out is stratification, cf. Klein and Moeschberger [43, Section 9.3]. Another approach is to allow β to vary in time. Murphy and Sen [56] assumed β to be some deterministic function and developed a sieve estimation procedure. The method of sieves is also used by Murphy [55] in order to derive a test for the hypothesis of proportional hazard. Verweij and van Houwelingen [70] assume the coefficient β to be a function on a discrete time domain and proposed some estimation procedure using penalized likelihoods. Sargent [62] allows the coefficients to vary in time and bases his method on a dynamic linear model. The model is fitted to the data using Markov Chain Monte Carlo Methods. Models using time-dependent coefficients were also investigated by Martinussen, Scheike and Skovgaard [52]. They use a kernel smoother for their estimation procedure.

The assumptions of the Cox Regression Model are often violated in practise, therefore goodness-of-fit methods for the CRM were developed, cf. Andersen *et al.* [4, Section VII.3]. A similar idea as in our approach is used by Lin [50] to construct a goodness-of-fit procedure. Lin used weighted score functions instead of the normal score functions derived from the partial likelihood – a proceeding that is motivated by the commonly used log-rank tests. Kauermann and Berger [42] apply a related strategy and use the local partial score to construct a goodness-of-fit procedure. The idea of introducing weights can also be found in Grambsch and Therneau [23]. A very general non-linear regression model was considered by McKeague and Utikal [54]. Under this model, they derive a test for independence of survival time and covariate and give as an example of a goodness-of-fit test for the proportional hazard model.

A parametric generalization of the CRM was investigated by Lin and Ying [49], they assume that the dual predictable projection of $N^{(i)}$ is given by

$$A^{(i)}_{\beta,\alpha}(t) = \int_{[0,t]} Y^{(i)}(s) \Big(g\big(\beta_1^{\mathrm{T}} Z_{1,i}(s)\big) + h\big(\beta_2^{\mathrm{T}} Z_{2,i}(s)\big)\, \alpha(s) \Big) \,\mathrm{d}s, \quad t \in \mathbb{R}_+,$$

$\beta = (\beta_1^{\mathrm{T}}, \beta_2^{\mathrm{T}})^{\mathrm{T}}$, $Z_i = (Z_{1,i}^{\mathrm{T}}, Z_{2,i}^{\mathrm{T}})^{\mathrm{T}}$, where g, and h are known link functions. A consistent estimator for β is derived and the weak convergence of an Aalen-Breslow-type estimator for $\int_{[0,t]} \alpha(s)\,\mathrm{d}s$, $t \in \mathbb{R}_+$, is also proved. Moreover, they present some adaptive estimators that achieve the semi-parametric information bounds.

A lot of research was also done concerning non-parametric extensions of the

CRM. Dabrowska [15] considered the very general model

$$A^{(i)}_{\beta,\alpha}(t) = \int_{[0,t]} Y^{(i)}(s)\, \exp\bigl(\beta^{\mathrm{T}} Z_{1,i}(s)\bigr)\, \alpha\bigl(s, Z_{2,i}(s)\bigr)\, \mathrm{d}s, \quad t \in \mathbb{R}_+,$$

$Z_i = (Z^{\mathrm{T}}_{1,i}, Z^{\mathrm{T}}_{2,i})^{\mathrm{T}}$. In that model a kernel smoothing technique is used for the estimation of β and the function α. LeBlanc and Crowley [46] consider a model, where the dual predictable projection of $N^{(i)}$ is given by

$$A^{(i)}_{\eta}(t) = \int_{[0,t]} Y^{(i)}(s)\, \exp\bigl(\eta(Z_i)\bigr)\, \alpha(s)\, \mathrm{d}s, \qquad t \in \mathbb{R}_+$$

and η is some spline. They demonstrate some adaptive technique for the estimation of η. Under the partly linear additive Cox model of Huang [29], one supposes that the dual predictable projection is defined by

$$A^{(i)}_{\beta,\phi_1,\ldots,\phi_{p_2},\alpha}(t) = \int_{[0,t]} Y^{(i)}(s)\, \exp\Bigl(\beta^{\mathrm{T}} Z_{1,i} + \sum_{j=1}^{p_2} \phi_j\bigl(Z^{(j)}_{2,i}\bigr)\Bigr)\, \alpha(s)\, \mathrm{d}s, \quad t \in \mathbb{R}_+,$$

where ϕ_j, $j = 1, \ldots, p_2$, are some smooth functions that are estimated with the help of splines. The rate of convergence is considered and it is shown that the estimator of β attains the semi-parametric information bound. Heller [27] investigated the more general model

$$A^{(i)}_{\beta,g,\alpha}(t) = \int_{[0,t]} Y^{(i)}(s)\, \exp\bigl(\beta^{\mathrm{T}} Z_{1,i} + g(Z_{2,i})\bigr)\, \alpha(s)\, \mathrm{d}s, \quad t \in \mathbb{R}_+,$$

where g is some unknown smooth real-valued function. The interesting parameter β is estimated by maximization of a profile partial likelihood, profiling out g using a kernel function.

In a series of papers the models

$$A^{(i)}_{\beta,g}(t) = \int_{[0,t]} Y^{(i)}(s)\, \exp\bigl(\beta^{\mathrm{T}} Z_{1,i}(s) + g^{\mathrm{T}}(s) Z_{2,i}(s)\bigr)\, \mathrm{d}s, \quad t \in \mathbb{R}_+,$$

cf. Martinussen *et al.* [52],

$$A^{(i)}_{\beta,g,\alpha}(t) = \int_{[0,t]} Y^{(i)}(s)\, \Bigl(\exp\bigl(\beta^{\mathrm{T}} Z_{1,i}(s)\bigr)\, \alpha(s) + g^{\mathrm{T}}(s) Z_{2,i}(s)\Bigr)\, \mathrm{d}s, \quad t \in \mathbb{R}_+,$$

cf. Scheike and Martinussen [51], and

$$A^{(i)}_{\beta,\alpha}(t) = \int_{[0,t]} Y^{(i)}(s)\, \exp\bigl(\beta^{\mathrm{T}} Z_{1,i}(s)\bigr)\bigl(Z^{\mathrm{T}}_{2,i}(s)\alpha(s)\bigr)\,\mathrm{d}s, \quad t \in \mathbb{R}_+,$$

cf. Scheike and Zhang [64], were considered. The authors obtain efficient estimation procedures depending on kernel smoothers for these non-parametric extensions of the CRM. Using the model of Martinussen *et al.*, Scheike and Martinussen [63] proposed tests for checking, whether or not a covariate effect varies in time. Kraus [45] developed goodness-of-fit tests for the additive-multiplicative intensity model introduced by Scheike and Zhang [64] using a stratified martingale residual process.

It is seen that many generalizations of the CRM aim to extend the model with some non-parametric component, so that at least some of the covariates effects can vary in time. This approach to overcome the assumption of proportional hazards has become quite popular in recent years. For these fairly general models kernel smoothers or estimators based on splines are used to detect a possible influence of the covariates on the survival times. A potential problem with methods based on kernel smoothers can be that relatively large sample sizes are often needed, if one wants to rely on asymptotic results. However, in survival analysis sample sizes are quite often comparatively small. Thus, it might be worth considering a different method to treat the CRM, if one intends to investigate dependencies between survival time and covariates. In this dissertation we want to use an approach of rank test theory for extending the CRM. Therefore, some remarks on the literature on the testing of right censored life time data need to be made.

Aalen [1] introduced counting processes and martingale methods to survival analysis. These methods were popularized by Gill [22] who investigated the two-sample problem and weighted log-rank statistics in great generality and detail in his PhD thesis. Martingale methods were also used by Jones and Crowley, [39] and [40], to consider the asymptotic properties of a general class of non-parametric tests for survival analysis. Using a generalized version of the test statistics of Crowley and Jones, Lin and Kosorok [48] consider function-indexed tests to receive testing procedures that are sensitive for a wider range of alternatives. This approach, that came to the author's attention just before finishing this dissertation, uses empirical process theory to derive limit theorems. Even though there are some connection to work presented in this dissertation, our approach is an approach in its own right. For more information on testing in survival analysis see Andersen and Borgan [2], Andersen *et al.* [3] and [4] as well as Jones and Crowley [40].

A different approach extending the classical rank test theory of Hájek and Šidák [26] to censored data was considered by Neuhaus [57] and Janssen [33]. They use local asymptotic normal approximations (LAN theory) to construct (asymptotically) distribution free tests for right censored data under the two-sample model. Additionally, these tests are asymptotically optimal under certain contiguous alternatives. Janssen [34] also investigated optimal k-sample tests for randomly censored data.

Since rank tests are optimal only in one direction of contiguous alternatives Behnen and Neuhaus [7] proposed rank tests with estimated scores that are distribution free and sensitive to a broader range of alternatives. They also apply their ideas to right censored survival data, see Behnen and Neuhaus [8]. Mayer [53] generalized their proceeding to weighted log-rank tests under the two-sample problem deriving asymptotically admissible tests. In this dissertation, it is intended to take up the idea of Behnen and Neuhaus in order to modify the CRM and to develop tests for an influence of covariates on the distribution of the survival times, see Section 1.3.

Combinations of k-sample tests and the CRM were considered by Shen and Fleming [65], who proposed a weighted mean survival test statistic for the two-sample problem that also considers additional, concomitant covariates, and by Heller and Venkatraman [28], who consider the k-sample problem with covariate adjustment in a extended CRM. They later use a kernel smoother to derive non-parametric test statistics. The test statistics derived in this dissertation also include k-sample tests under consideration of additional covariates as special cases.

Summarizing, one can say that the CRM model has become very popular in practical applications and there exists an enormous amount of literature on the CRM and possible generalizations. Survival analysis in general is a field of intensive research. Therefore, it is impossible to review and summarize the literature on that subject. The previous account of the literature is supposed to help to be sort the approach of treating the CRM suggested in this dissertation. In the next Section the MCRM is introduced and discussed in detail.

1.3 The Modified Cox Regression Model

In the previous Section a brief account of the literature on the CRM and some extensions of the CRM was given. Under the CRM the influence of a certain value of the covariate on the baseline hazard is constant in time. For time-independent covariates this means that the conditioned cumulative hazard functions given covariate are proportional. This property of the CRM has been regarded as one of its major drawbacks and has been the starting point for many extensions of the CRM, see Section 1.2. In this Section, the two-sample problem will be the starting point of our modification of the CRM.

1.3.1 Example (Continuation of Example 1.2.1). Let us consider the two-sample problem under the CRM and explicitly assume that (T_i, C_i, Z_i), $i = 1, \ldots, n$ are stochastically independent and that $\mathbb{F} = \big(\mathcal{F}_t \mid t \in \mathbb{R}_+\big)$, $\mathcal{F}_t = \sigma\big(N^{(i)}(s), \widetilde{N}^{(i)}(s) \mid i = 1, \ldots, n,\ s \leq t\big)$. The conditional hazard rate of the survival time T_i given $Z_i = z$ is given by

$$\begin{aligned}\lambda_{\beta,\alpha}(t \mid z) &= \lim_{h \downarrow 0} \frac{1}{h} \frac{P_{\beta,\alpha}\big(t < T_i \leq t + h \mid Z_i = z\big)}{P_{\beta,\alpha}\big(t < T_i \mid Z_i = z\big)} \\ &= \exp\big(\beta \cdot z\big)\, \alpha(t) = \big(1 + \beta \cdot z + o(\beta)\big)\alpha(t)\end{aligned}$$

where $o(\cdot)$ denotes the Landau symbol. Consequently, the probability of an individual dying in the small time interval $(t, t + h]$, if it survives longer than

t, is given by

$$P_{\beta,\alpha}\big(t < T_i \le t + h \mid T_i > t,\ Z_i = z\big) \approx h \cdot (1 + \beta \cdot z)\alpha(t),$$

so we see that the influence of the covariate, *i.e.* the treatment effect, is constant in time. As we have already mentioned, this assumption is often violated in practical applications. A way out can be the introduction of weight functions, that determine the influence of the treatment with respect to time. This could be done by considering the model

$$P_{\beta,\alpha}\big(t < T_i \le t + h \mid T_i > t,\ Z_i = z\big) \approx h \cdot \big(1 + \beta \cdot z \cdot \gamma(t)\big)\alpha(t), \tag{1.2}$$

where γ denotes a weight function. This model is quite handy to interpret. Assume that β is non-negative, a positive value of $\gamma(t)$ increases, a negative value decreases the probability of failure in the time interval $(t, t+h]$, given that the failure occurs after t. Moreover, one can argue that statisticians should have an idea, if they expect short term or long term differences in the survival times, so that they should be able to choose at least approximately a suitable weight function γ.

Such models are well known in the theory of rank statistics, where they play an important role in proving asymptotic optimality of linear rank tests. To simplify matters let us assume that no right censoring is present and that the covariates are deterministic. Let $n_1 = \sum_{k=1}^{n} Z_i$ and $n_2 = n - n_1$ denote the sizes of the first and the second sample and assume that $\frac{n_1}{n} \longrightarrow \nu \in (0,1)$ as $n \to \infty$. Given any two absolute continuous distributions P_1 and P_2 on $(\mathbb{R}, \mathbb{B})$, one can find a parametric family of distributions $\mathfrak{Q} = \big\{Q_\vartheta \mid \vartheta \in [\nu - 1, \nu]\big\}$ that comprises these distributions, more precisely $Q_{1-\nu} = P_1$ and $Q_\nu = P_2$. The probability measure Q_0 is defined by the cumulative distribution function

$$F_0(x) = \nu P_1\big\{(-\infty, x]\big\} + (1-\nu)P_2\big\{(-\infty, x]\big\}, \qquad t \in \mathbb{R}.$$

and arbitrary Q_ϑ is given by Q_0-densities of the form

$$f_\vartheta(x) = 1 + \vartheta \cdot b \circ F_0(x) \qquad [Q_0],$$

where

$$b(u) = \left(\frac{\mathrm{d}P_2}{\mathrm{d}Q_0} - \frac{\mathrm{d}P_1}{\mathrm{d}Q_0}\right) \circ F_0^{-1}(u), \qquad u \in [0,1].$$

F_0^{-1} denotes the pseudo inverse of F_0. The function b is bounded and it holds that $\int_{(0,1)} b(s)\,\mathrm{d}s = 0$. This approach is described in greater detail by Behnen and Neuhaus [7, pp. 18]. The distribution Q_0 is the so-called foot-point of the distribution family and the function b describes the direction of the alternatives.

Furthermore, the distribution family $\mathfrak{Q}$ is $\mathbb{L}_2$-differentiable at $\vartheta = 0$ with $\mathbb{L}_2$-derivative $b \circ F_0$, cf. Witting [71, Definition 1.187, Beispiel 1.200]. Let us define the sequence of rank statistics

$$S_n = \sum_{i=1}^{n} c_{n,i} \cdot b_{n,R_i}, \qquad n \in \mathbb{N},$$

where $c_{n,i}$ are regression coefficients given by

$$c_{n,i} = \sqrt{\frac{n_1 \cdot n_2}{n}} \cdot \left\{ n_2^{-1} \mathbb{1}_{\{0\}}(Z_i) - n_1^{-1} \mathbb{1}_{\{1\}}(Z_i) \right\}, \qquad i = 1, \ldots, n.$$

$R_i = \sum_{j=1}^{n} \mathbb{1}(T_j \le T_i)$ are the ranks of the survival times, and $b_{n,i}$, $i = 1, \ldots, n$ are scores. If one assumes that the scores satisfy the condition

$$\int_{(0,1)} \left(b_{n,\lfloor ns \rfloor} - b(s)\right)^2 \mathrm{d}s \longrightarrow 0, \qquad \text{as } n \to \infty,$$

where $\lfloor ns \rfloor$ denotes the integer part of ns, one can show that the sequence of tests

$$\phi_n = \left\{ \begin{array}{ll} 1, & \\ 0, & \end{array} S_n \begin{array}{c} > \\ \le \end{array} u_\alpha \cdot \left(\int_{(0,1)} b^2(s)\,\mathrm{d}s \right)^{1/2} \right., \qquad n \to \infty,$$

is asymptotically optimal for the testing problem $\vartheta = 0$ versus $\vartheta > 0$ which is a sub-problem of the testing problem $\mathcal{H}$ versus $\mathcal{K}$, if P_2 is stochastically larger than P_1 with respect to the standard stochastic ordering. $u_\alpha = \Phi^{-1}(1-\alpha)$, where Φ denotes the cumulative distribution function of a normal distribution with mean 0 and variance 1. Note that the statistic S_n is distribution-free under the hypothesis of randomness $\mathcal{H}$ and that the optimality of the sequence of tests does not depend on the foot-point of the parametric family, but merely depends on the direction of the alternatives. Thus, one can consider the rank tests ϕ_n as a non-parametric procedure that in certain cases is asymptotically optimal. More comprehensive information can be found in Neuhaus [58, Chapter 17].

In the presence of right censoring the situation becomes more delicate, since the optimal scores also depend on the censoring distribution. Choosing optimal scores for the rank tests proposed by Neuhaus [57] and Janssen [33] can only be done, if one knows the distribution of the censored survival times $X_i = T_i \wedge C_i$ under $\vartheta = 0$, *i.e.* the foot-point, however, their tests are distribution-free under the hypothesis of randomness. Scores for rank tests whose optimality only depends on the direction of the alternatives despite the presence of right censoring were proposed by Brendel [11].

To simplify matters let us stick to the situation of no right censoring. The $\mathbb{L}_2$-differentiability of $\mathfrak{Q}$ implies the existence of a hazard ratio derivate at $\vartheta = 0$. More precisely, it holds that

$$\frac{1}{\vartheta}\left(\frac{\mathrm{d}\Lambda_\vartheta}{\mathrm{d}\Lambda_0} - 1\right) \longrightarrow_{Q_0} R(b \circ F_0), \qquad \text{as } n \to \infty,$$

where $\Lambda_\vartheta(x) = \int_{(-\infty,x]} \big\{Q_\vartheta\big([s,\infty)\big)\big\}^{-1}\,\mathrm{d}Q_\vartheta$ and the hazard ratio derivative is given by

$$R(b \circ F_0)(x) = b \circ F_0(x) - \frac{\int_{[x,\infty)} b \circ F_0(u)\,\mathrm{d}Q_0(u)}{Q_0\big([x,\infty)\big)}, \qquad x \in \mathbb{R}.$$

The operator R establishes an isometry between tangents and hazard ratio derivatives, cf. Janssen [35]. Using this isometry we can also see, how the model introduced in equation (1.2) and the rank test theory approach are linked. Considering that equation, one can identify α as the hazard rate of Q_0 and γ as $R(b \circ F_0)$.

The optimality of the rank tests for certain contiguous alternatives depends on the right choice of b. Even though a statistician might have an idea about the direction of the alternatives, it is impossible to know the right direction of the alternative. Therefore, Neuhaus and Behnen [7] proposed both kernel estimators and projection estimators for the score function b. The resulting rank test proved to be sensitive for different directions of alternatives. Behnen and Neuhaus [8] also applied their approach to right censored data. Mayer [53] considered some primitive projection estimator for the hazard ratio derivatives under the two-sample model deriving projective-type tests generalizing the well-known weighted log-rank tests.

In this dissertation it is aimed to extend the testing procedure proposed by Mayer [53] to arbitrary covariates. And it is shown that new test statistics depend on a multivariate generalization of statistics considered by Jones and Crowley [39]. Basically, our test statistic is the squared norm of some primitive projective-type estimator for the direction of the alternatives. In order to construct the new test statistics, it is assumed that instead of the weight function $\gamma \equiv 1$ determining the direction of the alternatives under the CRM there are a finite number of weight functions specifying possible directions of the alternatives, where we also allow the weight functions to depend on the baseline hazard and the distribution of the censoring times. This very simple idea leads to the modification of the CRM subject of the following text.

1.3.2 Definition (**Modified Cox Regression Model).** Consider the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathfrak{C})$, where $\mathfrak{C} = \{P_{\beta,\alpha} \mid \beta \in \mathbb{R}^r,\ \alpha \in \aleph\}$. The observations are given by the tuples (X_i, Δ_i, Z_i), $i = 1, \dots, n$, where $X_i : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}_+, \mathbb{B}_+)$ denotes a censored survival time and $\Delta_i : (\Omega, \mathcal{F}) \longrightarrow (\{0,1\}, \mathcal{P}\{0,1\})$ the corresponding censoring indicator. $Z_i = \{Z_i(t) \mid t \in \mathbb{R}_+\}$, $Z_i(t) : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}^p, \mathbb{B}^p)$, is the predictable covariate process associated with the i-th observation. This process is obviously $\mathbb{F}$-adapted. Furthermore, it is supposed that the counting processes $N^{(i)} = \{N^{(i)}(t) \mid t \in \mathbb{R}_+\}$, $N^{(i)}(t) = \Delta_i \cdot \mathbb{1}(X_i \leq t)$, $i = 1, \dots, n$, are also adapted to the filtration $\mathbb{F}$.

Assume that $\gamma_\alpha^{(u,v)} : \mathbb{R}_+ \longrightarrow \mathbb{R}$, $v = 1, \dots, r_u$, $u = 1, \dots, p$, are some measurable functions. More precisely, it supposed that $\gamma_\alpha^{(u,v)} = \gamma_0^{(u,v)} \circ H_\alpha$, where

$$\gamma_0^{(u,v)} : [0,1] \longrightarrow \mathbb{R}, \qquad v = 1, \dots, r_u,\ \ u = 1, \dots, p,$$

are measurable functions and H_α is a cumulative distribution function that can depend on the baseline hazard α and the distribution of the censoring times. Set $r = \sum_{u=1}^{p} r_u$. The stochastic process $Z_i \circledcirc \gamma_\alpha = \big\{Z_i \circledcirc \gamma_\alpha(t) \mid t \in \mathbb{R}_+\big\}$, defined by

$$Z_i \circledcirc \gamma_\alpha(t) = \Big(\big(Z_i^{(1)}(t) \cdot \gamma_{\alpha,1}(t)\big)^{\mathrm{T}}, \ldots, \big(Z_i^{(p)}(t) \cdot \gamma_{\alpha,p}(t)\big)^{\mathrm{T}}\Big)^{\mathrm{T}},$$

is called weighted covariate process belonging to the i-th observations, $i = 1, \ldots, n$, where the abbreviations $\gamma_{\alpha,u} = \big(\gamma_\alpha^{(u,1)}, \ldots, \gamma_\alpha^{(u,r_u)}\big)^{\mathrm{T}}$, $u = 1, \ldots, p$, and $\gamma_\alpha = \big(\gamma_{\alpha,1}^{\mathrm{T}}, \ldots, \gamma_{\alpha,p}^{\mathrm{T}}\big)^{\mathrm{T}}$ are used.

Under the modified Cox Regression Model (MCRM) the dual predictable projection of $N^{(i)}$ under $P_{\beta,\alpha} \in \mathfrak{C}$ is given by $A_{\beta,\alpha}^{(i)} = \big\{A_{\beta,\alpha}^{(i)}(t) \mid t \in \mathbb{R}_+\big\}$, where

$$A_{\beta,\alpha}^{(i)}(t) = \int_{[0,t]} Y^{(i)}(s)\, \exp\big(\beta^{\mathrm{T}} Z_i \circledcirc \gamma_\alpha(s)\big)\, \alpha(s)\, \mathrm{d}s, \tag{1.3}$$

$Y^{(i)}(s) = \mathbb{1}(X_i \geq s)$, $i = 1, \ldots, n$, see Remark 1.3.3.b for a different representation of the predictable dual projection. We also use the abbreviations $N = (N^{(1)} \ldots, N^{(n)})$ and $A_{\beta,\alpha} = (A_{\beta,\alpha}^{(1)}, \ldots, A_{\beta,\alpha}^{(n)})$.

1.3.3 Remark. a) Under this model, every component of the covariate vector is multiplied by a vector of weight functions determining the direction of the alternatives. If one choose H_α independently of α then we receive the Cox Regression Model, whereas the covariate processes are given by $Z_i \circledcirc \gamma_\alpha$, $i = 1, \ldots, n$.

b) A perhaps more intuitive reprensentation of the predictable projection of $N^{(i)}$ under $P_{\beta,\alpha}$ is given by

$$A_{\beta,\alpha}^{(i)}(t) = \int_{[0,t]} Y^{(i)}(s)\, \exp\left(\sum_{u=1}^{p} Z_i^{(u)}(s) \sum_{v=1}^{r_u} \overline{\beta}^{(u,v)} \gamma_\alpha^{(u,v)}(s)\right) \alpha(s)\, \mathrm{d}s,$$

$t \in \mathbb{R}_+$, where $\overline{\beta}^{(u,v)} = \beta^{(\sum_{l=1}^{u-1} r_l + v)}$, $v = 1, \ldots, r_u$, $u = 1, \ldots, p$. Nevertheless, the representation of predictable dual projection given in Definition 1.3.2 is used, as it is intended to apply the methods provided by linear algebra.

c) As the weight functions $\gamma_0^{(u,v)}$ are defined on the interval $[0,1]$ a statistician can easily choose functions, such the influence of the covariates is appropriately weighted with respect to time. The "right" transformation of $\gamma_0^{(u,v)}$ onto $\mathbb{R}_+$ is provided by the cumulative distribution function H_α. For example, if one thinks that a covariate mainly effects the survival time, shortly after a subject entered the study, *e.g.* disease complications, then a possible choice of weight functions could be

$$\gamma_0^{(u,v)}(t) = \big(1-t\big)^k, \quad t \in [0,1], \qquad k \geq 1.$$

Furthermore, this procedure also guarantees that the testing procedures to be developed are independent of the underlying time scale of our data *i.e.* if we consider $\big(f(X_i), \Delta_i, Z_i^{(f)}\big)$, $Z_i^{(f)} = \big\{Z_i\big(f(t)\big) \mid t \in \mathbb{R}_+\big\}$, $i = 1, \ldots, n$, instead of (X_i, Δ_i, Z_i), $i = 1, \ldots, n$, where $f : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ is a strictly increasing function, our tests give the same result. In other words, if the weight functions γ_α are chosen independently of the baseline hazard then a transformation of the time scale could lead to a different outcome of the analysis.

d) Results by Janssen [36] suggest that any test keeping the level on the hypothesis can have reasonable power only for a finite number of orthogonal directions of alternatives. So, considering just a finite number of weight functions is no restriction in practice. If a testing procedure is based on kernel estimators the directions of alternatives are implicitly given by the kernel. So, an advantage of the approach discussed in this dissertation is that the directions of the alternatives are directly chosen by the statistician. Additionally, the number of different directions of the alternatives can be adjusted to the sample size.

The MCRM is a semi-parametric statistical model. The interesting parameter β is the parametric part and the infinitely dimensional α together with the distribution of the censoring times form the non-parametric part. For the further development we want to consider sequences of parametric sub-models of the MCRM that we localize. Like in rank test theory it is intended to apply results from LAN theory in order to derive some reasonable testing procedures for the statistical questions introduced in Section 1.1. We will show that there is no harm in considering parametric models, if these models are big enough. This will lead to the notion of sequences of hardest parametric sub-models in Chapter 3.

1.3.4 Definition (Parametric Sub-Model). Assume that our observations are given by the tuples (X_i, Δ_i, Z_i), $i = 1, \ldots, n$, where X_i, Δ_i and $Z_i(t)$, $t \in \mathbb{R}_+$, $i = 1, \ldots, n$, are measurable mapping on $\big(\Omega, \mathcal{F}\big)$. The filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathfrak{P})$, is called q-dimensional parametric sub-model of the MCRM with nuisance direction $\widetilde{\gamma}$ and foot-point α_0, if the following conditions hold:

i) $\widetilde{\gamma} = \big(\widetilde{\gamma}^{(1)}, \ldots, \widetilde{\gamma}^{(q)}\big)$, where $\widetilde{\gamma}^{(u)} : \mathbb{R}_+ \longrightarrow \mathbb{R}$, $u = 1, \ldots, q$, are measurable functions.

ii) $\mathfrak{P} = \big\{P_\xi \mid \xi = (\beta^{\mathrm{T}}, \eta^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{r+q}\big\}$ is a $r + q$-dimensional distribution family.

iii) The dual predictable projection of the counting process N under P_ξ is given by $A_\xi = \big(A_\xi^{(1)}, \ldots, A_\xi^{(n)}\big)$, where

$$A_\xi^{(i)}(t) = \int_{[0,t]} \exp\Big(\beta^{\mathrm{T}} Z_i \circledcirc \gamma_{\alpha_\eta}(s)\Big) Y_i(s)\, \alpha_\eta(s)\, \mathrm{d}s, \quad t \in \mathbb{R}_+, \tag{1.4}$$

and $\alpha_\eta(s) = \exp\big(\eta^{\mathrm{T}}\widetilde{\gamma}(s)\big)\, \alpha_0(s)$, $s \in \mathbb{R}_+$. The hazard rate $\alpha_0 \in \aleph$ is fixed.

If instead of (1.4) it holds that

$$A_\xi^{(i)}(t) = \int_{[0,t]} \exp\Big(\frac{1}{\sqrt{n}} \cdot \beta^{\mathrm{T}} Z_i \circledcirc \gamma(s) + \frac{1}{\sqrt{n}} \cdot \eta^{\mathrm{T}}\widetilde{\gamma}(s)\Big) Y_i(s)\, \alpha_0(s)\, \mathrm{d}s, \quad t \in \mathbb{R}_+,$$

where $\gamma = \gamma_{\alpha_0}$, we call the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathfrak{P})$ a $1/\sqrt{n}$-localized, q-dimensional parametric sub-model of the MCRM.

In the further treatment, we will concentrate on $1/\sqrt{n}$-localized, q-dimensional parametric sub-model of the MCRM.

1.3.5 Remark. a) Note that under a q-dimensional parametric sub-model the nuisance parameter α is restricted to the space $\{\alpha_0 \exp(\eta^{\mathrm{T}}\widetilde{\gamma}) \mid \eta \in \mathbb{R}^q\}$. Later we will see that completely fixing the nuisance parameter α would not be an appropriate approach, since those sub-models do not share the properties of the MCRM. They are too small. A complete fixing of the nuisance parameter is equivalent to knowing the correct baseline hazard α. This is clearly not the case under the MCRM.

b) Note that under the localized, q-dimensional parametric sub-model of the MCRM we fix the weight functions. More precisely, they only depend on the foot-point α_0 and the distribution of the censoring times. This procedure is justified by the localization.

c) The function $(\eta, t) \mapsto \exp(\eta^{\mathrm{T}}\widetilde{\gamma}(t))$ can be replaced by any function g that is two times continuously differentiable with respect to η and that satisfies

$$\frac{\partial g}{\partial \eta^{(i)}}_{\mid \eta = 0, t = t_0} = \widetilde{\gamma}^{(i)}(t_0), \qquad i = 1, \ldots, n.$$

As we want to localize, only the derivatives at $\eta = 0$ are important. However, the specific form for the parametrisation is very convenient for further treatment.

In the next Chapter sequences of localized, parametric sub-models are considered. Asymptotic normality for counting process models is introduced and conditions implying asymptotic normality are discussed.

2 Asymptotic Normality

In this chapter the theoretical framework is established and some important results from martingale theory, that are frequently applied in this dissertation, are stated. Furthermore, asymptotic normality for counting process models is introduced and a general theorem on local asymptotic normality is presented and discussed. Subsequently, this result is applied to sequences of localized, q-dimensional parametric sub-models of the MCRM.

2.1 Important Results and Concepts

As we intend to consider sequences of models, the concept of convergence in distribution plays an important role. Let the tuple (E, d) denote a metric space, where E is some set and d is a metric on E. Furthermore, let $\mathcal{E}$ be the Borel-σ-algebra on (E, d), *i.e.* the smallest σ-algebra that contains all open sets. Assume that P_n, $n \in \mathbb{N}_0$, are probability measures on $\mathcal{E}$. We say that P_n, $n \in \mathbb{N}$, converges weakly to P_0, if

$$\int f(x)\,\mathrm{d}P_n(x) \longrightarrow \int f(x)\,\mathrm{d}P_0(x), \qquad \text{as } n \to \infty\ ,$$

for all real-valued, bounded and continuous functions f. Equivalent definitions of convergence in distribution are summarized in the so-called Portmanteau Theorem, cf. Billingsley [9, Theorem 2.1].

However, in this dissertation we are interested in sequences of random variables. Consider a sequence of probability spaces $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$, $n \in \mathbb{N}$, and assume that

$X_n : \Omega_n \longrightarrow E$, $n \in \mathbb{N}_0$, are some measurable mappings. Consequently,

$$P_n(A) = \mathbb{P}_n\big\{X_n \in A\big\}, \qquad A \in \mathcal{E}, \quad n \in \mathbb{N},$$

are probability measures on $\mathcal{E}$. We say X_n converges in distribution to X_0 on E, in short $X_n \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} X_0$, as $n \to \infty$, if the sequence P_n, $n \in \mathbb{N}$, converges weakly to P_0. Note that the notation $\longrightarrow_{\mathbb{P}_n}$ means convergence in probability, see also Appendix B.4.

Before we can state Rebolledo's Central Limit Theorem for Local Martingales, some more notation has to be introduced. $D\big(\mathbb{R}_+, \mathbb{R}\big)$ denotes the set of all functions $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$ that are right continuous with left hand limits. These functions are called cadlag functions. Additionally, the function space $D\big(\mathbb{R}_+, \mathbb{R}\big)$ can be endowed with a metrizable topology, such that $D\big(\mathbb{R}_+, \mathbb{R}\big)$ is a Polish space. This topology is called Skorokhod topology. More detailed information on this subject can be found in Jacod and Shiryaev [32, Chapter VI]. Analogously, one defines $D\big([0,\tau], \mathbb{R}\big)$, as the set of all cadlag functions $f : [0,\tau] \longrightarrow \mathbb{R}$. The space $D\big([0,\tau], \mathbb{R}\big)$ can be identified with the space $D\big([0,1], \mathbb{R}\big)$, which is also a polish space, if it is endowed with the Skorokhod topology, see Billingsley [9, Chapter 3] for a detailed discussion. Note that $\tau = \infty$ is also possible. One easily shows that if the sequence of processes $\big\{X_n(t) \mid t \in \mathbb{R}_+\big\}$, $n \in \mathbb{N}$, converges in distribution to $\big\{X(t) \mid t \in \mathbb{R}_+\big\}$ on $D\big(\mathbb{R}_+, \mathbb{R}\big)$ then $\big\{X_n(t) \mid t \in [0,\tau]\big\}$ converges in distribution $\big\{X(t) \mid t \in [0,\tau]\big\}$ on $D\big([0,\tau], \mathbb{R}\big)$, $\tau < \infty$. Furthermore, the sequence of stopped processes $\big\{X_n(t \wedge \tau) \mid t \in [0,\infty]\big\}$ converges in distribution to $\big\{X(t \wedge \tau) \mid t \in [0,\infty]\big\}$ on $D\big([0,\infty], \mathbb{R}\big)$, $\tau < \infty$. These implications are applied in the proof of Theorem 2.2.7.

An important result from weak convergence theory, which is applied again and again, is the following replacement result.

2.1.1 Theorem. Let $\big(\Omega_n, \mathcal{F}_n, \mathbb{P}_n\big)$, $n \in \mathbb{N}$, be some sequence of probability spaces and let (E,d) be a polish space. Assume that

$$X_{n,k} : \Omega_n \longrightarrow E, \quad k,n \in \mathbb{N}, \qquad \text{and} \qquad Y_n : \Omega_n \longrightarrow E, \quad n \in \mathbb{N},$$

are $\mathcal{F}_n$–$\mathcal{E}$ measurable, where $\mathcal{E}$ denotes the Borel-σ-algebra on E, and that

$$X_{n,k} \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} X_k, \quad \text{as } n \to \infty, \quad \text{and} \quad X_k \xrightarrow{\mathfrak{D}} X, \quad \text{as } k \to \infty.$$

If for all $\varepsilon > 0$ the condition

$$\lim_{k\to\infty} \lim_{n\to\infty} \mathbb{P}_n\big(d(X_{n,k}, Y_n) \geq \varepsilon\big) = 0$$

holds then it also holds that $Y_n \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} X$, as $n \to \infty$.

Proof. Cf. Billingsley [9, Theorem 4.2] □

Furthermore, it is assumed that all (local) (sub- / super-) martingales have cadlag paths. For any (local) square integrable martingales M, N the processes $\langle M \rangle$ and $\langle M, N \rangle$ denote the dual predictable variation and covariation, *i.e.* $\langle M \rangle$ and $\langle M, N \rangle$ are predictable processes, $\langle M \rangle(0) = \langle M, N \rangle(0) = 0$, and the processes $M^2 - \langle M \rangle$ and $M \cdot N - \langle M, N \rangle$ are (local) martingales. The following result is one standard tool for analysing counting process models.

2.1.2 Theorem (**Rebolledo's Central Limit Theorem**)**.** Consider a sequence of filtered probability spaces $(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathbb{P}_n)$, $n \in \mathbb{N}$, satisfying the usual conditions, *i.e.* the filtration $\mathbb{F}_n$ is increasing, right continuous and $\mathbb{P}_n$-complete. The latter means that $\mathcal{F}_n$ and $\mathcal{F}_{n,t}$, $t \in \mathbb{R}_+$, contain all subsets of $\mathbb{P}_n$ negligible sets. And let $\{M_n(t) \mid t \in \mathbb{R}_+\}$, $n \in \mathbb{N}$, be a sequence of local, square integrable $\mathbb{F}$-martingales, where M_n is defined on $(\Omega_n, \mathcal{F}_n)$. For $\varepsilon > 0$ we define the jump process

$$J^\varepsilon[M_n](t) = \sum_{s \leq t} \Delta M_n(s) \mathbb{1}\big(|\Delta M_n(s)| \geq \varepsilon\big), \quad t \in \mathbb{R}_+,$$

where $\Delta M_n(s) = M_n(s) - M_n(s-)$, $M_n(s-) = \lim_{u \uparrow s} M_n(u)$ and $M_n(0-) = M_n(0)$. Because of the Doob-Meyer decomposition, cf. *e.g.* Jacod and Shiryaev [32, Theorem I.3.18], there exists a predictable, up to an evanescent set unique process $A^\varepsilon[M_n] = \{A^\varepsilon[M_n](t) \mid t \in \mathbb{R}_+\}$, such that the processes

$$M_{n,1}^\varepsilon = J_n^\varepsilon[M_n] - A_n^\varepsilon[M_n] \quad \text{and} \quad M_{n,2}^\varepsilon = M_n - M_{n,1}^\varepsilon$$

are local, square integrable $\mathbb{F}$-martingales. If

$$\langle M_{n,1}^{\varepsilon}\rangle(t) + \sup_{s\le t}\big|\langle M_{n,1}^{\varepsilon}, M_{n,2}^{\varepsilon}\rangle(s)\big| \longrightarrow_{\mathbb{P}_n} 0, \qquad \text{as } n \to \infty,$$

for all $t \geq 0$ and $\varepsilon > 0$, and

$$\langle M_n\rangle(t) \longrightarrow_{\mathbb{P}_n} A(t), \qquad \text{as } n \to \infty,$$

for all $t \geq 0$, where A is a continuous, non-decreasing function with $A(0) = 0$, then

$$M_n \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} \big\{\mathbb{W} \circ A(t) \mid t \in \mathbb{R}_+\big\}, \qquad \text{as } n \to \infty, \quad \text{on } D(\mathbb{R}_+, \mathbb{R}).$$

$\mathbb{W}$ denotes a standard Wiener motion (Brownian motion).

Proof. See Rebolledo [61, Theorem V.1]. □

In Chapter 6 this result is used to investigate the asymptotic properties of permutation tests. The following Corollary, see *e.g.* Fleming and Harrington [19, Theorem 5.1.1] is an easy consequence of Rebolledo's Central Limit Theorem. It will play a crucial role in deriving a criterion for asymptotic normality in counting process models. In the following we always assume that a counting process N defined on a filtered probability space $\big\{\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}\big\}$ is adapted to the filtration, *i.e.* $N(t)$ is $\mathcal{F}_t$-measurable for all $t \in \mathbb{R}_+$.

2.1.3 Corollary. Let $(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathbb{P}_n)$, $n \in \mathbb{N}$, be a sequence of filtered probability spaces satisfying the usual conditions and let $N_n = \big(N_n^{(1)}, \ldots, N_n^{(k_n)}\big)$, $N_n^{(i)} = \big\{N_n^{(i)}(t) \mid t \in \mathbb{R}_+\big\}$, $i = 1, \ldots, k_n$, be a multivariate counting process, *i.e.* none of the $N_n^{(i)}$, $i = 1, \ldots, n$, jump at the same time. The dual predictable projection $A_n = \big(A_n^{(1)}, \ldots, A_n^{(k_n)}\big)$ of N, where $A_n^{(i)} = \big\{A_n^{(i)}(t) \mid t \in \mathbb{R}_+)\big\}$, is assumed to have continuous paths. Furthermore, suppose that $H_n^{(i)}$, $i = 1, \ldots, k_n$, $n \in \mathbb{N}$, are real-valued, predictable, locally bounded processes and that for all $t \in \mathbb{R}_+$ the following conditions hold:

$$\sum_{i=1}^{k_n} \int_{[0,t]} \big(H_n^{(i)}(s)\big)^2 \,\mathrm{d}A_n^{(i)}(s) \longrightarrow_{\mathbb{P}_n} A(t), \quad \text{as } n \to \infty, \tag{2.1}$$

where $A : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous, non-decreasing function with $A(0) = 0$, and

$$\sum_{i=1}^{k_n} \int_{[0,t]} \big(H_n^{(i)}(s)\big)^2 \, \mathbb{1}\big(\big|H_n^{(i)}(s)\big| \geq \varepsilon\big) \, \mathrm{d}A_n^{(i)}(s) \longrightarrow_{\mathbb{P}_n} 0, \quad \text{as } n \to \infty, \tag{2.2}$$

for all $\varepsilon > 0$. Using the abbreviation $M_n^{(i)} = N_n^{(i)} - A_n^{(i)}$ it holds that

$$\left\{ \sum_{i=1}^{k_n} \int_{[0,t]} H_n^{(i)}(s) \, \mathrm{d}M_n^{(i)}(s) \;\middle|\; t \in \mathbb{R}_+ \right\} \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} \big\{\mathbb{W} \circ A(t) \mid t \in \mathbb{R}_+\big\}, \text{ as } n \to \infty,$$

on $D(\mathbb{R}_+, \mathbb{R})$.

Proof. See Appendix A.1. □

A very useful tool which is applied again and again is Lenglart's domination property. It is also essential for proving Rebolledo's Central Limit Theorem.

2.1.4 Definition (Lenglart's Domination Property). Assume that $X = \big\{X(t) \mid t \in \mathbb{R}_+\big\}$ and $Y = \big\{Y(t) \mid t \in \mathbb{R}_+\big\}$ are two processes on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, such that X is optional and Y is a predictable, non-negative, increasing process, where $Y(0) = 0$. If it holds that $\mathbb{E}\big(X(T)\big) \leq \mathbb{E}\big(Y(T)\big)$ for all bounded stopping times T then X is Lenglart-dominated by Y.

2.1.5 Theorem (Lenglart's Inequality). Let X be a cadlag process which is Lenglart-dominated by Y. For any stopping time T and any $\varepsilon, \eta > 0$, it holds that

$$\mathbb{P}\Big(\sup\nolimits_{0 \leq t \leq T} \big|X(t)\big| \geq \varepsilon\Big) \leq \frac{\eta}{\varepsilon} + \mathbb{P}\big(Y(T) \geq \eta\big).$$

Proof. Cf. Jacod and Shiryaev [32, Lemma I.3.30]. □

2.1.6 Corollary. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions. Suppose $N = (N^{(1)}, \ldots, N^{(n)})$ is a multivariate counting process with dual predictable projection $A = (A^{(1)}, \ldots, A^{(n)})$ having continuous paths. Moreover, let $H^{(i)}$, $i = 1, \ldots, n$, be locally bounded and predictable processes. For any stopping time T, such that $\mathbb{P}\{T < \infty\} = 1$, and any $\varepsilon, \eta > 0$, it holds that

$$\mathbb{P}\left(\sup_{0\le t\le T}\left(\sum_{i=1}^{n}\int_{[0,t]} H^{(i)}(s)\,\mathrm{d}M^{(i)}(s)\right)^2 \ge \varepsilon\right) \le \frac{\eta}{\varepsilon} + \mathbb{P}\Big(\sum_{i=1}^{n}\int_{[0,T]} \big(H^{(i)}(s)\big)^2\,\mathrm{d}A^{(i)}(s) \ge \eta\Big)$$

Proof. See Appendix A.2. □

2.1.7 Corollary. a) In the situation of Theorem 2.1.5, it holds that

$$\mathbb{P}\Big(\sup_{0\le t<\infty}\big|X(t)\big| \ge \varepsilon\Big) \le \frac{\eta}{\varepsilon} + \mathbb{P}\big(Y(\infty) \ge \eta\big),$$

where $Y(\infty) = \lim_{t\to\infty} Y(t)$, for all $\eta, \varepsilon > 0$.

b) In the situation of Corollary 2.1.6, it holds that

$$\mathbb{P}\left(\sup_{0\le t<\infty}\left(\sum_{i=1}^{n}\int_{[0,t]} H^{(i)}(s)\,\mathrm{d}M^{(i)}(s)\right)^2 \ge \varepsilon\right) \le \frac{\eta}{\varepsilon} + \mathbb{P}\Big(\sum_{i=1}^{n}\int_{[0,\infty)} \big(H^{(i)}(s)\big)^2\,\mathrm{d}A^{(i)}(s) \ge \eta\Big)$$

for all $\eta, \varepsilon > 0$.

Proof. In both cases, choose the stopping times $T_n \equiv \tau_n \in \mathbb{R}_+$, $n \in \mathbb{N}$, such that $T_n \uparrow \infty$ and apply Theorem 2.1.5 or Corollary 2.1.6 with T_n. Considering Corollary 2.1.7.a this means

$$p_{n,1} = \mathbb{P}\Big(\sup_{0\le t<T_n}\big|X(t)\big| \ge \varepsilon\Big) \le \frac{\eta}{\varepsilon} + \mathbb{P}\big(Y(T_n) \ge \eta\big).$$

Using the Monotone Convergence Theorem yields $p_{n,1} \le \frac{\eta}{\varepsilon} + \mathbb{P}\big(Y(\infty) \ge \eta\big)$. Applying the Monotone Convergence Theorem again establishes the result. Corollary 2.1.7.b is shown completely analogously. □

To show the asymptotic equivalence of certain sequences of random variables, the following Lemma will play a crucial role. It is an immediate consequence of the previous results.

2.1.8 Lemma. Let $(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathbb{P}_n)$, $n \in \mathbb{N}$, be a sequence of filtered probability spaces satisfying the usual conditions and let $I(t)$ denote the interval $[0,t]$, if $t < \infty$, or $[0,\infty)$, if $t = \infty$. The dual predictable projection $A_n = (A_n^{(1)}, \ldots, A_n^{(k_n)})$ of the multivariate counting process $N_n = (N_n^{(1)}, \ldots, N_n^{(k_n)})$ is assumed to have continuous paths. Moreover, let $H_n^{(i)}$, $i = 1, \ldots, k_n$ be locally bounded and predictable processes. The condition

$$\sum_{i=1}^{k_n} \int_{I(\tau)} \big(H_n^{(i)}(s)\big)^2 \,\mathrm{d}A_n^{(i)}(s) \longrightarrow_{\mathbb{P}_n} 0, \qquad \text{as } n \to \infty,$$

implies

$$\sum_{i=1}^{k_n} \int_{I(\tau)} H_n^{(i)}(s) \,\mathrm{d}M_n^{(i)}(s) \longrightarrow_{\mathbb{P}_n} 0 \qquad \text{as } n \to \infty,$$

where $M_n^{(i)} = N_n^{(i)} - A_n^{(i)}$, $i = 1, \ldots, k_n$, $n \in \mathbb{N}$.

Proof. Assume $\varepsilon, \eta > 0$. Applying Corollary 2.1.6 or Corollary 2.1.7 gives the estimate

$$\begin{aligned}
\mathbb{P}_n &\left(\left| \sum_{i=1}^{k_n} \int_{I(\tau)} H_n^{(i)}(s) \,\mathrm{d}M_n^{(i)}(s) \right| \geq \varepsilon \right) \\
&\leq \mathbb{P}_n \left(\sup\nolimits_{t \in I(\tau)} \left(\sum\nolimits_{i=1}^{k_n} \int_{I(t)} H_n^{(i)}(s) \,\mathrm{d}M_n^{(i)}(s) \right)^2 \geq \varepsilon^2 \right) \\
&\leq \frac{\eta}{\varepsilon^2} + \mathbb{P}_n \left(\sum\nolimits_{i=1}^{k_n} \int_{I(\tau)} \big(H_n^{(i)}(s)\big)^2 \,\mathrm{d}A_n^{(i)}(s) \geq \eta \right).
\end{aligned}$$

Consequently, $\limsup_{n\to\infty} \mathbb{P}_n\big(|\sum_{i=1}^{k_n} \int_{[0,\tau]} H_n^{(i)}(s) \,\mathrm{d}M_n^{(i)}(s)| \geq \varepsilon\big) \leq \frac{\eta}{\varepsilon^2}$, $\eta \downarrow 0$ establishes the result. $\square$

2.2 A General Result on Asymptotic Normality

In this Section, asymptotic normality for counting process models is established. First, let us introduce some general premises.

2.2.1 Assumption. Let $\big(\Omega_n, \mathcal{F}_n, \mathbb{F}_n = \{\mathcal{F}_{n,t} \mid t \in \mathbb{R}_+\}, \mathfrak{P}_n\big)$, $n \in \mathbb{N}$, be a sequence of filtered probability spaces and suppose that the following conditions hold.

i) The filtrations satisfy the condition $\mathcal{F}_n = \mathcal{F}_{n,\infty} = \bigvee_{t\geq 0} \mathcal{F}_{n,t}$.

ii) $\mathfrak{P}_n = \big\{P_{n,\xi} \mid \xi \in \mathbb{R}^m\big\}$ is a family of probability measures defined on the σ-algebra $\mathcal{F}_n$. Let $P^{(t)}_{n,\xi}$ denote the restriction of the probability measure $P_{n,\xi}$ to the sigma-algebra $\mathcal{F}_{n,t}$, $t \in [0,\infty]$, *i.e.* $P^{(t)}_{n,\xi}(F) = P_{n,\xi}(F)$, $F \in \mathcal{F}_{n,t}$. It is assumed that $P^{(0)}_{n,\xi} = P^{(0)}_{n,0}$ for all $\xi \in \mathbb{R}^m$.

iii) $\mathcal{F}_n$ and $\mathcal{F}_{n,t}$, $t \in \mathbb{R}_+$ are $P_{n,0}$-complete.

iv) $P_{n,\xi} \ll P_{n,0}$, for all $\xi \in \mathbb{R}^m$.

v) Let $N_n = \big(N^{(1)}_n, \dots N^{(k_n)}_n\big)^{\mathrm{T}}$ and $\widetilde{N}_n = \big(\widetilde{N}^{(1)}_n, \dots \widetilde{N}^{(k_n)}_n\big)^{\mathrm{T}}$ be two counting processes, such that $(N^{\mathrm{T}}, \widetilde{N}^{\mathrm{T}})^{\mathrm{T}}$ is a multivariate counting process.

vi) $N^{(i)}_n + \widetilde{N}^{(i)}_n \leq 1$ $P_{n,0}$-almost surely and $N^{(i)}_n(0) = \widetilde{N}^{(i)}_n(0) = 0$, $i = 1, \dots, k_n$. The latter means that no events occur at time 0.

vii) Let $\mathcal{G}_n$ be some σ-algebra and assume that $\mathcal{F}_{n,t} = \mathcal{G}_n \vee \mathcal{N}_{n,t}$, where $\mathcal{N}_{n,t} = \sigma\big\{N(s),\, \widetilde{N}_n(s) \mid s \leq t\big\}$, $t \in \mathbb{R}_+$. Usually, $\mathcal{G}_n$ contains all subsets of $P_{n,0}$ negligible sets and the information on the covariates.

viii) The dual predictable projection of the counting processes $N^{(i)}_n$ under the probability measure $P_{n,\xi}$ is given by

$$A^{(i)}_{n,\xi}(t) = \int_{[0,t]} Y^{(i)}_n(s)\, \alpha^{(i)}_{n,\xi}(s)\, \mathrm{d}s, \qquad t \in \mathbb{R}_+,$$

where $Y^{(i)}_n = \{Y^{(i)}_n(s) \mid s \in \mathbb{R}_+\}$, $Y^{(i)}_n(s) = 1 - \big(N^{(i)}_n(s-) + \widetilde{N}^{(i)}_n(s-)\big)$ and $\alpha^{(i)}_{n,\xi} Y^{(i)}_n = \big\{\alpha^{(i)}_{n,\xi}(t)\, Y^{(i)}_n(t) \mid t \in \mathbb{R}_+\big\}$, $i = 1, \dots, n$, are some non-negative, predictable processes. Furthermore, suppose that $\alpha^{(i)}_{n,0}(s) = 0$ implies $\alpha^{(i)}_{n,\xi}(s) = 0$ for all $\xi \in \mathbb{R}^m$.

ix) The dual predictable projection of the counting processes $\widetilde{N}^{(i)}_n$ under the probability measure $P_{n,\xi}$ is given by

$$\widetilde{A}^{(i)}_n(t) = \int_{[0,t]} Y^{(i)}_n(s)\, \widetilde{\alpha}^{(i)}_n(s)\, \mathrm{d}s, \qquad t \in \mathbb{R}_+,$$

where $\widetilde{\alpha}_n^{(i)} Y_n^{(i)} = \{\widetilde{\alpha}_n^{(i)}(t) Y_n^{(i)}(t) \mid t \in \mathbb{R}_+\}$, $i = 1, \ldots, n$, are non-negative, predictable processes, and $Y_n^{(i)}$ is defined in viii).

The existence of filtered probability spaces satisfying these assumptions is discussed later in the context of examples, see Section 5.1.

The processes N_n and $\widetilde{N}_n$ are supposed to coincide with the processes defined in Section 1.2 (page 6). $N_n^{(i)}$ is a counting process associated with some survival time. A jump only occurs, if the survival time is observed. Analogously, $\widetilde{N}_n^{(i)}$ is a counting process associated with the corresponding censoring time. The requirement $N_n^{(i)} + \widetilde{N}_n^{(i)} \leq 1$ guarantees that only the survival time or the censoring time is observed.

The conditions imposed on the processes $\alpha_{n,\xi}^{(i)}$ and $\widetilde{\alpha}_n^{(i)}$ are very natural, if one keeps in mind that the paths of these processes are supposed to be hazard rates of some measure on $\mathbb{B}_+$. The assumption that the processes are predictable guarantees that every path is a $\mathbb{B}_+$–$\mathbb{B}_+$ measurable function. The implication stated in Assumption 2.2.1.viii reflects Assumption 2.2.1.iv. Before we give some more remarks on these assumptions, let us introduce a notion of asymptotic normality for counting process models.

2.2.2 Definition (Asymptotic Normality). a) A sequence of filtered probability spaces $(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{P}_n)$, $n \in \mathbb{N}$, is said to be asymptotically normal restricted to the time $t \in (0, \infty]$ with positive semi-definite asymptotic information matrix $\mathscr{I}(t) \in \mathbb{R}^{m \times m}$, if

$$\log \frac{\mathrm{d}P_{n,\xi}^{(t)}}{\mathrm{d}P_{n,0}^{(t)}} - \xi^{\mathrm{T}} S_n(t) + \frac{1}{2}\xi^{\mathrm{T}} \mathscr{I}(t)\xi \longrightarrow_{P_{n,0}} 0, \quad \text{for all } \xi \in \mathbb{R}^m,$$

where $S_n(t)$ is $\mathfrak{F}_{n,t}$-measurable and $S_n(t) \xrightarrow{\mathfrak{D}}_{P_{n,0}} \mathcal{N}(0, \mathscr{I}(t))$ as $n \to \infty$. $S_n(t)$, $n \in \mathbb{N}$, is called central sequence.

b) A sequence of filtered probability spaces is said to be asymptotically normal, if it is asymptotically normal restricted to the time ∞.

2.2.3 Remark. a) Restricting the sequence of filtered probability spaces to time t means for a statistical experiment that we only consider the information up to time t. At this point Definition 2.2.2 pays tribute to the fact that it is often easier to consider counting process models on compact intervals $[0,t]$ rather than on $\mathbb{R}_+$.

b) If X_n is $\mathcal{F}_{n,t}$-measurable for all $n \in \mathbb{N}$, then $X_n \xrightarrow{\mathfrak{D}}_{P_{n,\xi}} X$, as $n \to \infty$, and $X_n \xrightarrow{\mathfrak{D}}_{P_{n,\xi}^{(t)}} X$, as $n \to \infty$ are equivalent. If additionally Y_n is also $\mathcal{F}_{n,t}$-measurable for all $n \in \mathbb{N}$, then $X_n - Y_n \longrightarrow_{P_{n,\xi}} 0$, as $n \to \infty$, and $X_n - Y_n \longrightarrow_{P_{n,\xi}^{(t)}} 0$, as $n \to \infty$ are also equivalent.

c) Asymptotic normality yields that

$$\log \mathrm{d}P_{n,\xi}^{(t)}/\mathrm{d}P_{n,0}^{(t)} \xrightarrow{\mathfrak{D}}_{P_{n,0}^{(t)}} \mathcal{N}\big(-\frac{1}{2}\xi^{\mathrm{T}}\mathscr{I}(t)\xi, \xi^{\mathrm{T}}\mathscr{I}(t)\xi\big) \qquad \text{for all } \xi \in \mathbb{R}^m.$$

Moreover, the First Le Cam Lemma, cf. Witting and Müller-Funk [72, Korollar 6.124], gives that the sequences of probability measures $\{P_{n,\xi}^{(t)} \mid n \in \mathbb{N}\}$ and $\{P_{n,0}^{(t)} \mid n \in \mathbb{N}\}$ are mutual contiguous. In particular convergence in $P_{n,0}^{(t)}$-probability implies convergence in $P_{n,\xi}^{(t)}$-probability.

One can think of the quantities depending on t in the previous definition, especially $\mathrm{d}P_{n,\xi}^{(t)}/\mathrm{d}P_{n,0}^{(t)}$ and $S_n(t)$, $t \in \mathbb{R}_+$, as stochastic processes. This leads to the introduction of the density process.

2.2.4 Definition (Density Process). The process

$$\Upsilon_{n,\xi} = \left\{ \Upsilon_{n,\xi}(t) = \frac{\mathrm{d}P_{n,\xi}^{(t)}}{\mathrm{d}P_{n,0}^{(t)}} \;\Big|\; t \in \mathbb{R}_+ \right\}$$

is called the density process of $P_{n,\xi}$ with respect to $P_{n,0}$.

Assumption 2.2.1.i guarantees that the probability measure $P_{n,\xi}$ can be approximated by $P_{n,\xi}^{(t)}$ for sufficiently large t. More precisely, this condition enables us to approximate the likelihood $\mathrm{d}P_{n,\xi}/\mathrm{d}P_{n,0}$ by the density process, see Proposition 2.2.5. This is an essential step in proving asymptotic normality.

In the case of external covariates, Assumption 2.2.1.ii means that the distribution of the covariates does not depend on the underlying probability distribution. This property reflects the notion of external covariates presented in Section 1.1. In this context Assumption 2.2.1.vii seems to be very restrictive, but it is necessary to ensure that the density process can be represented by the processes N_n and $\widetilde{N}_n$. However, this type of filtration is only needed for the proofs. The following Proposition is a specialisation of a well-known result by Jacod, cf. Jacod [30] or Jacod and Shiryaev [32, Theorem III.5.19].

2.2.5 Proposition (Jacod's Formula for the Density Process). Let $I(t)$ be the interval $[0, t]$, if $t \in \mathbb{R}_+$, or $[0, \infty)$, if $t = \infty$. Under Assumption 2.2.1, it holds that

$$\Upsilon_{n,\xi}(t) = \frac{\exp\Big(-\sum_{i=1}^{k_n} \int_{I(t)} Y_n^{(i)}(s)\, \alpha_{n,\xi}^{(i)}(s)\, \mathrm{d}s\Big)}{\exp\Big(-\sum_{i=1}^{k_n} \int_{I(t)} Y_n^{(i)}(s)\, \alpha_{n,0}^{(i)}(s)\, \mathrm{d}s\Big)} \times \exp\left(\sum_{i=1}^{k_n} \int_{I(t)} \log\left(\frac{\alpha_{n,\xi}^{(i)}(s)}{\alpha_{n,0}^{(i)}(s)}\right) \mathrm{d}N_n^{(i)}(s)\right)$$

for the density process $\Upsilon_{n,\xi}$

Proof. First of all we note that the filtration $\mathbb{F}_n$ is increasing and right continuous, see Proposition B.5.1. Let us start with finite t. We use the notions presented in Jacod and Shiryaev [32, Chapter II, Chapter III], especially we intend to apply Theorem III.5.19. Assumption 2.2.1.iv gives that $P_{n,\xi}^{(t)} \ll P_{n,0}^{(t)}$ for all $t \in \mathbb{R}_+$.

The first and the second characteristic of the process $X_n = \big(N_n^{\mathrm{T}}, \widetilde{N}_n^{\mathrm{T}}\big)^{\mathrm{T}}$ can be chosen identically as 0, cf. Jacod and Shiryaev [32, Definition II.2.6, Formula II.3.22]. Setting $e_{n,j} = (\delta_{1,j}, \ldots, \delta_{2k_n,j})$, where $\delta_{i,j}$ denotes the Kronecker symbol ($\delta_{i,j} = 1$ (0), if and only if $i = j$ $(i \neq j)$), we can represent the multivariate point process associated to the jumps of X_n as follows:

$$\mu_n\big([0, t],\, B\big) = \sum_{i=1}^{k_n} \mathbb{1}_B\big(e_{n,i}\big) \cdot N_n^{(i)}(t) + \sum_{i=1}^{k_n} \mathbb{1}_B\big(e_{n,k_n+i}\big) \cdot \widetilde{N}_n^{(i)}(t), \quad B \in \mathbb{B}^{2k_n},$$

where we use the fact that X_n is a multivariate counting process. See Jacod and Shiryaev [32, Proposition II.1.16 and Definition III.1.23] for detailed information.

Clearly, μ_n is a integer-valued random measure, cf. Jacod and Shiryaev [32, Definition II.1.13, Proposition II.1.16]. Using the fact that $N_n^{(i)} - A_{n,\xi}^{(i)}$ and $\widetilde{N}_n^{(i)} - \widetilde{A}_n^{(i)}$ are local martingales, we get that the third characteristic of X_n under $P_{n,\xi}$ is given by

$$\nu_{n,\xi}\big([0,t], B\big) = \sum_{i=1}^{k_n} \mathbb{1}_B\big(e_{n,i}\big) \cdot A_{n,\xi}^{(i)}(t) + \sum_{i=1}^{k_n} \mathbb{1}_B\big(e_{n,k_n+i}\big) \cdot \widetilde{A}_n^{(i)}(t), \quad B \in \mathbb{B}^{2k_n},$$

where we use Theorem I.3.18 and Theorem II.1.8.ii of Jacod and Shiryaev [32]. By Girsanov's Theorem, cf. Jacod and Shiryaev [32, Theorem III.3.24], we know that there exists a function $U_{n,\xi}$, such that $U_{n,\xi}(\omega,t,x) \cdot \nu_{n,0}(\omega, \mathrm{d}s, \mathrm{d}x) = \nu_{n,\xi}(\omega, \mathrm{d}s, \mathrm{d}x)$. This function is obviously given by

$$U_{n,\xi}(\omega,t,x) = \sum_{i=1}^{k_n} \frac{\alpha_{n,\xi}^{(i)}(\omega,t)}{\alpha_{n,0}^{(i)}(\omega,t)} \cdot \mathbb{1}(x = e_{n,i}) + \mathbb{1}\big(x \in \{e_{n,j} \mid j = k_n+1, \dots, 2k_n\}\big).$$

We see that $\nu_{n,0}\big(\omega, \{t\}, \mathbb{R}^{2k_n}\big) = 0$. Evaluating formula III.5.7 in Jacod and Shiryaev [32] gives that

$$\begin{aligned} H_{n,\xi}(\omega,\, t) &= \int_{[0,t]\times\mathbb{R}^{2k_n}} \big(1 - \sqrt{U_{n,\xi}(\omega,s,x)}\big)\, \nu_{n,\xi}(\omega,\, \mathrm{d}s,\, \mathrm{d}x) \\ &= \sum_{i=1}^{n} \int_{[0,t]} \Big(1 - \sqrt{\alpha_{n,\xi}^{(i)}(\omega,s)/\alpha_{n,0}^{(i)}(\omega,s)}\Big)^2\, Y_n^{(i)}(\omega,\, s)\; \alpha_{n,0}^{(i)}(\omega,s)\, \mathrm{d}s. \end{aligned}$$

The process $H_{n,\xi}$ does not jump to infinity, cf. Jacod and Shiryaev [32, Definition III.5.8], and therefore the condition (ii) of Corollary III.5.22 in Jacod and Shiryaev [32] is satisfied.

In particular, all local martingales have representation property relative to μ_n, cf. Jacod and Shiryaev [32, Condition III.1.25, Equation III.4.35]. It holds that $\Upsilon_{n,\xi}(0) = 1$ because of Assumption 2.2.1.ii. Evaluation of formula III.5.23 in Jacod and Shiryaev [32] gives the assertion for finite t.

In the next step the result is extended to $t = \infty$. Because of Assumption 2.2.1.iv, Jacod and Shiryaev [32, Proposition III.3.5] give that the density process is a uniformly integrable martingale, *i.e.* there exists a integrable random variable $\Upsilon_{n,\xi}(\infty)$, such that $\mathbb{E}\big[\Upsilon_{n,\xi}(\infty) \mid \mathcal{F}_{n,t}\big] = \Upsilon_{n,\xi}(t)$. Moreover, it holds that $\Upsilon_{n,\xi}(\infty) = \lim_{t\to\infty} \Upsilon_{n,\xi}(t)$ $P_{n,0}$-almost surely, cf. Jacod and Shiryaev [32, Theorem I.1.42]. $\mathcal{F}_n = \mathcal{F}_{n,\infty}$ is generated by $\bigcup_{t\geq 0} \mathcal{F}_{n,t}$. Note that for all $A, B \in \bigcup_{t\geq 0} \mathcal{F}_{n,t}$ we have $A \cap B \in \bigcup_{t\geq 0} \mathcal{F}_{n,t}$. For $B \in F_t$ it holds that

$$
\begin{aligned}
\int_B \Upsilon_{n,\xi}(\infty)\,\mathrm{d}P_{n,0} &= \int_B \Upsilon_{n,\xi}(t)\,\mathrm{d}P_{n,0} \\
&= \int_B \Upsilon_{n,\xi}(t)\,\mathrm{d}P_{n,0}^{(t)} = \int_B 1\,\mathrm{d}P_{n,\xi}^{(t)} = \int_B 1\,\mathrm{d}P_{n,\xi}.
\end{aligned}
$$

Thus, $\Upsilon_{n,\xi}(\infty)$ is a version of the density of $P_{n,\xi}$ with respect to $P_{n,0}$. $\square$

2.2.6 Remark. In particular, this result means that the censoring mechanism is non-informative in the sense of Andersen *et al.* [4, Definition III.2.2]. But now let us concentrate on the main result of this section.

2.2.7 Theorem. Let $I(t)$ be the interval $[0, t]$, if $t < \infty$, or the interval $[0, \infty)$, if $t = \infty$. For fixed $\tau \in (0, \infty]$ let $\mathscr{J} : I(\tau) \to \mathbb{R}^{m\times m}$ be a continuous function, such that $\mathscr{J}(0) = 0$ and that the matrix $\mathscr{J}(t)$ is positive semi-definite and symmetric for all $t \in I(\tau)$. Moreover, assume that the mappings $\mathscr{J}_\xi : I(\tau) \longrightarrow \mathbb{R}_+$, $t \mapsto \xi^{\mathrm{T}} \mathscr{J}(t)\xi$, $\xi \in \mathbb{R}^m$, are non-decreasing and continuous. In the case of $\tau = \infty$, it is supposed that $\mathscr{J}(\infty) = \lim_{t\to\infty} \mathscr{J}(t) < \infty$. Clearly, $\mathscr{J}(\infty)$ is positive semi-definite. Abbreviating $\lambda_{n,0}^{(i)} = Y_n^{(i)}\,\alpha_{n,0}^{(i)}$ and

$$
f_{n,\xi}^{(i)}(s) = \left(\sqrt{\frac{\alpha_{n,\xi}^{(i)}(s)}{\alpha_{n,0}^{(i)}(s)}} - 1\right) Y_n^{(i)}(s), \quad i = 1, \ldots, k_n,
$$

we suppose that the processes $\big\{\alpha_{n,\xi}^{(i)}(t)/\alpha_{n,0}^{(i)}(t)\, Y_n^{(i)}(t) \mid t \in \mathbb{R}_+\big\}$, $i = 1, \ldots, k_n$, are predictable and locally bounded and that Assumption 2.2.1 and the following conditions hold:

$$\sum_{i=1}^{k_n} \int_{I(t)} f_{n,\xi}^{(i)}(s)\, f_{n,\xi'}^{(i)}(s)\, \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s \longrightarrow_{P_{n,0}} \frac{1}{4}\xi^{\mathrm{T}} \mathscr{I}(t)\xi' \tag{2.3}$$

and

$$\sum_{i=1}^{k_n} \int_{I(t)} \big(f_{n,\xi}^{(i)}(s)\big)^2 \mathbb{1}\big(\big|f_{n,\xi}^{(i)}(s)\big| > \varepsilon\big)\, \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s \longrightarrow_{P_{n,0}} 0 \tag{2.4}$$

for all $t \in [0,\tau]$, $\xi, \xi' \in \mathbb{R}^m$ and $\varepsilon > 0$. Then we have

$$\sup_{t\in I(\tau)} \left| \log \Upsilon_{n,\xi}(t) + \frac{1}{2}\mathscr{I}_\xi(t) - 2\sum_{i=1}^{k_n} \int_{I(t)} f_{n,\xi}^{(i)}(s)\,\mathrm{d}M_{n,0}^{(i)} \right| \longrightarrow_{P_{n,0}} 0, \tag{2.5}$$

where $M_{n,0}^{(i)}(s) = N_n^{(i)}(s) - A_{n,0}^{(i)}(s)$, $i = 1,\dots,k_n$, and

$$\left\{ 2\sum_{i=1}^{k_n} \int_{I(t\wedge\tau)} f_{n,\xi}^{(i)}(s)\,\mathrm{d}M_{n,0}^{(i)} \;\middle|\; t \in \mathbb{R}_+ \right\} \xrightarrow{\mathfrak{D}}_{P_{n,0}} \big\{\mathbb{W} \circ \mathscr{I}_\xi(t\wedge\tau) \mid t \in \mathbb{R}_+\big\} \tag{2.6}$$

on $D(\mathbb{R}_+, \mathbb{R})$ implying

$$\big\{\log\Upsilon_{n,\xi}(t\wedge\tau)\,|\, t \in \mathbb{R}_+\big\} \xrightarrow{\mathfrak{D}}_{P_{n,0}} \left\{\mathbb{W} \circ \mathscr{I}_\xi(t\wedge\tau) - \frac{1}{2}\mathscr{I}_\xi(t\wedge\tau) \;\middle|\; t \in \mathbb{R}_+\right\} \tag{2.7}$$

on $D(\mathbb{R}_+, \mathbb{R})$. Especially, it holds that

$$\left| \log \Upsilon_{n,\xi}(\tau) + \frac{1}{2}\mathscr{I}_\xi(\tau) - 2\sum_{i=1}^{k_n} \int_{I(\tau)} f_{n,\xi}^{(i)}(s)\,\mathrm{d}M_{n,0}^{(i)} \right| \longrightarrow_{P_{n,0}} 0. \tag{2.8}$$

and

$$2\sum_{i=1}^{k_n} \int_{I(\tau)} f_{n,\xi}^{(i)}(s)\,\mathrm{d}M_{n,0}^{(i)} \xrightarrow{\mathfrak{D}}_{P_{n,0}} \mathcal{N}\big(0, \mathscr{I}_\xi(\tau)\big). \tag{2.9}$$

As a consequence of the previous implications we get that

$$\big(\log \Upsilon_{n,\xi_1}(\tau), \dots, \log \Upsilon_{n,\xi_r}(\tau)\big)^{\mathrm{T}} \xrightarrow{\mathfrak{D}}_{P_{n,0}} \mathcal{N}\Big(-\frac{1}{2}\varsigma(\tau), \mathscr{S}(\tau)\Big), \tag{2.10}$$

where

$$\mathscr{S}(\tau) = \big(\xi_1,\dots,\xi_r\big)^{\mathrm{T}} \mathscr{I}(\tau)\big(\xi_1,\dots,\xi_r\big), \qquad \xi_1,\dots,\xi_r \in \mathbb{R}^m,$$

$\varsigma(\tau) = \big(\mathscr{S}^{(1,1)}(\tau), \dots, \mathscr{S}^{(r,r)}(\tau)\big)^{\mathrm{T}}$ and $r \in \mathbb{N}$.

Proof. See Appendix A.3. □

2.2.8 Remark. a) Theorem 2.2.7 is a version of a Theorem stated by Andersen *et al.* [4, Theorem VIII.2.1], (2.3) and (2.4) imply (2.10). Unfortunately, they do not present a proof of this theorem, but they accredit this Theorem to Jacod and Shiryaev [32, Theorem X.1.12]. The Theorem of Jacod and Shiryaev gives a sufficient and necessary condition for (2.7). However, the conditions stated there do not coincide with (2.3) and (2.4). Therefore, a proof of Theorem 2.2.7 is presented in the Appendix.

b) Equation (2.10) could also be used as a definition of the asymptotic normality, since equation (2.10) implies that the sequence of filtered probability spaces is asymptotically normal restricted to time τ in the sense of Definition 2.2.2. This can be see as follows. Set $e_u = \big(\delta_{u,v},\, v = 1, \ldots, m\big)^{\mathrm{T}}$, $u = 1, \ldots, m$, where $\delta_{u,v}$ denotes the Kronecker symbol. If one sets

$$S_n(\tau) = \Big(\log \Upsilon_{n,e_1}(\tau), \ldots, \log \Upsilon_{n,e_m}(\tau)\Big)^{\mathrm{T}} + \frac{1}{2}\varsigma(\tau), \qquad n \in \mathbb{N},$$

one readily shows using equation (2.10) that $S_n(\tau)$ is a central sequence, where one uses that convergence in distribution to some constant implies convergence in probability to that constant cf. Witting and Müller-Funk [72, Hilfssatz 5.82].

c) If (2.3) only holds for $\xi = \xi'$, then Theorem 2.2.7 stays valid except that assertion (2.10) only holds for $r = 1$.

2.3 Asymptotic Normality for Parametric Sub-Models

In the next Section the main result of Section 2.2 is applied to sequences of parametric sub-models providing the first key result for the further analysis of the MCRM. Initially, let us introduce some general premises.

2.3.1 Assumption. i) The measurable, non-negative function $\alpha_0 \in \aleph$ is called baseline hazard. Moreover, we set

$$\tau_0 = \sup\Big\{ t \in \mathbb{R}_+ \;\Big|\; \int_{[0,t]} \alpha_0(s)\,\mathrm{d}s < \infty \Big\}.$$

ii) Suppose that $\alpha_{n,\xi}^{(i)}(s) = \exp\Big(\frac{1}{\sqrt{n}}\,\xi^{\mathrm{T}}\Psi_{n,i}(s)\Big)\,\alpha_0(s)$, $i = 1, \ldots, n$, and that the processes $\big\{\Psi_{n,i}(s)\,Y_n^{(i)}(s) \mid s \in \mathbb{R}_+\big\}$, $i = 1, \ldots, n$ are predictable and locally bounded.

iii) $\sup_{t\in[0,t]}\big|\frac{1}{n}\sum_{i=1}^{n} \Psi_{n,i}^{(u)}(t)\,\Psi_{n,i}^{(v)}(t)\,Y_n^{(i)}(t) - \Psi^{(u,v)}(t)\big| \longrightarrow_{P_{n,0}} 0$, as $n \to \infty$, where $\Psi^{(u,v)}$ is some measurable function that is bounded on every interval of the form $[0,t]$, $u, v = 1, \ldots, m$, for all $t < \tau_0$.

iv) $\sup_{i\in\{1,\ldots,n\}, s\in[0,t]}\Big\{\frac{1}{\sqrt{n}} \max_{1\le u\le m}\big|\Psi_{n,i}^{(u)}(s)\,Y_n^{(i)}(s)\big|\Big\} \longrightarrow_{P_{n,0}} 0$, as $n \to \infty$, for all $t < \tau_0$.

2.3.2 Remark. In Assumption 2.3.1 we consider the supremum of uncountably many random variables. This supremum is not necessarily a random variable, *i.e.* a measurable mapping. In the following we always assume that all suprema of uncountable many random variables are measurable. This question will be discussed in Chapter 5. See Proposition B.5.5.a for some condition guaranteeing the measurability.

2.3.3 Theorem. Let $I(t)$ denote the interval $[0,t]$, if $t < \infty$, or $[0,\infty)$, if $t = \infty$. Under Assumption 2.2.1 and Assumption 2.3.1 the sequence of filtered spaces $\big(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{P}_n\big)$, $n \in \mathbb{N}$, is asymptotically normal restricted to time τ with asymptotic information matrix $\mathscr{I}(\tau) = \big(\mathscr{I}^{(u,v)}(\tau)\big)$, where

$$\mathscr{I}^{(u,v)}(\tau) = \int_{I(\tau)} \Psi^{(u,v)}(s)\,\alpha_0(s)\,\mathrm{d}s, \quad u, v = 1, \ldots, m,$$

for all $\tau < \tau_0$. Moreover, it holds that

$$\log \Upsilon_{n,\xi}(\tau) + \frac{1}{2}\xi^{\mathrm{T}}\mathscr{I}(\tau)\xi - 2\sum_{i=1}^{n}\int_{I(\tau)} R_n^{(i)}(s,\xi) - 1\,\mathrm{d}M_{n,0}^{(i)}(s) \longrightarrow_{P_{n,0}} 0, \quad (2.11)$$

where $R_n^{(i)}(s,\xi) = \exp\big(\frac{1}{2\sqrt{n}}\xi^{\mathrm{T}}\Psi_{n,i}(s)\,Y_n^{(i)}(s)\big)$ and $M_{n,0}^{(i)} = N_n^{(i)} - A_{n,0}^{(i)}$, $i = 1,\dots,n$. A central sequence is given by

$$S_n(\tau) = \frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n}\int_{I(\tau)} \Psi_{n,i}^{(u)}(s)\,\mathrm{d}M_{n,0}^{(i)}(s),\ u = 1,\dots,m\right)^{\mathrm{T}}. \tag{2.12}$$

If additionally $\mathscr{I}(\tau_0) = \lim_{t\to\tau_0}\mathscr{I}(t)$ exists and the conditions

$$\lim_{t\to\tau_0}\limsup_{n\to\infty} P_{n,0}\left(\left|\sum_{i=1}^{n}\int_{(t,\tau_0)} V_n^{(i)}(s,\xi,\xi')\ \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s\right| \geq \varepsilon\right) = 0, \tag{2.13}$$

where $V_n^{(i)}(s,\xi,\xi') = \big(R_n^{(i)}(s,\xi)-1\big)\big(R_n^{(i)}(s,\xi')-1\big)$, $\lambda_{n,0}^{(i)}(s) = Y_n^{(i)}(s)\,\alpha_0(s)$, and

$$\lim_{t\to\tau_0}\limsup_{n\to\infty} P_{n,0}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\int_{(t,\tau_0)} \big(\xi^{\mathrm{T}}\Psi_{n,i}(s)\big)^2\ \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s\right| \geq \varepsilon\right) = 0 \tag{2.14}$$

hold for all $\varepsilon > 0$, then the above assertions also hold for τ_0.

Proof. Note that the processes $\big\{\alpha_{n,\xi}^{(i)}(t)\mid t\in\mathbb{R}_+\big\}$, $i = 1,\dots,n$, $\xi\in\mathbb{R}^m$, are predictable and that the processes

$$\left\{\frac{\alpha_{n,\xi}^{(i)}(t)}{\alpha_{n,0}^{(i)}(t)}\,Y_n^{(i)}(t)\ \middle|\ t\in\mathbb{R}_+\right\}, \qquad i = 1,\dots,n,\ \xi\in\mathbb{R}^m,$$

are predictable and locally bounded, cf. Proposition B.5.3, where we set $0/0 = 0$. We want to apply Theorem 2.2.7. First, it is verified that the condition (2.3), *i.e.*

$$\sum_{i=1}^{n}\int_{I(t)} V_n^{(i)}(s,\xi,\xi')\ \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s \longrightarrow_{P_{n,0}} \frac{1}{4}\xi^{\mathrm{T}}\mathscr{I}(t)\xi'$$

holds for all $\xi,\xi'\in\mathbb{R}^m$ and $t\in[0,\tau]$.

Proof of (2.3). For the functions $R_n^{(i)}$ we want to compute a Taylor-expansion in ξ with fixed s. The Jacobian and the Hessian matrix of $R_n^{(i)}$ are given by $\nabla_\xi R_n^{(i)}(s,\xi) = \frac{1}{2\sqrt{n}}R_n^{(i)}(s,\xi)\cdot\Psi_{n,i}(s)\,Y_n^{(i)}(s)$ and $\nabla_\xi^2 R_n^{(i)}(s,\xi) = \frac{1}{4n}R_n^{(i)}(s,\xi)\cdot\Psi_{n,i}(s)\,\Psi_{n,i}^{\mathrm{T}}(s)\,Y_n^{(i)}(s)$. Therefore, a Taylor-expansion at the point $\xi = 0$ gives

$$R_n^{(i)}(s,\xi)-1 = \left(\frac{\xi^{\mathrm{T}}\Psi_{n,i}(s)}{2\sqrt{n}} + \frac{1}{8n}\,R_n^{(i)}\big(s,\xi_{n,i}(s)\big)\big(\xi^{\mathrm{T}}\Psi_{n,i}(s)\big)^2\right)Y_n^{(i)}(s), \tag{2.15}$$

$\big|\xi_{n,i}(s)\big| \in [0,\xi] = \bigtimes_{i=1}^m \big[0,|\xi^{(i)}|\big]$, and that the right hand side of (2.3) is equal to $T_n^{(1)}(t) + T_n^{(2)}(t) + T_n^{(3)}(t) + T_n^{(4)}(t)$, where

$$\begin{aligned}
T_n^{(1)}(t) &= \int_{I(t)} \frac{1}{4n} \sum_{i=1}^n \xi^{\mathrm{T}} \Psi_{n,i}(s) \cdot \Psi_{n,i}^{\mathrm{T}}(s)\, \xi'\, \lambda_{n,0}^{(i)}(s)\, \mathrm{d}s, \\
T_n^{(2)}(t) &= \int_{I(t)} \frac{1}{16n^{3/2}} \sum_{i=1}^n \xi'^{\mathrm{T}} \Psi_{n,i}(s)\, Q_{n,\xi}^{(i)}(s)\, \lambda_{n,0}^{(i)}(s)\, \mathrm{d}s, \\
T_n^{(3)}(t) &= \int_{I(t)} \frac{1}{16n^{3/2}} \sum_{i=1}^n \xi^{\mathrm{T}} \Psi_{n,i}(s)\, Q_{n,\xi'}^{(i)}(s)\, \lambda_{n,0}^{(i)}(s)\, \mathrm{d}s, \\
T_n^{(4)}(t) &= \int_{I(t)} \frac{1}{64n^2} \sum_{i=1}^n Q_{n,\xi}^{(i)}(s)\, Q_{n,\xi'}^{(i)}(s)\, \lambda_{n,0}^{(i)}(s)\, \mathrm{d}s
\end{aligned}$$

using the abbreviation

$$Q_{n,\xi}^{(i)}(s) = R_n^{(i)}\big(s, \xi_{n,i}(s)\big)\, \big(\xi^{\mathrm{T}} \Psi_{n,i}(s)\big)^2\, Y_n^{(i)}(s).$$

In the following we are going to use the abbreviations $\mathscr{D}_n(t) = \big(\mathscr{D}_n^{(u,v)}(t)\big)$ $u,v = 1,\dots,m$, where

$$\mathscr{D}_n^{(u,v)}(t) = \sup_{s\in[0,t]} \left| \frac{1}{n} \sum_{i=1}^n \Psi_{n,i}^{(u)}(s)\, \Psi_{n,i}^{(v)}(s)\, Y_n^{(i)}(s) - \Psi^{(u,v)}(s) \right|$$

and $|\xi| = \big(|\xi^{(1)}|,\dots,|\xi^{(m)}|\big)^{\mathrm{T}}$. Note that Assumption 2.3.1.iii implies

$$\zeta^{\mathrm{T}} \mathscr{D}_n(t) \zeta' \longrightarrow_{P_{n,0}} 0 \quad \text{for all } \zeta,\zeta' \in \mathbb{R}^m. \tag{2.16}$$

Because of the estimate

$$\begin{aligned}
&\left| T_n^{(1)}(t) - \frac{1}{4} \xi^{\mathrm{T}} \mathscr{I}(t) \xi' \right| \le \\
&\frac{1}{4} \sum_{u,v=1}^m \big|\xi^{(u)} \xi'^{(v)}\big| \int_{I(t)} \left| \frac{1}{n} \sum_{i=1}^n \Psi_{n,i}^{(u)}(s)\, \Psi_{n,i}^{(v)}(s)\, Y_n^{(i)}(s) - \Psi^{(u,v)}(s) \right| \alpha_0(s)\, \mathrm{d}s \\
&\qquad\qquad\qquad \le \frac{1}{4} |\xi|^{\mathrm{T}} \mathscr{D}_n(t) |\xi'| \cdot \int_{I(t)} \alpha_0(s)\, \mathrm{d}s,
\end{aligned}$$

Assumption 2.3.1.i and (2.16) directly imply $T_n^{(1)}(t) \longrightarrow_{P_{n,0}} \frac{1}{4}\xi^{\mathrm{T}} \mathscr{I}(t)\xi'$. Therefore, it remains to be shown that $T_n^{(i)}(t)$, $i = 2, 3, 4$, converge to 0 in $P_{n,0}$-probability. Note that in the following $\sup_{i,s} = \sup_{1\le i\le n, s\in[0,t]}$. Because of $\xi_{n,i}(s) \in \big[0, \xi\big]$ it holds that

$$\begin{aligned}
\sup_{i,s}\Big|\frac{1}{2\sqrt{n}}\xi_{n,i}(s)^{\mathrm{T}}\Psi_{n,i}(s)\,Y_n^{(i)}(s)\Big| &\le \sup_{i,s}\frac{1}{2\sqrt{n}}\sum_{u=1}^{m}\big|\xi_{n,i}^{(u)}(s)\big|\cdot\big|\Psi_{n,i}^{(u)}(s)\,Y_n^{(i)}(s)\big| \\
&\le c\cdot\sup_{i,s}\frac{1}{\sqrt{n}}\sum_{u=1}^{m}\big|\Psi_{n,i}^{(u)}(s)\,Y_n^{(i)}(s)\big| \\
&\le m\cdot c\cdot\sup_{i,s}\frac{1}{\sqrt{n}}\max_{1\le u\le m}\big|\Psi_{n,i}^{(u)}(s)\,Y_n^{(i)}(s)\big|,
\end{aligned}$$

where $c = \max_{1\le u\le m}|\xi^{(u)}|$. Assumption 2.3.1.iv gives

$$\sup_{i,s}\Big|\frac{1}{2\sqrt{n}}\xi_{n,i}(s)^{\mathrm{T}}\Psi_{n,i}(s)\,Y_n^{(i)}(s)\Big| \longrightarrow_{P_{n,0}} 0 \quad\text{and}\quad \sup_{i,s} R_n^{(i)}(s,\xi_{n,i}(s)) \longrightarrow_{P_{n,0}} 1. \tag{2.17}$$

One easily obtains that

$$\begin{aligned}
\big|T_n^{(2)}(t)\big| &\le \sup_{i,s}\Big|\frac{1}{2\sqrt{n}}\xi'^{\mathrm{T}}\Psi_{n,i}(s)\,Y_n^{(i)}(s)\Big|\cdot\int_{I(t)}\frac{1}{n}\sum_{i=1}^{n}Q_{n,\xi}^{(i)}(s)\,\lambda_{n,0}^{(i)}(s)\,\mathrm{d}s \\
&\le \sup_{i,s} R_n^{(i)}(t,\xi_{n,i}(t))\sup_{i,s}\Big|\frac{1}{2\sqrt{n}}\xi'^{\mathrm{T}}\Psi_{n,i}(s)\,Y_n^{(i)}(s)\Big| \\
&\qquad\times\int_{I(t)}\frac{1}{n}\sum_{i=1}^{n}\big(\xi^{\mathrm{T}}\Psi_{n,i}(s)\big)^2\,\lambda_{n,0}^{(i)}(s)\,\mathrm{d}s \\
&\le \sup_{i,s} R_n^{(i)}(s,\xi_{n,i}(s))\cdot\sup_{i,s}\Big|\frac{1}{2\sqrt{n}}\xi'^{\mathrm{T}}\Psi_{n,i}(s)\,Y_n^{(i)}(s)\Big| \\
&\qquad\times\left(|\xi|^{\mathrm{T}}\mathscr{D}_n|\xi|\cdot\int_{I(t)}\alpha_0(s)\,\mathrm{d}s + |\xi|^{\mathrm{T}}\mathscr{I}(t)|\xi|\right).
\end{aligned}\tag{2.18}$$

Assumption 2.3.1.i, (2.16) and (2.17) imply $T_n^{(2)}(t) \longrightarrow_{P_{n,0}} 0$. By a similar estimate one shows $T_n^{(3)}(t) \longrightarrow_{P_{n,0}} 0$.

In a last step we show $T_n^{(4)} \longrightarrow_{P_{n,0}} 0$. First, note that (2.17) implies

$$\sup_{i,s}\Big|\frac{1}{2\sqrt{n}}\xi_{n,i}(s)^{\mathrm{T}}\Psi_{n,i}(s)\,Y_n^{(i)}(s)\Big|^2 \longrightarrow_{P_{n,0}} 0, \tag{2.19}$$

then (2.16), (2.17) and (2.19) using the estimate

$$
\begin{aligned}
|T_n^{(4)}(t)| \leq \frac{1}{64} \sup_{i,s} R_n^{(i)}(s, \xi'_{n,i}(s)) \cdot \sup_{i,s} R_n^{(i)}(s, \xi_{n,i}(s)) \\
\times \int_{I(t)} \frac{1}{n^2} \sum_{i=1}^{n} (\xi'^{\mathrm{T}} \Psi_{n,i}(s))^2 \, (\xi^{\mathrm{T}} \Psi_{n,i}(s))^2 \, \lambda_{n,0}^{(i)}(s) \, \mathrm{d}s \leq \\
\frac{1}{16} \sup_{i,s} R_n^{(i)}(t, \xi'_{n,i}(s)) \; \sup_{i,s} R_n^{(i)}(s, \xi_{n,i}(s)) \; \sup_{i,s} \left| \frac{1}{2\sqrt{n}} \xi'^{\mathrm{T}} \Psi_{n,i}(s) \, Y_n^{(i)}(s) \right|^2 \\
\times \left(|\xi|^{\mathrm{T}} \mathscr{D}_n(t) |\xi| \cdot \int_{I(t)} \alpha_0(s) \, \mathrm{d}s + |\xi|^{\mathrm{T}} \mathscr{J}(t) |\xi| \right)
\end{aligned}
\tag{2.20}
$$

give the assertion and therefore condition (2.3) holds. The second condition to verify is (2.4), *i.e.*

$$
\sum_{i=1}^{n} \int_{I(t)} \left(R_n^{(i)}(s, \xi) - 1 \right)^2 \mathbb{1}\left(\left| R_n^{(i)}(s, \xi) - 1 \right| > \varepsilon \right) \lambda_{n,0}^{(i)}(s) \, \mathrm{d}s \longrightarrow_{P_{n,0}} 0
$$

for all $t \in [0, \tau]$ and for all $\varepsilon > 0$. Again, a Taylor-expansion gives that the right hand side of (2.4) equals $T_{n,\varepsilon}^{(1)}(t) + 2T_{n,\varepsilon}^{(2)}(t) + T_{n,\varepsilon}^{(3)}(t)$, where

$$
\begin{aligned}
T_{n,\varepsilon}^{(1)}(t) &= \int_{I(t)} \frac{1}{4n} \sum_{i=1}^{n} \left(\xi^{\mathrm{T}} \Psi_{n,i}(s) \right)^2 \mathbb{1}\left(\left| R_n^{(i)}(s, \xi) - 1 \right| > \varepsilon \right) \lambda_{n,0}^{(i)}(s) \, \mathrm{d}s, \\
T_{n,\varepsilon}^{(2)}(t) &= \int_{I(t)} \frac{2}{16 n^{3/2}} \sum_{i=1}^{n} Q_{n,\xi}^{(i)}(s) \, \xi^{\mathrm{T}} \Psi_{n,i}(s) \; \mathbb{1}\left(\left| R_n^{(i)}(s, \xi) - 1 \right| > \varepsilon \right) \lambda_{n,0}^{(i)}(s) \, \mathrm{d}s, \\
T_{n,\varepsilon}^{(3)}(t) &= \int_{I(t)} \frac{1}{64 n^2} \sum_{i=1}^{n} \left(Q_{n,\xi}^{(i)}(s) \right)^2 \mathbb{1}\left(\left| R_n^{(i)}(s, \xi) - 1 \right| > \varepsilon \right) \lambda_{n,0}^{(i)}(s) \, \mathrm{d}s.
\end{aligned}
$$

It holds the estimates

$$
\left| T_{n,\varepsilon}^{(2)}(t) \right| \leq \sup_{i,s} \left| \frac{1}{2\sqrt{n}} \xi^{\mathrm{T}} \Psi_{n,i}(s) \; Y_n^{(i)}(s) \right| \cdot \int_{I(t)} \frac{1}{n} \sum_{i=1}^{n} Q_{n,\xi}^{(i)}(s) \, \lambda_{n,0}^{(i)}(s) \, \mathrm{d}s,
$$

and

$$
\left| T_{n,\varepsilon}^{(3)}(t) \right| \leq \int_{I(t)} \frac{1}{64 n^2} \sum_{i=1}^{n} \left(Q_{n,\xi}^{(i)}(s) \right)^2 \lambda_{n,0}^{(i)}(s) \, \mathrm{d}s.
$$

Now, (2.18) and (2.20) (with $\xi' = \xi$) give that $T_{n,\varepsilon}^{(i)}(t) \longrightarrow_{P_{n,0}} 0$, $i = 2, 3$. Thus, it merely remains to be proved that $T^{(1)}_{n,\varepsilon}(t) \longrightarrow_{P_{n,0}} 0$. Applying a Taylor-expansion to the function in the indicator function and using the abbreviation

$$C = \sup_{i,s}\Big|\frac{1}{2\sqrt{n}}\xi^{\mathrm{T}}\Psi_{n,i}(s)\ Y^{(i)}_n(s)\Big| \\ + \sup_{i,s} R^{(i)}_n(s,\xi_{n,i}(s)) \cdot \sup_{i,s}\Big|\frac{1}{2\sqrt{n}}\xi^{\mathrm{T}}\Psi_{n,i}(s)\ Y^{(i)}_n(s)\Big|^2$$

one gets

$$\big|T^{(1)}_{n,\varepsilon}(t)\big| \le \mathbb{1}\big(C > \varepsilon\big) \cdot \int_{I(t)} \frac{1}{n}\sum_{i=1}^{n}\big(\xi^{\mathrm{T}}\Psi_{n,i}(s)\big)^2\ \lambda^{(i)}_{n,0}(s)\,\mathrm{d}s. \tag{2.21}$$

Note that $\mathbb{1}\big(C > \varepsilon\big) \longrightarrow_{P_{n,0}} 0$ because of (2.17) and (2.19). The estimate (2.21) gives $T^{(1)}_{n,\varepsilon}(t) \longrightarrow_{P_{n,0}} 0$. Thus, (2.4) also holds. The matrix $\mathscr{J}(t)$ is obviously symmetric. Because of

$$\sum_{i=1}^{n}\int_{I(t)}\big(R^{(i)}_n(s,\xi)-1\big)^2\,\lambda^{(i)}_{n,0}(s)\,\mathrm{d}s \ge 0 \quad \text{for all } n \in \mathbb{N} \text{ and } \xi \in \mathbb{R}^m$$

and the convergence in (2.3), it holds that $\mathscr{J}(t)$ is positive semi-definite. Theorem 2.2.7 implies (2.11).

We show that S_n is indeed a central sequence. A Taylor expansion, cf. (2.15), gives

$$2\sum_{i=1}^{n}\int_{I(\tau)} R^{(i)}_n(s,\xi) - 1\,\mathrm{d}M^{(i)}_{n,0}(s) - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_{I(\tau)}\xi^{\mathrm{T}}\Psi_{n,i}(s)\,Y^{(i)}_n(s)\,\mathrm{d}M^{(i)}_{n,0}(s) \\ = -\frac{1}{4n}\sum_{i=1}^{n}\int_{I(\tau)} R_{n,i}(s,\xi_{n,i}(s)) \cdot \big(\xi^{\mathrm{T}}\Psi_{n,i}(s)\big)^2\,Y^{(i)}_n(s)\,\mathrm{d}M^{(i)}_{n,0}(s).$$

We want to apply Lemma 2.1.8. $R_{n,i}(s,\xi_{n,i}(s)) \cdot \big(\xi^{\mathrm{T}}\Psi_{n,i}(s)\big)^2\,Y^{(i)}_n(s)$, $i = 1,\dots,n$, are predictable and locally bounded because of (2.15).

The estimate

$$0 \le \int_{I(\tau)} \frac{1}{16n^2}\sum_{i=1}^{n}\big(R^{(i)}_n(s,\xi_{n,i}(s))\big)^2\ \big(\xi^{\mathrm{T}}\Psi_{n,i}(s)\big)^4\lambda^{(i)}_{n,0}(s)\,\mathrm{d}s \le 4\,T^{(4)}_n(\tau)$$

and $T_n^{(4)}(\tau) \longrightarrow_{P_{n,0}} 0$ $(\xi' = \xi)$ yield that $2\sum_{i=1}^n \int_{I(\tau)} R_n^{(i)}(s,\xi) - 1 \,\mathrm{d}M_{n,0}^{(i)}(s)$ and $\xi^{\mathrm{T}} S_n(\tau)$ are asymptotically equivalent. We know that $\xi^{\mathrm{T}} S_n(\tau) \xrightarrow{\mathfrak{D}}_{P_{n,0}} \mathcal{N}\big(0, \xi^{\mathrm{T}} \mathscr{I}(\tau)\xi\big)$, for all $\xi \in \mathbb{R}^m$. By applying the Cramér-Wold-device, cf. Billingsley [9, Theorem 7.7], one obtains that S_n is a central sequence. Up to now, it was shown that $\big(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathfrak{P}_n\big)$, $n \in \mathbb{N}$, restricted to time τ is asymptotically normal.

In order to proof the last assertion one uses Theorem 2.1.1 to show that (2.3) and (2.4) also hold for τ_0. Assume that $\tau_k < \tau_0$, $k \in \mathbb{N}$, and $\lim_{k\to\infty} \tau_k = \tau_0$. Set

$$\begin{aligned} X_{n,k} &= \sum_{i=1}^n \int_{I(\tau_k)} \big(R_n^{(i)}(s,\xi) - 1\big)\big(R_n^{(i)}(s,\xi') - 1\big)\, \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s, \\ \widetilde{X}_n &= \sum_{i=1}^n \int_{I(\tau_0)} \big(R_n^{(i)}(s,\xi) - 1\big)\big(R_n^{(i)}(s,\xi') - 1\big)\, \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s, \end{aligned}$$

$X_k = \frac{1}{4}\xi^{\mathrm{T}} \mathscr{I}(\tau_k)\xi'$ and $X = \frac{1}{4}\xi^{\mathrm{T}} \mathscr{I}(\tau_0)\xi'$. We showed that $X_{n,k} \xrightarrow{\mathfrak{D}}_{P_{n,0}} X_k$, as $n \to \infty$, and obviously it holds that $X_k \xrightarrow{\mathfrak{D}}_{P_{n,0}} X$ as $n \to \infty$. As we assumed that

$$\lim_{k\to\infty} \limsup_{n\to\infty} P_{n,0}\big(\big|X_{n,k} - \widetilde{X}_n\big| \geq \delta\big) = 0 \tag{2.22}$$

for all $\delta > 0$, it follows that (2.3) also holds for τ_0. With a similar consideration one shows that (2.4) holds. Set

$$\begin{aligned} X'_{n,k} &= \sum_{i=1}^n \int_{I(\tau_k)} \big(R_n^{(i)}(s,\xi) - 1\big)^2 \mathbb{1}\big(\big|R_n^{(i)}(s,\xi) - 1\big| \geq \varepsilon\big)\, \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s, \\ \widetilde{X}'_n &= \sum_{i=1}^n \int_{I(\tau_0)} \big(R_n^{(i)}(s,\xi) - 1\big)^2 \mathbb{1}\big(\big|R_n^{(i)}(s,\xi) - 1\big| \geq \varepsilon\big)\, \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s, \end{aligned}$$

and $X'_k = 0$ and $X' = 0$. Because of

$$P_{n,0}\big(|X'_{n,k} - \widetilde{X}'_n| \geq \delta\big) \leq P_{n,0}\big(|X_{n,k} - \widetilde{X}_n| \geq \delta\big) \quad (\text{with } \xi = \xi'),$$

(2.22) and Theorem 2.1.1 give that (2.4) also holds for τ_0. Applying Theorem 2.2.7 yields that the assertions also hold for τ_0 except (2.12).

In the last step we show that $P_{n,0}\big(\big|\widetilde{S}_{n,\xi}(\tau_0) - \xi^{\mathrm{T}} S_n(\tau_0)\big| \geq \varepsilon\big) \to 0$, as $n \to \infty$. It holds that

$$\begin{aligned}P_{n,0}\big(\big|\widetilde{S}_{n,\xi}(\tau_0) - \xi^{\mathrm{T}} S_{n,\xi}(\tau_0)\big| \geq \varepsilon\big) &\leq P_{n,0}\big(\big|\widetilde{S}_{n,\xi}(\tau_0) - \widetilde{S}_{n,\xi}(\tau)\big| \geq \varepsilon/3\big) + \\ &P_{n,0}\big(\big|\widetilde{S}_{n,\xi}(\tau) - \xi^{\mathrm{T}} S_n(\tau)\big| \geq \varepsilon/3\big) + P_{n,0}\big(\big|\xi^{\mathrm{T}} S_n(\tau) - \xi^{\mathrm{T}} S_n(\tau_0)\big| \geq \varepsilon/3\big),\end{aligned}$$

where $\widetilde{S}_{n,\xi}(t) = 2\sum_{i=1}^{n} \int_{I(t)} R_n^{(i)}(s,\xi) - 1 \,\mathrm{d}M_{n,0}^{(i)}(s)$. It was already shown that the second summand on the right hand side converges to 0 as $n \to \infty$. Therefore, there remains to be shown that the first and the third summand on the right hand side get arbitrarily small. For all $\eta > 0$ we can choose $\tau \in \mathbb{R}_+$, such that $h_\xi = \eta/4 - \xi^{\mathrm{T}}\big(\mathscr{I}(\tau_0) - \mathscr{I}(\tau)\big)\xi/4 > 0$ and

$$\limsup_{n\to\infty} P_{n,0}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\int_{(\tau,\tau_0)} \big(\xi^{\mathrm{T}}\Psi_{n,i}(s)\big)^2 \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s\right| \geq \eta\right) \leq 9\eta/\varepsilon^2, \tag{2.23}$$

where we use equation (2.14).

Applying Corollary 2.1.7 to the first summand gives

$$\begin{aligned}&P_{n,0}\big(\big|\widetilde{S}_{n,\xi}(\tau_0) - \widetilde{S}_{n,\xi}(\tau)\big| \geq \varepsilon/3\big) \leq \\ &\frac{9\eta}{\varepsilon^2} + P_{n,0}\left(\sum_{i=1}^{n}\int_{(\tau,\tau_0)} \big(R_n^{(i)}(s,\xi) - 1\big)^2 \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s \geq \eta/4\right) \leq \\ &\frac{9\eta}{\varepsilon^2} + P_{n,0}\left(\left|\sum_{i=1}^{n}\int_{I(\tau)} \big(R_n^{(i)}(s,\xi) - 1\big)^2 \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s - \xi^{\mathrm{T}}\mathscr{I}(\tau)\xi/4\right| \geq h_\xi/2\right) \\ &\quad + P_{n,0}\left(\left|\sum_{i=1}^{n}\int_{I(\tau_0)} \big(R_n^{(i)}(s,\xi) - 1\big)^2 \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s - \xi^{\mathrm{T}}\mathscr{I}(\tau_0)\xi/4\right| \geq h_\xi/2\right).\end{aligned}$$

We get that $\limsup_{n\to\infty} P_{n,0}\big(\big|\widetilde{S}_{n,\xi}(\tau_0) - \widetilde{S}_{n,\xi}(\tau)\big| \geq \varepsilon/3\big) \leq 9\eta/\varepsilon^2$, since (2.3) holds for all $t \in [0,\tau_0]$. Again, applying Corollary 2.1.7 to the third summand gives

$$\begin{aligned}&P_{n,0}\big(\big|\xi^{\mathrm{T}} S_n(\tau) - \xi^{\mathrm{T}} S_n(\tau_0)\big| \geq \varepsilon/3\big) \\ &\qquad\qquad \leq \frac{9\eta}{\varepsilon^2} + P_{n,0}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\int_{(\tau,\tau_0)} \big(\xi^{\mathrm{T}}\Psi_{n,i}(s)\big)^2 \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s\right| \geq \eta\right).\end{aligned}$$

Thus, it holds that $\limsup_{n\to\infty} P_{n,0}\big(|\xi^{\mathrm{T}} S_n(\tau) - \xi^{\mathrm{T}} S_n(\tau_0)| \geq \varepsilon/3\big) \leq 18\eta/\varepsilon^2$, where we use (2.23). All in all, we proved that

$$\limsup_{n\to\infty} P_{n,0}\Big(\big|\widetilde{S}_{n,\xi}(\tau_0) - \xi^{\mathrm{T}} S_{n,\xi}(\tau_0)\big| \geq \varepsilon\Big) \leq \frac{27\eta}{\varepsilon^2},$$

using the fact that η was arbitrarily chosen, the result is our assertion. □

The next result gives the asymptotic distribution of a central sequence under alternatives.

2.3.4 Corollary. In the situation of Theorem 2.3.3, it holds that

$$S_n(\tau) \xrightarrow{\mathfrak{D}}_{P_{n,\xi}} \mathcal{N}\big(\mathscr{I}(\tau)\xi, \mathscr{I}(\tau)\big).$$

Proof. As $S_n(\tau)$, $n \in \mathbb{N}$ is a central sequence, we know by applying the Cramér-Wold device, cf. Witting and Müller-Funk [72, Korollar 5.69], that $\vartheta^{\mathrm{T}} S_n(\tau) \xrightarrow{\mathfrak{D}}_{P_{n,0}} \mathcal{N}(0, \vartheta^{\mathrm{T}} \mathscr{I} \vartheta)$ for all $\vartheta \in \mathbb{R}^m$, therefore it holds that

$$\begin{aligned}
&\zeta_1\Big(\xi^{\mathrm{T}} S_n(\tau) - \frac{1}{2}\xi^{\mathrm{T}} \mathscr{I}(\tau)\xi\Big) + \zeta_2 \vartheta^{\mathrm{T}} S_n(\tau)\\
&\quad = \big(\zeta_1 \cdot \xi + \zeta_2 \cdot \vartheta\big)^{\mathrm{T}} S_n(\tau) - \frac{1}{2}\zeta_1 \xi^{\mathrm{T}} \mathscr{I}(\tau)\xi\\
&\qquad\quad \xrightarrow{\mathfrak{D}}_{P_{n,0}} \mathcal{N}\Big(-\frac{1}{2}\zeta_1\xi^{\mathrm{T}} \mathscr{I}(\tau)\xi,\ (\zeta_1 \cdot \xi + \zeta_2 \cdot \vartheta)^{\mathrm{T}} \mathscr{I}(\tau)(\zeta_1 \cdot \xi + \zeta_2 \cdot \vartheta)\Big)
\end{aligned}$$

for all $(\zeta_1, \zeta_2)^{\mathrm{T}} \in \mathbb{R}^2$, where we applied Slutsky's Lemma, cf. Witting and Müller-Funk [72, Korollar 5.84]. Noting that

$$\zeta_1 \log \Upsilon_{n,\xi}(\tau) + \zeta_2 \vartheta^{\mathrm{T}} S_n - \Big(\zeta_1\Big(\xi^{\mathrm{T}} S_n - \frac{1}{2}\xi^{\mathrm{T}} \mathscr{I}\xi\Big) + \zeta_2 \vartheta^{\mathrm{T}} S_n\Big) \longrightarrow_{P_{n,0}} 0$$

and applying the Cramér-Wold device yields

$$\begin{pmatrix} \log \Upsilon_{n,\xi}(\tau) \\ \vartheta^{\mathrm{T}} S_n(\tau) \end{pmatrix} \xrightarrow{\mathfrak{D}}_{P_{n,0}} \mathcal{N}\left(\begin{pmatrix} -\frac{1}{2}\xi^{\mathrm{T}} \mathscr{I}\xi \\ 0 \end{pmatrix}, \begin{pmatrix} \xi^{\mathrm{T}} \mathscr{I}\xi & \vartheta^{\mathrm{T}} \mathscr{I}\xi \\ \vartheta^{\mathrm{T}} \mathscr{I}\xi & \vartheta^{\mathrm{T}} \mathscr{I}\vartheta \end{pmatrix}\right).$$

Le Cam's Third Lemma, cf. Witting and Müller-Funk [72, Korollar 6.139], implies $\vartheta^{\mathrm{T}} S_n(\tau) \xrightarrow{\mathfrak{D}}_{P_{n,\xi}} \mathcal{N}(\vartheta^{\mathrm{T}} \mathscr{I}\xi, \vartheta^{\mathrm{T}} \mathscr{I}\vartheta)$ for all $\vartheta \in \mathbb{R}^m$. Applying the Cramér-Wold device gives the second assertion. □

2.3.5 Remark. In the proof of Theorem 2.3.3 we merely use the fact that that the exponential function is two times continuously differentiable with the Taylor expansion $\exp(x) = 1+x+O(x^2)$ at the point $x_0 = 0$. Therefore, we can substitute the exponential function by any strictly positive function g that is two times continuously differentiable with the Taylor expansion $g(x) = 1+x+O(x^2)$ in the point $x_0 = 0$. In order to receive some suitable stochastic ordering one should also demand that this function is monotone. In the subsequent considerations we only employ the just mentioned properties of the exponential function.

Now, it is intended to apply the previous result to localized, q-dimensional parametric sub-models, *i.e.* Definition 1.3.4 holds for all $n \in \mathbb{N}$. In particular, this means that

$$\alpha_{n,\xi}^{(i)}(s) = \exp\left(\frac{1}{\sqrt{n}} \cdot \beta^{\mathrm{T}} Z_{n,i} \circledcirc \gamma(s) + \frac{1}{\sqrt{n}} \cdot \eta^{\mathrm{T}} \widetilde{\gamma}(s)\right)\alpha_0(s) \qquad s \in \mathbb{R}_+,$$

$i = 1, \ldots, n$, where $Z_{n,i}$ denotes the covariate process of the i-th observation. However, we have to introduce some more notions and assumptions. The following Definition will play a central role in Chapter 3.

2.3.6 Definition (Parametric Sub-Sub-Model). a) Let

$$\mathfrak{Q}_n = \left\{Q_{n,\bar{\xi}} \mid \bar{\xi} = (\bar{\beta}^{\mathrm{T}}, \bar{\eta}^{\mathrm{T}}),\ \bar{\beta} \in \mathbb{R}^{\bar{r}},\ \bar{\eta} \in \mathbb{R}^{\bar{q}}\right\} \subset \mathfrak{P}_n$$

be a family of probability measures, $1 \leq \bar{r} \leq r$ and $1 \leq \bar{q} \leq q$. If there exists two matrices $\mathscr{T}_1 \in \mathbb{R}^{r \times \bar{r}}$ and $\mathscr{T}_2 \in \mathbb{R}^{q \times \bar{q}}$, such that under $Q_{n,\bar{\xi}}$ the counting process N_n has the $\mathbb{F}_n$-compensator $A_{n,\mathscr{T}\bar{\xi}}$, where

$$\mathscr{T} = \begin{pmatrix} \mathscr{T}_1 & 0 \\ 0 & \mathscr{T}_2 \end{pmatrix} \in \mathbb{R}^{(r+q)\times(\bar{r}+\bar{q})},$$

for all $\bar{\xi} \in \mathbb{R}^{\bar{r}+\bar{q}}$, then we call $\left(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{Q}_n\right)$ a $(\bar{r}, \bar{q})$-sub-sub-model of the localized, q-dimensional parametric sub-model $\left(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{P}_n\right)$ with transformation-matrix $\mathscr{T}$.

b) Assume that $(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{Q}_n)$ is a $(\bar{r}, \bar{q})$-sub-sub-model with transformation-matrix $\mathscr{T}$ for all $n \in \mathbb{N}$. Then the sequence $(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{Q}_n)$, $n \in \mathbb{N}$, is called a sequence of $(\bar{r}, \bar{q})$-sub-sub-models with transformation-matrix $\mathscr{T}$.

In this Definition we allow the interesting parameter β and the nuisance parameter η only to vary in the subspaces given by the matrices $\mathscr{T}_1$ and $\mathscr{T}_2$. In other words some additional restrictions are imposed on the parameters. These restrictions mean that we also have some more information on the parameter. Considering sequences of sub-sub-models turns out to be important for determining sequences of so-called hardest parametric sub-models of the MCRM. These sequences share some nice properties and enable us to derive a suitable statistic on which we can base the testing on β. However, before we can proceed we need some more notation.

2.3.7 Definition. a) We agree that

$$\dot{u} = l, \quad \text{if} \sum_{v=1}^{l-1} r_v < u \leq \sum_{v=1}^{l} r_v,\ l \in \{1, \ldots, p\}, \quad \text{and} \quad \ddot{u} = u - \sum_{v=1}^{\dot{u}-1} r_v,$$

for $1 \leq u \leq r$ where the r_u's were introduced in Definition 1.3.2. The functions $\dot{u}$ and $\ddot{u}$ are called index functions.

b) In the following we use the abbreviations

$$\begin{aligned}
\widehat{\mu}_{n,0}(s) &= \frac{1}{n}\sum_{i=1}^{n} Y_n^{(i)}(s), \\
\widehat{\mu}_{n,1}^{(u)}(s) &= \frac{1}{n}\sum_{i=1}^{n} Z_{n,i}^{(u)}(s)\, Y_n^{(i)}(s), && u = 1, \ldots, p, \\
\widehat{\mu}_{n,2}^{(u,v)}(s) &= \frac{1}{n}\sum_{i=1}^{n} Z_{n,i}^{(u)}(s)\, Z_{n,i}^{(v)}(s)\, Y_n^{(i)}(s), && u, v = 1, \ldots, p.
\end{aligned}$$

2.3.8 Remark. a) Definition 2.3.7.a enables us to easily determine the weight function $\gamma^{(\cdot,\cdot)}$ corresponding with $\beta^{(u)}$, $u \in \{1, \ldots, r\}$. The index functions help us to keep the notation simple, see also Remark 1.3.3.b.

b) The quantities of Definition 2.3.7.b will be crucial for the further treatment of the MCRM. $\widehat{\mu}_{n,0}$ is an estimator of the survival function of the censored survival times. $\widehat{\mu}_{n,1}$ and $\widehat{\mu}_{n,2}$ are basically estimators of the moments of covariates. However, these estimators only use the covariates of the individuals that are still at-risk, *i.e.* they only depend on observations that are available. Note that $\widehat{\mu}_{n,0}(s) = 0$ always implies that $\widehat{\mu}_{n,1}^{(u)}(s) = 0$, $u = 1, \dots, p$. Using the definition $0/0 = 0$ gives that the fraction $\widehat{\mu}_{n,1}^{(u)}/\widehat{\mu}_{n,0}$ is always defined.

2.3.9 Assumption. The functions μ_0, $\mu_1^{(u)}$ and $\mu_2^{(u,v)}$, $u, v = 1, \dots, p$, are measurable and real-valued. Furthermore, it is assumed that

i) $\tau_0 = \sup\{t \in \mathbb{R}_+ \mid \int_{[0,t]} \alpha_0(s)\,\mathrm{d}s < \infty\}$, where $\alpha_0 \in \aleph$ is a measurable, non-negative function called baseline hazard.

ii) The processes $\{Z_{n,i}^{(u)}(s)\, Y_n^{(i)}(s) \mid s \in \mathbb{R}_+\}$, $u = 1, \dots, p$, $i = 1, \dots, n$, $n \in \mathbb{N}$, are predictable and locally bounded.

iii) μ_0, $\mu_1^{(u)}$ and $\mu_2^{(u,v)}$, $u, v = 1, \dots, p$, are bounded on every interval $[0, t]$ for all $t < \tau_0$.

iv) $\sup_{s \in [0,t]} \left|\widehat{\mu}_{n,0}(s) - \mu_0(s)\right| \longrightarrow_{P_{n,0}} 0$ for all $t < \tau_0$.

v) $\sup_{s \in [0,t]} \left|\widehat{\mu}_{n,1}^{(u)}(s) - \mu_1^{(u)}(s)\right| \longrightarrow_{P_{n,0}} 0$ for all $u = 1, \dots, p$ and $t < \tau_0$.

vi) $\sup_{s \in [0,t]} \left|\widehat{\mu}_{n,2}^{(u,v)}(s) - \mu_2^{(u,v)}(s)\right| \longrightarrow_{P_{n,0}} 0$ for all $u, v = 1, \dots, p$ and $t < \tau_0$,

vii) $\sup_{i \in \{1,\dots,n\},\, s \in [0,t]} \left\{ \frac{1}{\sqrt{n}} \max_{1 \le u \le p} \left|Z_{n,i}^{(u)}(s)\ Y_n^{(i)}(s)\right| \right\} \longrightarrow_{P_{n,0}} 0$ for all $t < \tau_0$.

viii) The real-valued functions $\gamma^{(\dot{u},\ddot{u})}$, $u = 1, \dots, r$, are bounded on every interval $[0, t]$ for all $t < \tau_0$.

ix) The real-valued functions $\widetilde{\gamma}^{(u)}$, $u = 1, \dots, q$, are bounded on every interval $[0, t]$ for all $t < \tau_0$.

Andersen and Gill, who consider the asymptotic properties of the partial likelihood estimator for β under the CRM, have to assume conditions similar to Assumption 2.3.9 for deriving their results, cf. Andersen *et al.* [4, Condition VII.2.1] or Andersen and Gill [5]. After this preparations we can state the main result of this section.

2.3.10 Theorem. a) Let $I(t)$ denote the interval $[0,t]$, if $t<\infty$, or $[0,\infty)$, if $t=\infty$. Under Assumption 2.2.1 and Assumption 2.3.9, the sequence of localized, q-dimensional parametric sub-models of the MCRM $(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathfrak{P}_n)$, $n\in\mathbb{N}$, is asymptotically normal restricted time τ, for all $\tau<\tau_0$. The asymptotic information matrix

$$\mathscr{I}(\tau) = \begin{pmatrix} \mathscr{I}_{1,1}(\tau) & \mathscr{I}_{1,2}(\tau) \\ \mathscr{I}_{2,1}(\tau) & \mathscr{I}_{2,2}(\tau) \end{pmatrix},$$

is given by the matrices

$$\begin{aligned} \mathscr{I}_{1,1}(\tau) &= \big(\mathscr{I}_{1,1}^{(u,v)}(\tau)\big) \in \mathbb{R}^{r\times r}, \\ \mathscr{I}_{2,2}(\tau) &= \big(\mathscr{I}_{2,2}^{(u,v)}(\tau)\big) \in \mathbb{R}^{q\times q}, \\ \mathscr{I}_{1,2}(\tau) = \mathscr{I}_{2,1}^{\mathrm{T}}(\tau) &= \big(\mathscr{I}_{1,2}^{(u,v)}(\tau)\big) \in \mathbb{R}^{r\times q}, \end{aligned}$$

where

$$\begin{aligned} \mathscr{I}_{1,1}^{(u,v)}(\tau) &= \int_{I(\tau)} \gamma^{(\dot u,\ddot u)}(s)\,\gamma^{(\dot v,\ddot v)}(s)\,\mu_2^{(\dot u,\dot v)}(s)\,\alpha_0(s)\,\mathrm{d}s, \\ \mathscr{I}_{2,2}^{(u,v)}(\tau) &= \int_{I(\tau)} \widetilde\gamma^{(u)}(s)\,\widetilde\gamma^{(v)}(s)\,\mu_0(s)\,\alpha_0(s)\,\mathrm{d}s, \\ \mathscr{I}_{1,2}^{(u,v)}(\tau) &= \int_{I(\tau)} \gamma^{(\dot u,\ddot u)}(s)\,\widetilde\gamma^{(v)}(s)\,\mu_1^{(\dot u)}(s)\,\alpha_0(s)\,\mathrm{d}s. \end{aligned}$$

A central sequence is given by $S_n(\tau) = \big(S_{n,1}^{\mathrm{T}}(\tau), S_{n,2}^{\mathrm{T}}(\tau)\big)^{\mathrm{T}}$, where

$$\begin{aligned} S_{n,1}(\tau) &= \frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n}\int_{I(\tau)} \gamma^{(\dot u,\ddot u)}(s)\,Z_{n,i}^{(\dot u)}(s)\,\mathrm{d}M_{n,0}^{(i)}(s),\ u=1,\dots,r\right)^{\mathrm{T}}, \\ S_{n,2}(\tau) &= \frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n}\int_{I(\tau)} \widetilde\gamma^{(u)}(s)\,\mathrm{d}M_{n,0}^{(i)}(s),\ u=1,\dots,q\right)^{\mathrm{T}}. \end{aligned}$$

Assume that $\lim_{t\to\tau_0} \mathscr{I}(t) = \mathscr{I}(\tau_0)$ exists and set

$$\Psi_{n,i}(t) = Y_n^{(i)}(t) \cdot \Big(\big(Z_{n,i} \circledcirc \gamma(t)\big)^{\mathrm{T}}, \widetilde{\gamma}^{\mathrm{T}}(t) \Big)^{\mathrm{T}}, \qquad t \in \mathbb{R}_+, \quad i = 1, \dots, n.$$

If the condition

$$\lim_{t\to\tau_0} \limsup_{n\to\infty} P_{n,0}\left(\left| \sum_{i=1}^{n} \int_{(t,\tau_0)} V_n^{(i)}(s,\xi,\xi')\ \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s \right| \geq \varepsilon \right) = 0, \qquad (2.24)$$

where

$$V_n^{(i)}(s,\xi,\xi') = \left(\exp\!\left(\frac{1}{2\sqrt{n}} \xi^{\mathrm{T}} \Psi_{n,i}(s) \right) - 1 \right) \left(\exp\!\left(\frac{1}{2\sqrt{n}} {\xi'}^{\mathrm{T}} \Psi_{n,i}(s) \right) - 1 \right),$$

and $\lambda_{n,0}^{(i)}(s) = Y_n^{(i)}(s)\,\alpha_0(s)$, $s \in \mathbb{R}_+$, and the condition

$$\lim_{t\to\tau_0} \limsup_{n\to\infty} P_{n,0}\left(\left| \frac{1}{n} \sum_{i=1}^{n} \int_{(t,\tau_0)} \big(\xi^{\mathrm{T}} \Psi_{n,i}(s)\big)^2\ \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s \right| \geq \varepsilon \right) = 0 \qquad (2.25)$$

hold for all $\xi, \xi' \in \mathbb{R}^{r+q}$ and $\varepsilon > 0$, then the previous assertions also hold for $\tau = \tau_0$.

b) Let $\big(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathfrak{Q}_n\big)$, $n \in \mathbb{N}$, be a sequence of $(\bar{r},\bar{q})$-sub-sub-models with transformation matrix $\mathscr{T}$ in the sense of Definition 2.3.6. Under Assumption 2.2.1 and Assumption 2.3.9, the sequence $\big(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathfrak{Q}_n\big)$, $n \in \mathbb{N}$, is asymptotically normal restricted to time τ with asymptotic information matrix $\mathscr{T}^{\mathrm{T}} \mathscr{I}(\tau) \mathscr{T}$ for all $\tau < \tau_0$. A central sequence is given by $\mathscr{T}^{\mathrm{T}} S_n(\tau)$. In particular, we have

$$\mathscr{T}^{\mathrm{T}} S_n(\tau) \xrightarrow{\mathfrak{D}}_{Q_{n,\bar{\xi}}} \mathcal{N}\big(\mathscr{T}^{\mathrm{T}} \mathscr{I}(\tau) \mathscr{T} \bar{\xi},\ \mathscr{T}^{\mathrm{T}} \mathscr{I}(\tau) \mathscr{T}\big).$$

Assume that $\lim_{t\to\tau_0} \mathscr{T}^{\mathrm{T}} \mathscr{I}(t) \mathscr{T} = \mathscr{T}^{\mathrm{T}} \mathscr{I}(\tau_0) \mathscr{T}$ exists and set $\overline{\Psi}_{n,i} = \mathscr{T}^{\mathrm{T}} \Psi_{n,i}$, $i = 1, \dots, n$, $n \in \mathbb{N}$. If the condition (2.24), where

$$V_n^{(i)}(s,\xi,\xi') = \left(\exp\!\left(\frac{1}{2\sqrt{n}} \xi^{\mathrm{T}} \overline{\Psi}_{n,i}(s) \right) - 1 \right) \left(\exp\!\left(\frac{1}{2\sqrt{n}} {\xi'}^{\mathrm{T}} \overline{\Psi}_{n,i}(s) \right) - 1 \right),$$

and the condition

$$\lim_{t\to\tau_0} \limsup_{n\to\infty} P_{n,0}\left(\left| \frac{1}{n} \sum_{i=1}^{n} \int_{(t,\tau_0)} \big(\xi^{\mathrm{T}} \overline{\Psi}_{n,i}(s)\big)^2\ \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s \right| \geq \varepsilon \right) = 0 \qquad (2.26)$$

hold for all $\xi, \xi' \in \mathbb{R}^{\bar{r}+\bar{q}}$, then the previous assertions of Theorem 2.3.10.b also hold for $\tau = \tau_0$.

Proof. We want to apply Theorem 2.3.3 to prove the assertions of b). Choose $\tau < \tau_0$. As a first step, we show that the processes

$$\overline{\Psi}_{n,i,\tau} = \big\{\overline{\Psi}_{n,i}(t \wedge \tau) \mid t \in \mathbb{R}_+\big\}, \qquad i = 1, \dots, n,$$

are predictable and locally bounded and that Assumption 2.3.1 holds. In the following we set $\overline{m} = \bar{r} + \bar{q}$ and $m = r + q$.

Using Assumption 2.3.9.viii and Assumption 2.3.9.ix, one can easily see that $\big\{\gamma^{(\dot{u},\ddot{u})}(t \wedge \tau) \mid t \in \mathbb{R}_+\big\}$, $u = 1, \dots, r$, and $\big\{\widetilde{\gamma}^{(v)}(t \wedge \tau) \mid t \in \mathbb{R}_+\big\}$, $v = 1, \dots, q$, are predictable and locally bounded processes for all $\tau < \tau_0$. Assumption 2.3.9.ii and Proposition B.5.3 yield that the processes $\big\{\Psi_{n,i}(t \wedge \tau) \mid t \in \mathbb{R}_+\big\}$, $i = 1, \dots, n$ are predictable and locally bounded for all $\tau < \tau_0$. Again, Proposition B.5.3 gives that $\big\{\overline{\Psi}_{n,i}(t \wedge \tau) \mid t \in \mathbb{R}_+\big\}$, $i = 1, \dots, n$, are predictable and locally bounded. Therefore, Assumption 2.3.1.ii holds.

Assumption 2.3.1.i is exactly Assumption 2.3.9.i. Moreover, we set

$$T = \max\Big\{\big|\mathscr{T}^{(u,v)}\big| \;\Big|\; u \in \{1, \dots, m\},\, v \in \{1, \dots, \overline{m}\}\Big\}$$

and choose $C \in \mathbb{R}_+$, such that

$$\max_{1 \le u \le r}\Big\{\sup_{t \in [0,\tau]}\big|\gamma^{(\dot{u},\ddot{u})}(t)\big|\Big\} \le C \quad \text{and} \quad \max_{1 \le u \le q}\Big\{\sup_{t \in [0,\tau]}\big|\widetilde{\gamma}^{(u)}(t)\big|\Big\} \le C. \tag{2.27}$$

Proof that Assumption 2.3.1.iii holds. We consider 3 cases. For $\bar{r} < u \le v \le \overline{m}$, it holds that

$$\overline{\Psi}^{(u)}_{n,i}(s) = \sum_{k=1}^{m} (\mathscr{T}^{\mathrm{T}})^{(u,k)} \Psi^{(k)}_{n,i}(s) = \sum_{k=r+1}^{m} \mathscr{T}^{(k,u)} \Psi^{(k)}_{n,i}(s) = \sum_{k=1}^{q} \mathscr{T}_2{}^{(k,u-\bar{r})} \Psi^{(k+r)}_{n,i}(s).$$

Using (2.27) and Assumption 2.3.9.iv we get that

$$\sup_{s\in[0,t]}\Big|\frac{1}{n}\sum_{i=1}^{n}\overline{\Psi}_{n,i}^{(u)}(s)\,\overline{\Psi}_{n,i}^{(v)}(s)\,Y_n^{(i)}(s) - \sum_{k_1,k_2=1}^{q}\mathscr{T}_2^{(k_1,u-\bar r)}\,\mathscr{T}_2^{(k_2,v-\bar r)}\widetilde{\gamma}^{(k_1)}(s)\widetilde{\gamma}^{(k_2)}(s)\,\mu_0(s)\Big| \le C^2\,T^2\,q^2\cdot\sup_{s\in[0,t]}\big|\widehat{\mu}_{n,0}(s)-\mu_0(s)\big| \longrightarrow_{P_{n,0}} 0.$$

In the case $1\le u\le \bar r < v \le \overline{m}$, we get that

$$\overline{\Psi}_{n,i}^{(u)}(s) = \sum_{k=1}^{m}(\mathscr{T}^{\mathrm{T}})^{(u,k)}\Psi_{n,i}^{(k)}(s) = \sum_{k=1}^{r}\mathscr{T}_1^{(k,u)}\Psi_{n,i}^{(k)}(s),$$

(2.27) and Assumption 2.3.9.v yield

$$\sup_{s\in[0,t]}\Big|\frac{1}{n}\sum_{i=1}^{n}\overline{\Psi}_{n,i}^{(u)}(s)\,\overline{\Psi}_{n,i}^{(v)}(s)\,Y_n^{(i)}(s) - \sum_{k_1=1}^{r}\sum_{k_2=1}^{q}\mathscr{T}_1^{(k_1,u)}\,\mathscr{T}_2^{(k_2,v)}\gamma^{(\dot k_1,\ddot k_1)}(s)\,\widetilde{\gamma}^{(k_2)}(s)\,\mu_1^{(\dot k_1)}(s)\Big| \le C^2\,T^2\,q\cdot\sum_{k_1=1}^{r}\sup_{s\in[0,t]}\big|\widehat{\mu}_{n,1}^{(\dot k_1)}(s)-\mu_1^{(\dot k_1)}(s)\big| \longrightarrow_{P_{n,0}} 0.$$

Last but not least, if $1\le u\le v\le r$, Assumption 2.3.9.vi and (2.27) give that

$$\sup_{s\in[0,t]}\Big|\frac{1}{n}\sum_{i=1}^{n}\bar{\Psi}_{n,i}^{(u)}(s)\,\bar{\Psi}_{n,i}^{(v)}(s)\,Y_n^{(i)}(s) - \sum_{k_1=1}^{r}\sum_{k_2=1}^{r}\mathscr{T}_1^{(k_1,u)}\,\mathscr{T}_1^{(k_2,v)}\gamma^{(\dot k_1,\ddot k_1)}(s)\,\gamma^{(\dot k_2,\ddot k_2)}(s)\,\mu_2^{(\dot k_1,\dot k_2)}(s)\Big| \le T^2\,C^2\cdot\sum_{k_1=1}^{r}\sum_{k_2=1}^{r}\sup_{s\in[0,t]}\big|\widehat{\mu}_{n,2}^{(\dot k_1,\dot k_2)}(s)-\mu_2^{(\dot k_1,\dot k_2)}(s)\big| \longrightarrow_{P_{n,0}} 0.$$

Proof that Assumption 2.3.1.iv holds. Note that $\sup_{i\in\{1,\dots,n\},\,s\in[0,t]}$ is abbreviated to $\sup_{i,s}$. It holds that

$$\begin{aligned}\sup_{i,s} \frac{1}{\sqrt{n}} \max_{1\le u\le \overline{m}} \big|\overline{\Psi}_{n,i}^{(u)}(s)\big| &\le \sup_{i,s} \frac{1}{\sqrt{n}} \max_{1\le u\le \bar{r}} \big|\overline{\Psi}_{n,i}^{(u)}(s)\big| + \sup_{i,s} \frac{1}{\sqrt{n}} \max_{\bar{r}< u\le \overline{m}} \big|\overline{\Psi}_{n,i}^{(u)}(s)\big| \\ &= s_n^{(1)} + s_n^{(2)}.\end{aligned}$$

Because of (2.27) we have $s_n^{(2)} \le \frac{1}{\sqrt{n}} q\; T\; C \longrightarrow 0$, as $n \to \infty$, and

$$\begin{aligned}s_n^{(1)} &= \frac{1}{\sqrt{n}} \sup_{i,s} \max_{1\le u\le \bar{r}} \Big|\sum_{k=1}^{r} \mathscr{T}_1^{(k,u)} \gamma^{(\dot{k},\ddot{k})}(s)\, Z_{n,i}^{(\dot{k})}(s)\, Y_n^{(i)}(s)\Big| \\ &\le \frac{1}{\sqrt{n}} \sup_{i,s} \max_{1\le u\le \bar{r}} \sum_{k=1}^{r} \big|\mathscr{T}_1^{(k,u)}\big|\, \big|\gamma^{(\dot{k},\ddot{k})}(s)\big|\, \big|Z_{n,i}^{(\dot{k})}(s)\, Y_n^{(i)}(s)\big| \\ &\le r\, T\, C \cdot \frac{1}{\sqrt{n}} \sup_{i,s} \max_{1\le u\le p} \big|Z_{n,i}^{(u)}(s)\, Y_n^{(i)}(s)\big| \longrightarrow_{P_{n,0}} 0,\end{aligned}$$

where we use Assumption 2.3.9.vii. The formula for the central sequence results from (2.12). The asymptotic distribution of the central sequence under alternatives is a consequence of Corollary 2.3.4. The conditions stated to extend the result to the point τ_0 are exactly the conditions (2.13) and (2.14) of Theorem 2.3.3.

Assertion a) is a special case of assertion b), choose $\mathscr{T}$ as $(m \times m)$-unity matrix. □

2.3.11 Remark. a) If the condition (2.24) in Theorem 2.3.10.a holds then (2.24) also holds with the $V_n^{(i)}(\cdot,\xi,\xi')$, $i = 1,\dots,n$, ξ,ξ', $n \in \mathbb{N}$, given in Theorem 2.3.10.b. Moreover, the condition (2.25) implies the condition (2.26).

b) By construction all events occur before τ_0 under $P_{n,0}$, since

$$\begin{aligned}P_{n,0}\big(Y_n^{(i)}(\tau_0) = 1\big) &= \lim_{t\to\tau_0} P_{n,0}\big(Y_n^{(i)}(t) = 1\big) \\ &= \lim_{t\to\tau_0} \big(1 - G_{n,i}(t)\big)\big(1 - F_0(t)\big) = 0,\end{aligned}$$

where

$$F_0(t) = \exp\Big(-\int_{I(t)} \alpha_0(s)\, \mathrm{d}s\Big) \quad \text{and} \quad G_{n,i}(t) = \exp\Big(-\int_{I(t)} \widetilde{\alpha}_n^{(i)}(s)\, \mathrm{d}s\Big),$$

and we use Fleming and Harrington [19, Theorem 1.3.1] as well as

$$\big\{Y_n^{(i)}(\tau_0) = 1\big\} = \bigcap_{k \in \mathbb{N}} \big\{Y_n^{(i)}(t_k) = 1\big\}, \qquad t_k \uparrow \tau_0.$$

Thus, it holds that $P_{n,0} = P_{n,0}^{(\tau_0)}$ and $\Upsilon_{n,\xi}(\tau_0) = \Upsilon_{n,\xi}(\infty)$ $P_{n,0}-$almost surely, see Proposition 2.2.5, and $S_n(\tau_0) = S_n(\infty)$ $P_{n,0}$-almost surely. In particular this means that asymptotic normality restricted to time τ_0 is asymptotic normality restricted to time ∞, see Definition 2.2.2.

3 Sequences of Hardest Parametric Sub-Models

The development tests for detecting a possible influence of covariates on survival times is the objective of this dissertation. In Chapter 1 it was shown that the MCRM is a reasonable mathematical description for the interaction between covariates and survival times. However, instead of looking at the MCRM we considered sequences of localized, q-dimensional parametric sub-models of the MCRM and proved asymptotic normality, see Theorem 2.3.10. These parametric sub-models depend on the choice of the number of nuisance parameters q, the nuisance directions $\widetilde{\gamma}$ and the foot-point α_0. In contrast to the choice of γ, a statistician has no indication for a sensible choice of these quantities. From this point of view studying sequences of parametric sub-models seems to be a cul-de-sac, if one wants to obtain some test that is applicable under the MCRM, since one can suspect that any reasonable testing procedure derived from some parametric sub-model should depend on the above mentioned nuisance quantities. However, this is not the case, if the underlying localized parametric sub-models are "big enough", *i.e.* the sequence of localized, q-dimensional parametric sub-models is a sequence of hardest parametric sub-models (SHPSM).

In Section 3.1 we will discuss the notion of sequences of hardest parametric sub-models which leads to an algebraic Definition of SHPSM. In Section 3.2 the properties of SHPSM are investigated. Among others, we construct a test statistic that is independent of the sequence of the underlying parametric sub-models, as long as this sequence of parametric sub-models is a SHPSM. In

Chapter 4 it turns out that this test statistic is important for the analysis of our testing problems. All in all this chapter provides a justification why it is sufficient to consider localized parametric sub-models.

3.1 Primary Remarks

In this section the notion of sequences of hardest parametric sub-models is derived. In the following we assume that $\big(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathfrak{P}_n\big)$, $n \in \mathbb{N}$, is some sequence of localized, q-dimensional parametric sub-models of the MCRM. Moreover, it is assumed that this sequence is asymptotically normal restricted to time τ, where $0 < \tau \leq \infty$, with asymptotic information matrix $\mathscr{I}(\tau)$, that $m = q + r$ and that $\mathcal{F}_{n,\infty} = \mathcal{F}_n$.

3.1.1 Discussion. The sequence of statistical experiments $\big(\Omega_n, \mathcal{F}_n, \mathfrak{P}_n^{(\tau)}\big)$, $\mathfrak{P}_n^{(\tau)} = \{P_{n,\xi}^{(\tau)} \mid \xi = (\beta^{\mathrm{T}}, \eta^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^m\}$, $n \in \mathbb{N}$, converges weakly to some Gauss shift experiment $\big(\Omega, \mathcal{A}, \mathfrak{P}^{(\tau)}\big)$, $\mathfrak{P}^{(\tau)} = \{P_{\tau,\xi} \mid \xi = (\beta^{\mathrm{T}}, \eta^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^m\}$, with central random variable $S^{\mathrm{T}}(\tau) = \big(S_1^{\mathrm{T}}(\tau)\,, S_2^{\mathrm{T}}(\tau)\big)^{\mathrm{T}} : (\Omega, \mathcal{A}) \longrightarrow (\mathbb{R}^m, \mathbb{B}^m)$, *i.e.*

$$\frac{\mathrm{d}P_{\tau,\xi}}{\mathrm{d}P_{\tau,0}} = \exp\big(S^{\mathrm{T}}\xi - \frac{1}{2}\xi^{\mathrm{T}}\mathscr{I}(\tau)\xi\big) \quad P_{\tau,0}\text{-almost surely,}$$

$\mathfrak{L}\big(S \mid P_{\tau,0}\big) = \mathcal{N}\big(0, \mathscr{I}(\tau)\big)$, cf. Strasser [68, Theorem 80.2]. For the rest of this discussion we drop the index τ, if possible. Let us consider the limit experiment. It holds that

$$P^S_{\tau,\beta,\eta} = \mathcal{N}\left(\begin{pmatrix} \mathscr{I}_{1,1} & \mathscr{I}_{1,2} \\ \mathscr{I}_{2,1} & \mathscr{I}_{2,2} \end{pmatrix}\begin{pmatrix} \beta \\ \eta \end{pmatrix}, \begin{pmatrix} \mathscr{I}_{1,1} & \mathscr{I}_{1,2} \\ \mathscr{I}_{2,1} & \mathscr{I}_{2,2} \end{pmatrix}\right).$$

For the binary testing problem $\beta = 0$, $\eta = \widetilde{\eta}$ versus $\beta = \beta_0 \neq 0$, $\eta = \eta_0$ the Neyman-Pearson test to the level $\alpha \in (0, 1)$ is given by

$$\phi(s_1, s_2) = \begin{cases} 1, & T(s_1, s_2) \; > \; c(\alpha) \\ 0, & T(s_1, s_2) \; \leq \; c(\alpha) \end{cases},$$

where

$$T(s_1, s_2) = s_1^{\mathrm{T}}\beta_0 + s_2^{\mathrm{T}}(\eta_0 - \widetilde{\eta}) - \frac{1}{2}\xi_0^{\mathrm{T}}\mathscr{I}\xi_0 + \frac{1}{2}\widetilde{\eta}^{\mathrm{T}}\mathscr{I}_{2,2}\widetilde{\eta},$$

$\xi_0^{\mathrm{T}} = (\beta_0^{\mathrm{T}}, \eta_0^{\mathrm{T}})$, and $c(\alpha)$ is chosen, such that $P^S_{\tau,0,\eta}\big(T > c(\alpha)\big) = \alpha$. For the special choice $\widetilde{\eta} = \eta_0 + \mathscr{I}_{2,2}^{-1}\mathscr{I}_{2,1}\beta_0$ the test statistic T simplifies to

$$T(s_1, s_2) = \beta_0^{\mathrm{T}}\big(s_1 - \mathscr{I}_{1,2}\mathscr{I}_{2,2}^{-1}s_2\big) - \frac{1}{2}\beta_0^{\mathrm{T}}\mathscr{I}^*\beta_0,$$

where $\mathscr{I}^* = \mathscr{I}_{1,1} - \mathscr{I}_{1,2}\mathscr{I}_{2,2}^{-1}\mathscr{I}_{2,1}$. One readily checks that the resulting test ϕ keeps the level on the composite hypothesis $\beta = 0$, $\eta \in \mathbb{R}^q$. Moreover, ϕ is almost surely determined, since it is a Neyman-Pearson test.

Let Φ_α denote the set of all tests that keep the level on the composite hypothesis $\beta = 0$, $\eta \in \mathbb{R}^q$ and let

$$e(\beta, \eta) = \sup_{\phi \in \Phi_\alpha} \int \phi \,\mathrm{d}P^S_{\tau,\beta,\eta}$$

be the envelope power function. For every $\beta \in \mathbb{R}^r$, there exists a Neyman-Pearson test ϕ^*, such that $e(\beta, \eta) = \int \phi^* \,\mathrm{d}P^S_{\tau,\beta,\eta}$ for all $\eta \in \mathbb{R}^q$. Therefore, we can say that ϕ is an efficient test for the testing problem $\beta = 0$, $\eta \in \mathbb{R}^q$ versus $\beta = \beta_0$, $\eta \in \mathbb{R}^q$ which is a sub-problem of the testing problems considered in Chapter 4. A similar discussion can also be found in Janssen and Werft [38].

Let us now consider the sequence of tests $\phi_n = \mathbb{1}\big(T(S_{n,1}, S_{n,2}) > c(\alpha)\big)$, $n \in \mathbb{N}$, where $S_n^{\mathrm{T}} = \big(S_{n,1}^{\mathrm{T}}, S_{n,2}^{\mathrm{T}}\big)$, is a central sequence. This sequence of tests is asymptotically efficient for the testing problem $\beta = 0$, $\eta \in \mathbb{R}^q$ versus $\beta = \beta_0$, $\eta \in \mathbb{R}^q$. For any other sequence of tests ψ_n, $n \in \mathbb{N}$, of asymptotic level α that is efficient for this testing problem it holds that $\psi_n - \phi_n \longrightarrow_{P^{(\tau)}_{n,\xi}} 0$, cf. Strasser [68, Theorem 63.6] (The optimal test in limit experiment is uniquely determined by its distribution, ϕ_n, and ψ_n converge in distribution to the efficient test in limit experiment in the sense of Strasser [68, Definition 62.1]). That means any asymptotically efficient testing procedure for the test problem $\beta = 0$, $\eta \in \mathbb{R}^q$ versus $\beta = \beta_0$, $\eta \in \mathbb{R}^q$ is asymptotically equivalent to a sequence of tests that depends on the random variables

$$U_n(\tau) = S_{n,1}(\tau) - \mathscr{I}_{1,2}(\tau)\mathscr{I}_{2,2}^{-1}(\tau)S_{n,2}(\tau), \qquad n \in \mathbb{N}.$$

Thus, if one aims to construct procedures that keep the level on the composite hypothesis and that possibly attain the envelope power function at some point, one can restrict oneself to tests that depend on U_n. This consideration may justify, why we concentrate only on the statistics U_n in the following. Additionally, looking at the literature on testing problems involving nuisance parameters reveals that many efficient procedures depend on the statistics U_n, $n \in \mathbb{N}$, cf. Witting and Müller-Funk [72, Section 6.4.2].

3.1.2 Lemma. In the situation of Theorem 2.3.10.b, we abbreviate $\mathscr{S}(\tau) = \mathscr{T}^{\mathrm{T}} \mathscr{J}(\tau)\mathscr{T}$ and partition the matrix $\mathscr{S}(\tau)$ as follows

$$\mathscr{S}(\tau) = \begin{pmatrix} \mathscr{S}_{1,1}(\tau) & \mathscr{S}_{1,2}(\tau) \\ \mathscr{S}_{2,1}(\tau) & \mathscr{S}_{2,2}(\tau) \end{pmatrix} = \begin{pmatrix} \mathscr{T}_1^{\mathrm{T}} \mathscr{J}_{1,1}(\tau)\mathscr{T}_1 & \mathscr{T}_1^{\mathrm{T}} \mathscr{J}_{1,2}(\tau)\mathscr{T}_2 \\ \mathscr{T}_2^{\mathrm{T}} \mathscr{J}_{2,1}(\tau)\mathscr{T}_1 & \mathscr{T}_2^{\mathrm{T}} \mathscr{J}_{2,2}(\tau)\mathscr{T}_2 \end{pmatrix}.$$

It holds that

$$\mathscr{T}_1^{\mathrm{T}} S_{n,1}(\tau) - \mathscr{S}_{1,2}(\tau)\mathscr{S}_{2,2}^{-}(\tau)\mathscr{T}_2^{\mathrm{T}} S_{n,2}(\tau) \xrightarrow{\mathfrak{D}}_{Q_{n,\bar{\xi}}} \mathcal{N}\Big(\mathscr{S}^*(\tau)\bar{\beta}, -\mathscr{S}^*(\tau)\Big),$$

where $\mathscr{S}^*(\tau) = \mathscr{S}_{1,1}(\tau) - \mathscr{S}_{1,2}(\tau)\mathscr{S}_{2,2}(\tau)^{-}\mathscr{S}_{2,1}(\tau)$ and $\mathscr{S}_{2,2}^{-}(\tau)$ denotes the generalized inverse of $\mathscr{S}_{2,2}(\tau)$, cf. Definition B.1.1.

Proof. For a moment let us drop the index τ. The Cramér-Wold device, cf. Billingsley [9, Theorem 7.7], yields that $\zeta^{\mathrm{T}}\mathscr{T}^{\mathrm{T}} S_n \xrightarrow{\mathfrak{D}}_{Q_{n,\bar{\xi}}} \mathcal{N}\big(\zeta^{\mathrm{T}}\mathscr{S}\bar{\xi}, \zeta^{\mathrm{T}}\mathscr{S}\zeta\big)$ for all $\zeta \in \mathbb{R}^{\bar{r}+\bar{q}}$. Let $\mathscr{E}_{\bar{r}}$ denote the $(\bar{r}\times\bar{r})$ unity matrix. Choosing $\zeta^{\mathrm{T}} = \rho^{\mathrm{T}}\mathscr{A}$, where $\mathscr{A} = \begin{pmatrix} \mathscr{E}_{\bar{r}} & -\mathscr{S}_{1,2}\mathscr{S}_{2,2}^{-} \end{pmatrix}$, and $\rho \in \mathbb{R}^{\bar{r}}$ gives that

$$\rho^{\mathrm{T}}\big(\mathscr{T}_1^{\mathrm{T}} S_{n,1} - \mathscr{S}_{1,2}\mathscr{S}_{2,2}^{-}\mathscr{T}_2^{\mathrm{T}} S_{n,2}\big) = \rho^{\mathrm{T}}\mathscr{A}\,\mathscr{T}^{\mathrm{T}} S_n.$$

Clearly, it holds that

$$\rho^{\mathrm{T}}\mathscr{A}\,\mathscr{T}^{\mathrm{T}} S_n \xrightarrow{\mathfrak{D}}_{Q_{n,\bar{\xi}}} \mathcal{N}\big(\rho^{\mathrm{T}}\mathscr{A}\mathscr{S}\bar{\xi},\, \rho^{\mathrm{T}}\mathscr{A}\mathscr{S}\mathscr{A}^{\mathrm{T}}\rho\big),$$

where

$$\begin{aligned}
\mathscr{A}\mathscr{S} &= \begin{pmatrix} \mathscr{E}_{\bar{r}} & -\mathscr{S}_{1,2}\mathscr{S}_{2,2}^{-} \end{pmatrix} \begin{pmatrix} \mathscr{S}_{1,1} & \mathscr{S}_{1,2} \\ \mathscr{S}_{2,1} & \mathscr{S}_{2,2} \end{pmatrix} \\
&= \begin{pmatrix} \mathscr{S}^* & \mathscr{S}_{1,2} - \mathscr{S}_{1,2}\mathscr{S}_{2,2}^{-}\mathscr{S}_{2,2} \end{pmatrix} \\
&= \begin{pmatrix} \mathscr{S}^* & 0 \end{pmatrix}
\end{aligned}$$

and $\mathscr{A}\mathscr{S}\mathscr{A}^{\mathrm{T}} = \mathscr{S}^*$. In the last but one equation, Proposition B.3.4.b and $\mathscr{S}_{2,2}^{-}\mathscr{S}_{2,2}\mathscr{S}_{2,2}^{-} = \mathscr{S}_{2,2}^{-}$, cf. Definition B.1.1, and $(\mathscr{S}_{2,2}^{-})^{\mathrm{T}} = (\mathscr{S}_{2,2}^{\mathrm{T}})^{-}$, cf. Proposition B.1.4 are used. Applying the Cramér-Wold device gives the assertion. □

In the situation of Theorem 2.3.10.a the previous result especially means that

$$U_n(\tau) = S_{n,1}(\tau) - \mathscr{I}_{1,2}(\tau)\mathscr{I}_{2,2}^{-}(\tau)S_{n,2}(\tau) \xrightarrow{\mathfrak{D}}_{P_{n,\xi}} \mathcal{N}\big(\mathscr{I}^*(\tau)\beta,\, \mathscr{I}^*(\tau)\big), \quad (3.1)$$

where $\mathscr{I}^*(\tau) = \mathscr{I}_{1,1}(\tau) - \mathscr{I}_{1,2}(\tau)\mathscr{I}_{2,2}^{-}(\tau)\mathscr{I}_{2,1}(\tau)$. The asymptotic distribution of the sequence $U_n(\tau)$ depends only on the interesting parameter β, *i.e.* under the composite hypothesis $\beta = 0$, $\eta \in \mathbb{R}^q$ the statistic $U_n(\tau)$ has, asymptotically, always the same distribution. Of course this property is fairly useful, if one aims to construct tests, since one does not need to worry about the nuisance parameter η. However, the asymptotic distribution still depends on the choice of the nuisance direction $\widetilde{\gamma}$.

In the following it is aimed to find nuisance directions that satisfy some "optimality" condition. This leads to the notion of sequences of hardest parametric sub-models. Our approach generalizes a well known idea on the construction of non-parametric test statistics, see *e.g.* Neuhaus [60].

3.1.3 Proposition. Let $\big(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathfrak{Q}_n\big)$, $n \in \mathbb{N}$, be a sequence of $(1,1)$-sub-sub-models with the transformation matrix

$$\mathscr{T}_{\beta_0,\eta_0} = \begin{pmatrix} \beta_0 & 0 \\ 0 & \eta_0 \end{pmatrix}, \qquad \beta_0 \in \mathbb{R}^r,\ \eta_0 \in \mathbb{R}^q.$$

In the situation of Theorem 2.3.10.b, it holds that

$$\beta_0^{\mathrm{T}} S_{n,1}(\tau) - \mathscr{S}_{1,2}(\tau)\mathscr{S}_{2,2}^{-}(\tau)\eta_0^{\mathrm{T}} S_{n,2}(\tau) \xrightarrow{\mathfrak{D}}_{Q_{n,\bar{\xi}}} \mathcal{N}\big(\mathscr{S}_{\beta_0,\eta_0}^*(\tau)\bar{\beta},\, \mathscr{S}_{\beta_0,\eta_0}^*(\tau)\big),$$

where $\mathscr{S}_{\beta_0,\eta_0}^*(\tau) = \mathscr{S}_{\beta_0}(\tau) - \widetilde{\mathscr{S}}_{\beta_0,\eta_0}(\tau)$,

$$\mathscr{S}_{\beta_0}(\tau) = \int_{I(\tau)} \sum_{u=1}^{r}\sum_{v=1}^{r} \beta_0^{(u)}\beta_0^{(v)}\,\gamma^{(\dot u,\ddot u)}(s)\,\gamma^{(\dot v,\ddot v)}(s)\,\mu_2^{(\dot u,\dot v)}(s)\,\alpha_0(s)\,\mathrm{d}s$$

and

$$\widetilde{\mathscr{S}}_{\beta_0,\eta_0}(\tau) = \Big(\int_{I(\tau)} \sum_{u=1}^{r}\sum_{v=1}^{q} \beta_0^{(u)}\eta_0^{(v)}\gamma^{(\dot u,\ddot u)}(s)\,\widetilde{\gamma}^{(v)}(s)\,\mu_1^{(\dot u)}(s)\,\alpha_0(s)\,\mathrm{d}s\Big)^2 \times \Big(\int_{I(\tau)} \Big(\sum_{u=1}^{q}\eta_0^{(u)}\widetilde{\gamma}^{(u)}(s)\Big)^2 \mu_0(s)\,\alpha_0(s)\,\mathrm{d}s\Big)^{-}.$$

Moreover, we have the estimate

$$0 \le \widetilde{\mathscr{S}}_{\beta_0,\eta_0}(\tau) \le \int_{I(\tau)} \Big(\sum_{u=1}^{r}\beta_0^{(u)}\gamma^{(\dot u,\ddot u)}(s)\,\mu_1^{(\dot u)}(s)\Big)^2 \frac{\alpha_0(s)}{\mu_0(s)}\,\mathrm{d}s. \tag{3.2}$$

Proof. The convergence in distribution is an easy consequence of Lemma 3.1.2. Proof of (3.2). In the case that $\widetilde{\mathscr{S}}_{\beta_0,\eta_0}(\tau) = 0$ any non-negative number is an upper bound. Thus, let us assume that $\widetilde{\mathscr{S}}_{\beta_0,\eta_0}(\tau) > 0$. Applying the Cauchy-Schwarz-inequality, cf. Gänssler and Stute [20, Satz 1.13.3], one gets

$$\begin{aligned}\widetilde{\mathscr{S}}_{\beta_0,\eta_0} &= \frac{\Big(\int_{I(\tau)}\Big(\sum_u \beta_0^{(u)}\,\gamma^{(\dot u,\ddot u)}(s)\,\mu_1^{(\dot u)}(s)\Big)\Big(\sum_v \eta_0^{(v)}\widetilde{\gamma}^{(v)}(s)\Big)\sqrt{\frac{\mu_0(s)}{\mu_0(s)}}\,\alpha_0(s)\,\mathrm{d}s\Big)^2}{\int_{I(\tau)}\Big(\sum_v \eta_0^{(v)}\widetilde{\gamma}^{(v)}(s)\Big)^2\mu_0(s)\,\alpha_0(s)\,\mathrm{d}s}\\ &\le \Big(\int_{I(\tau)}\Big(\sum\nolimits_u \beta_0^{(u)}\gamma^{(\dot u,\ddot u)}(s)\,\mu_1^{(\dot u)}(s)\Big)^2\frac{\alpha_0(s)}{\mu_0(s)}\,\mathrm{d}s\Big)\\ &\qquad\times\frac{\Big(\int_{I(\tau)}\big(\sum_v \eta_0^{(v)}\widetilde{\gamma}^{(v)}(s)\big)^2\mu_0(s)\,\alpha_0(s)\,\mathrm{d}s\Big)}{\int_{I(\tau)}\big(\sum_v \eta_0^{(v)}\widetilde{\gamma}^{(v)}(s)\big)^2\mu_0(s)\,\alpha_0(s)\,\mathrm{d}s}.\end{aligned}\tag{3.3}$$

□

3.1.4 Discussion. In the situation of Theorem 2.3.10.a, consider the sequence of (1, 1)-sub-sub-models $\big(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathfrak{Q}_n\big)$, $n \in \mathbb{N}$, with transformation matrix $\mathscr{T}_{\beta_0,\eta_0}$. Under this sequence of experiments the parameter of interest is $\bar\beta$ and the nuisance parameter is $\bar\eta$. For the limit model $\big(\Omega, \mathcal{A}, \{Q_{\bar\beta,\bar\eta} \mid \bar\beta, \bar\eta \in \mathbb{R}\}\big)$ with the central random variable $S = \big(S^{(1)}, S^{(2)}\big)$, it holds that

$$Q^S_{\bar\beta,\bar\eta} = \mathcal{N}\left(\begin{pmatrix}\beta_0^{\mathrm{T}}\mathscr{I}_{1,1}\beta_0 & \beta_0^{\mathrm{T}}\mathscr{I}_{1,2}\eta_0\\ \eta_0^{\mathrm{T}}\mathscr{I}_{2,1}\beta_0 & \eta_0^{\mathrm{T}}\mathscr{I}_{2,2}\eta_0\end{pmatrix}\begin{pmatrix}\bar\beta\\ \bar\eta\end{pmatrix}, \begin{pmatrix}\beta_0^{\mathrm{T}}\mathscr{I}_{1,1}\beta_0 & \beta_0^{\mathrm{T}}\mathscr{I}_{1,2}\eta_0\\ \eta_0^{\mathrm{T}}\mathscr{I}_{2,1}\beta_0 & \eta_0^{\mathrm{T}}\mathscr{I}_{2,2}\eta_0\end{pmatrix}\right),$$

where we dropped the index τ.

Now, let us investigate the testing problem $\bar{\beta} = 0$, $\bar{\eta} \in \mathbb{R}$ versus $\bar{\beta} > 0$, $\bar{\eta} \in \mathbb{R}$. The uniformly best, unbiased test to the level α for this testing problem is

$$\phi(S^{(1)}, S^{(2)}) = \begin{cases} 1, \\ 0, \end{cases} S^{(1)} - \beta_0^{\mathrm{T}} \mathscr{I}_{1,2}\eta_0 \, (\eta_0^{\mathrm{T}} \mathscr{I}_{2,2}\eta_0)^{-} S^{(2)} \begin{matrix} > \\ \leq \end{matrix} c(\alpha) \; ,$$

where $Q_{0,0}\big(S^{(1)} - \beta_0^{\mathrm{T}} \mathscr{I}_{1,2}\eta_0 \, (\eta_0^{\mathrm{T}} \mathscr{I}_{2,2}\eta_0)^{-} S^{(2)} > c(\alpha)\big) = \alpha$, cf. Witting and Müller-Funk [72, Satz .6.184] . The power function of this test is given by

$$\Phi\Big(-u_\alpha + \bar{\beta} \cdot (\mathscr{S}^*_{\beta_0,\eta_0})^{1/2}\Big), \qquad \bar{\beta} \geq 0, \; \bar{\eta} \in \mathbb{R}$$

where Φ is the distribution function of a normal distribution with mean 0 and variance 1, and $u_\alpha = \Phi^{-1}\big(1-\alpha\big)$. The factor $\mathscr{S}^*_{\beta_0,\eta_0}$ determines the capability of the uniformly best, unbiased test to detect any fixed alternative, *i.e.* the smaller this factor, the less powerful the test. Therefore, we have to minimize $\mathscr{S}^*_{\beta_0,\eta_0}$ for finding a hardest model. Since we considered a fixed $(1,1)$-sub-sub-model, β_0 and η_0 cannot be subject to a minimization, consequently, minimizing $\mathscr{S}^*_{\beta_0,\eta_0}$ is equivalent to maximizing $\widetilde{\mathscr{S}}_{\beta_0,\eta_0}$ with respect to $\widetilde{\gamma}$. Note that $\widetilde{\mathscr{S}}_{\beta_0,\eta_0}$ is bounded, see Proposition 3.1.3. One easily checks that the upper bound of $\widetilde{\mathscr{S}}_{\beta_0,\eta_0}$, cf. equation (3.2), is attained, if

$$\sum_{v=1}^{q} \eta_0^{(v)} \widetilde{\gamma}^{(v)}(s) = \frac{c_0}{\mu_0(s)} \sum_{u=1}^{r} \beta_0^{(u)} \gamma^{(\dot{u},\ddot{u})}(s) \, \mu_1^{(\dot{u})}(s) \qquad \Lambda_0\text{-almost everywhere}$$

for some $c_0 \in \mathbb{R}$, $\Lambda_0^{(\tau_0)}(B) = \int_{B \cap I(\tau)} \alpha_0(s) \, \mathrm{d}s$, $B \in \mathbb{B}$. A similar discussion can be found in Neuhaus [60].

The previous discussion provides an idea how to generalize the notion of a hardest model to multi-dimensional models. Basically, it is asked that in the inequality (3.3) equality holds.

3.1.5 Definition (Sequence of Hardest Parametric Sub-Models). Let us consider a sequence of localized, q-dimensional parametric sub-models of the MCRM $(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathfrak{P}_n)$, $n \in \mathbb{N}$, that is asymptotically normal restricted to time τ with information matrix $\mathscr{J}(\tau)$. We say $(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathfrak{P}_n)$, $n \in \mathbb{N}$, is a sequence of hardest parametric sub-models restricted to time τ, if for every $\beta_0 \in \mathbb{R}^r$ there exists an $\eta_0 \in \mathbb{R}^q$, such that

$$\left(\beta_0^{\mathrm{T}} \mathscr{J}_{1,2}(\tau)\eta_0\right)^2 = \int_{I(\tau)} \left(\sum\nolimits_{u=1}^{r} \beta_0^{(u)} \gamma^{(\dot{u},\ddot{u})}(s)\, \mu_1^{(\dot{u})}(s)\right)^2 \frac{\alpha_0(s)}{\mu_0(s)}\, \mathrm{d}s \cdot \eta_0^{\mathrm{T}} \mathscr{J}_{2,2}(\tau)\eta_0 \tag{3.4}$$

and

$$\int_{I(\tau)} \Big(\sum_{u=1}^{r} \beta_0^{(u)} \gamma^{(\dot{u},\ddot{u})}(s)\, \mu_1^{(\dot{u})}(s)\Big)^2 \frac{\alpha_0(s)}{\mu_0(s)}\, \mathrm{d}s = 0 \iff \eta_0^{\mathrm{T}} \mathscr{J}_{2,2}(\tau)\eta_0 = 0. \tag{3.5}$$

η_0 is called a hardest nuisance parameter with respect to β_0.

3.1.6 Remark. The Definition 3.1.5 is based on the idea that for every fixed direction of the interesting parameter β_0 there should be a nuisance parameter η_0, such that the limit model of the sequence $(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathfrak{Q}_n)$, $n \in \mathbb{N}$, of $(1,1)$-sub-sub-models with transformation matrix $\mathscr{T}_{\beta_0,\eta_0}$ is a hardest model in the sense of Discussion 3.1.4. More precisely, this means that for every β_0 there exists an η_0, such that $\widetilde{\mathscr{I}}_{\beta_0,\eta_0}$ attains the right hand side of (3.2).

Condition (3.5) prevents the trivial solution of (3.4), namely $\eta_0 = 0$. Otherwise every sequence of parametric sub-model would be a sequence of hardest parametric models. Clearly, a hardest nuisance parameter is by no means unique, since if η_0 is a hardest nuisance direction then $c \cdot \eta_0$, $c \in \mathbb{R} \backslash \{0\}$, is also a hardest nuisance direction.

3.2 Properties of Sequences of Hardest Parametric Sub-models

In the previous section the notion of SHPSM was established. In this section the properties of such sequences of models are investigated. First of all, the existence of SHPSM is discussed. One easily sees that starting from a given

SHPSM, new sequences of hardest parametric sub-models can be generated by reparametrization and adding new nuisance direction. Considering this fact, one should regard (3.4) and (3.5) in Definition 3.1.5 as conditions guaranteeing that a sequence of localized, q-dimensional parametric sub-models is “big enough” to reflect the properties of the MCRM. However, sequences of hardest parametric sub-models share one important property: the statistic $U_n(\tau)$ defined in equation (3.1) is independent of the underlying SHPSM. We replace the asymptotic quantities of the statistic $U_n(\tau)$ by consistent estimators and prove asymptotic equivalence of these statistics. Last but not least, a weakly consistent variance estimator is introduced. But first a crucial premise for proving the main results of this chapter.

3.2.1 Assumption. Set $\tau_0^c = \sup\{s \mid \mu_0(s) > 0\}$. It is assumed that

$$P_{n,0}\big(Y_n^{(i)}(\tau_0^c) = 1\big) = 0 \quad \text{for all } i = 1, \ldots, n \text{ and } n \in \mathbb{N}.$$

3.2.2 Remark. a) Obviously, it holds that $\tau_0^c \leq \tau_0$, see Assumption 2.3.9.i. In particular, all observed survival times are smaller than τ_0^c.

b) In the case $\tau_0^c < \tau_0$, we do not have any information on the distribution of the survival times in the limit model after τ_0^c due to the right censoring. Assumption 3.2.1 guarantees that the limit model is a reasonable approximation of the models for finite $n \in \mathbb{N}$.

c) Remark 2.3.11.b still holds, if one replaces τ_0 by τ_0^c. Especially, asymptotic normality restricted to time τ_0^c implies asymptotic normality restricted to time ∞, see Definition 2.2.2. In the case that $\tau_0^c < \tau_0$, the conditions (2.24) and (2.25) trivially holds.

3.2.3 Theorem (Existence of SHPSM). Consider the sequence of localized, r-dimensional parametric sub-models $\big(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathfrak{P}_n\big)$, $n \in \mathbb{N}$, where the nuisance directions are given by

$$\widetilde{\gamma}^{(v)}(t) = \gamma^{(\dot{v},\ddot{v})}(t)\frac{\mu_1^{(\dot{v})}(t)}{\mu_0(t)}, \quad v = 1, \ldots, r.$$

Under Assumption 2.2.1, Assumption 2.3.9.i – Assumption 2.3.9.viii and Assumption 3.2.1, the sequence of parametric sub-models $(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{P}_n)$, $n \in \mathbb{N}$, is asymptotically normal restricted to time τ for all $\tau < \tau_0^c$. If additionally the conditions (2.24) and (2.25) hold with τ_0 replaced by τ_0^c, then the result extends to $\tau = \tau_0^c$.

Moreover, the sequence of localized parametric sub-models $(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{P}_n)$, $n \in \mathbb{N}$, is a sequence of hardest parametric sub-models in the sense of Definition 3.1.5. We call it the canonical sequence of hardest parametric sub-models.

Proof. Assumption 2.3.9.i – Assumption 2.3.9.viii imply Assumption 2.3.9.ix, therefore asymptotic normality is an immediate consequence of Theorem 2.3.10. For given β_0 choose $\eta_0 = \beta_0$. Using the basic Definition of the asymptotic information matrix $\mathscr{I}$, see Theorem 2.3.10.a, it holds that

$$\left(\beta_0^{\mathrm{T}} \mathscr{I}_{1,2} \eta_0\right)^2 = \left(\int_{I(\tau)} \left(\sum\nolimits_{u=1}^{r} \beta_0^{(u)} \gamma^{(\dot{u},\ddot{u})}(s)\, \mu_1^{(\dot{u})}(s)\right)^2 \frac{\alpha_0(s)}{\mu_0(s)}\, \mathrm{d}s\right)^2$$

and

$$\eta_0^{\mathrm{T}} \mathscr{I}_{2,2} \eta_0 = \int_{I(\tau)} \left(\sum\nolimits_{u=1}^{r} \beta_0^{(u)} \gamma^{(\dot{u},\ddot{u})}(s)\, \mu_1^{(\dot{u})}(s)\right)^2 \frac{\alpha_0(s)}{\mu_0(s)}\, \mathrm{d}s.$$

Consequently, the conditions (3.4) and (3.5) hold. □

3.2.4 Theorem. Under Assumption 2.2.1, Assumption 2.3.9.i – Assumption 2.3.9.viii and Assumption 3.2.1, let $(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{P}_n)$, $n \in \mathbb{N}$, be a sequence of localized, q-dimensional sub-models of the MCRM that is asymptotically normal restricted to time τ' with asymptotic information matrix $\mathscr{I}(\tau')$ and central sequence $S_n(\tau')$. Moreover, assume that $(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{P}_n)$, $n \in \mathbb{N}$, is a SHPSM in the sense of Definition 3.1.5.

a) The matrices $\mathscr{I}_{1,1}^{\mathrm{can}}(\tau') \in \mathbb{R}^{r \times r}$ and $\mathscr{I}_{2,2}^{\mathrm{can}}(\tau') \in \mathbb{R}^{r \times r}$, where

$$\mathscr{I}_{1,1}^{\mathrm{can}\,(u,v)}(\tau') = \int_{I(\tau')} \gamma^{(\dot{u},\ddot{u})}(s)\, \gamma^{(\dot{v},\ddot{v})}(s)\, \mu_2^{(\dot{u},\dot{v})}(s)\, \alpha_0(s)\, \mathrm{d}s$$

and

$$\mathscr{I}_{2,2}^{\mathrm{can}\,(u,v)}(\tau') = \int_{I(\tau')} \gamma^{(\dot u,\ddot u)}(s)\,\gamma^{(\dot v,\ddot v)}(s)\,\mu_1^{(\dot u)}(s)\,\mu_1^{(\dot v)}(s)\,\frac{\alpha_0(s)}{\mu_0(s)}\,\mathrm{d}s$$

are well defined. (Only in the case $\tau' = \tau_0^c$ is this not obvious.)

b) It holds that $\mathscr{I}_{1,2}(t)\mathscr{I}_{2,2}^{-}(t)\mathscr{I}_{2,1}(t) = \mathscr{I}_{2,2}^{\mathrm{can}}(t)$, $t \in I(\tau')$.

c) For the statistic $U_n(t) = S_{n,1}(t) - \mathscr{I}_{1,2}(t)\mathscr{I}_{2,2}(t)^{-}S_{n,2}(t)$ it holds that

$$U_n^{(u)}(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_{I(t)} \gamma^{(\dot u,\ddot u)}(s)\left(Z_{n,i}^{(\dot u)}(s) - \frac{\mu_1^{(\dot u)}(s)}{\mu_0(s)}\right)\mathrm{d}M_{n,0}^{(i)}(s) \tag{3.6}$$

$P_{n,0}$-almost surely, $u = 1,\ldots,r$, for all $t \le \tau'$. Moreover, we have

$$U_n(\tau') \xrightarrow{\mathfrak{D}}_{P_{n,\xi}} \mathcal{N}\big(\mathscr{I}^{*,\mathrm{can}}(\tau')\beta,\, \mathscr{I}^{*,\mathrm{can}}(\tau')\big),$$

where $\mathscr{I}^{*,\mathrm{can}}(\tau') = \mathscr{I}^{*}(\tau') = \mathscr{I}_{1,1}^{\mathrm{can}}(\tau') - \mathscr{I}_{2,2}^{\mathrm{can}}(\tau')$. $\mathscr{I}^{*,\mathrm{can}}(\tau')$ is called (asymptotic) information matrix of the MCRM and its components are given by

$$\int_{I(\tau')} \gamma^{(\dot u,\ddot u)}(s)\,\gamma^{(\dot v,\ddot v)}(s)\left(\mu_2^{(\dot u,\dot v)}(s) - \frac{\mu_1^{(\dot u)}(s)\mu_1^{(\dot v)}(s)}{\mu_0(s)}\right)\alpha_0(s)\,\mathrm{d}s,$$

$u, v = 1,\ldots,r$.

For the proof of the Theorem the following well-known results are needed.

3.2.5 Lemma. Let f_i, $i = 1,2$, be real-valued functions. If $0 < \left(\int f_1\cdot f_2\,\mathrm{d}\nu\right)^2 = \int f_1^2\,\mathrm{d}\nu \cdot \int f_2^2\,\mathrm{d}\nu < \infty$ then it holds that $f_1 = c_0\, f_2$ ν-almost surely for some $c_0 \in \mathbb{R}\backslash\{0\}$.

Proof. The functions f_i, $i = 1,2$, are obviously square-integrable. Set

$$a = \sqrt{\int f_2^2\,\mathrm{d}\nu} \quad\text{and}\quad b = \sqrt{\int f_1^2\,\mathrm{d}\nu}\cdot\begin{cases} 1, & \\ -1, & \end{cases} \text{if } \int f_1\cdot f_2\,\mathrm{d}\nu \begin{matrix} < \\ > \end{matrix} 0\,.$$

Note that $a, b \neq 0$. Since the equation

$$\begin{aligned}\int \big(a \cdot f_1 + b \cdot f_2\big)^2 \,\mathrm{d}\nu &= 2 \int f_1^2 \,\mathrm{d}\nu \cdot \int f_2^2 \,\mathrm{d}\nu + 2ab \int f_1 \cdot f_2 \,\mathrm{d}\nu \\ &= 2 \int f_1^2 \,\mathrm{d}\nu \cdot \int f_2^2 \,\mathrm{d}\nu - 2\sqrt{\int f_1^2 \,\mathrm{d}\nu \cdot \int f_2^2 \,\mathrm{d}\nu} \cdot \left| \int f_1 \cdot f_2 \,\mathrm{d}\nu \right| \\ &= (2-2) \cdot \int f_1^2 \,\mathrm{d}\nu \cdot \int_2^2 \mathrm{d}\nu = 0\end{aligned}$$

implies $a\, f = b\, g$ ν-almost surely, Lemma 3.2.5 holds. □

3.2.6 Lemma. Let $\phi_j : \mathbb{R}_+ \to \mathbb{R}$, $j = 1, 2$, be some measurable functions that are bounded on bounded intervals. Suppose that Assumption 2.2.1 holds. If $\phi_1 = \phi_2$ $\Lambda_0^{(\tau)}$-almost surely, $\Lambda_0^{(\tau)}(B) = \int_B \mathbb{1}_{I(\tau)}(s)\, \alpha_0(s) \,\mathrm{d}s$, $B \in \mathbb{B}$, then it holds that

$$\sum_{i=1}^n \int_{I(\tau)} \phi_1(s) \,\mathrm{d}M_{n,0}^{(i)}(s) = \sum_{i=1}^n \int_{I(\tau)} \phi_2(s) \,\mathrm{d}M_{n,0}^{(i)}(s) \quad P_{n,0}\text{-almost surely.}$$

Proof. Using the abbreviation $X^{(j)} = \sum_{i=1}^n \int_{I(\tau)} \phi_j(s) \,\mathrm{d}M_{n,0}^{(i)}(s)$, $j = 1, 2$, it holds that

$$\big\{X^{(1)} \neq X^{(2)}\big\} = \bigcup_{k=1}^{\infty} \big\{|X^{(1)} - X^{(2)}| \geq 1/k\big\},$$

where $\big\{|X^{(1)} - X^{(2)}| \geq 1/k\big\} \subset \big\{|X^{(1)} - X^{(2)}| \geq 1/(k+1)\big\}$. Corollary 2.1.7 implies

$$\begin{aligned}&P_{n,0}\big\{|X^{(1)} - X^{(2)}| \geq 1/k\big\} \leq \\ &\qquad k^2\varepsilon + P_{n,0}\left(\sum_{i=1}^n \int_{I(\tau)} \big(\phi_1(s) - \phi_2(s)\big)^2\, Y_n^{(i)}(s)\, \alpha_0(s) \,\mathrm{d}s \geq \varepsilon \right) = k^2\varepsilon.\end{aligned}$$

As $\varepsilon > 0$ was arbitrary, $\varepsilon \downarrow 0$ gives $P_{n,0}\big\{\big|X^{(1)} - X^{(2)}\big| \geq 1/k\big\} = 0$. We have $P_{n,0}\big(\big\{X^{(1)} \neq X^{(2)}\big\}\big) = \lim_{k\to\infty} P_{n,0}\big\{\big|X^{(1)} - X^{(2)}\big| \geq 1/k\big\} = 0$, where we use an elementary property of measures. Thus, the assertion holds. □

Proof of Theorem 3.2.4. First, note that we can assume that $\tau' \leq \tau_0^c$ because of Remark 3.2.2. Let us start with a). In the case $\tau' < \tau_0^c$ the assertion is trivial, since we assumed that all functions are bounded and μ_0 is bounded away from 0. In the case $\tau' = \tau_0^c$ this argument does not apply. However, it holds that $\mathscr{I}_{1,1}(\tau_0^c) = \mathscr{I}_{1,1}^{\text{can}}(\tau_0^c)$, so that we only have to worry about $\mathscr{I}_{2,2}^{\text{can}}(\tau_0^c)$. Set $\beta_w = (\delta_{1,w}, \ldots, \delta_{r,w})^{\mathrm{T}}$, $w = 1, \ldots, r$, where $\delta_{i,w}$ is the Kronecker symbol, and let η_w denote a hardest nuisance parameter with respect to β_w.

Using the fact that a SHPSM is considered, it holds that $\beta_w^{\mathrm{T}} \mathscr{I}_{2,2}^{\text{can}}(\tau_0^c)\beta_w \cdot \eta_w^{\mathrm{T}} \mathscr{I}_{2,2}(\tau_0^c)\eta_w = \big(\beta_w^{\mathrm{T}} \mathscr{I}_{1,2}(\tau_0^c)\eta_w\big)^2 < \infty$ implying

$$\mathscr{I}_{2,2}^{\text{can}\,(w,w)}(\tau_0^c) = \int_{I(\tau')} \left(\gamma^{(\dot{w},\ddot{w})}(s)\,\frac{\mu_1^{(\dot{w})}(s)}{\mu_0(s)}\right)^2 \mu_0(s)\,\alpha_0(s)\,\mathrm{d}s < \infty.$$

Consequently, $\gamma^{(\dot{u},\ddot{u})}\,\mu_1^{(\dot{u})}/\mu_0$ is a square-integrable functions with respect to $\widetilde{\Lambda}_0^{(\tau_0^c)}(B) = \int_B \mathbb{1}_{I(\tau_0^c)}(s)\,\mu_0(s)\,\alpha_0(s)\,\mathrm{d}s$, $B \in \mathbb{B}$. Thus,

$$\mathscr{I}_{2,2}^{\text{can}\,(u,v)}(\tau') = \int_{I(\tau')} \gamma^{(\dot{u},\ddot{u})}(s)\,\gamma^{(\dot{v},\ddot{v})}(s)\,\frac{\mu_1^{(\dot{u})}(s)\,\mu_1^{(\dot{v})}(s)}{\mu_0^2(s)}\,\mu_0(s)\,\alpha_0(s)\,\mathrm{d}s$$

exists and is well defined.

Proof of b). Let us drop the Index τ' for a moment. Assume that $\beta \in \ker(\mathscr{I}_{1,2}^{\mathrm{T}})$. As a first step we show that $\beta \in \ker(\mathscr{I}_{2,2}^{\text{can}})$. Using the properties of SHPSM we can find a hardest nuisance parameter η, such that

$$0 = \big(\beta^{\mathrm{T}}\,\mathscr{I}_{1,2}\eta\big)^2 = \beta^{\mathrm{T}}\,\mathscr{I}_{2,2}^{\text{can}}\beta \cdot \eta^{\mathrm{T}}\,\mathscr{I}_{2,2}\eta$$

and $\beta^{\mathrm{T}}\,\mathscr{I}_{2,2}^{\text{can}}\beta = 0$. Set $\widetilde{\eta} = \mathscr{I}_{2,2}^{\text{can}}\beta$. Applying the Cauchy-Schwarz inequality, cf. (3.3), gives

$$0 \leq \big(\beta^{\mathrm{T}}\,\mathscr{I}_{2,2}^{\text{can}}\,\mathscr{I}_{2,2}^{\text{can}}\beta\big)^2 = \big(\beta^{\mathrm{T}}\,\mathscr{I}_{2,2}^{\text{can}}\widetilde{\eta}\big)^2 \leq \beta^{\mathrm{T}}\,\mathscr{I}_{2,2}^{\text{can}}\beta \cdot \widetilde{\eta}^{\mathrm{T}}\,\mathscr{I}_{2,2}^{\text{can}}\widetilde{\eta} = 0.$$

As a second step we consider $\beta \in \ker(\mathscr{I}_{1,2}^{\mathrm{T}})^{\perp}$. Let η denote a corresponding hardest nuisance parameter. First we show that $(\beta^{\mathrm{T}}\,\mathscr{I}_{1,2}\eta)^2 > 0$. It holds that

$$0 < \beta^{\mathrm{T}}\,\mathscr{I}_{1,2}\,\mathscr{I}_{2,1}\beta \leq \beta^{\mathrm{T}}\,\mathscr{I}_{2,2}^{\text{can}}\beta \cdot \sum_{v=1}^{q} \int_{I(\tau')} \big(\widetilde{\gamma}^{(v)}(s)\big)^2\,\mu_0(s)\,\alpha_0(s)\,\mathrm{d}s,$$

where the Cauchy-Schwarz-inequality is used, again. Consequently, we get that $\beta^{\mathrm{T}} \mathscr{I}_{2,2}^{\mathrm{can}} \beta > 0$ and $\eta^{\mathrm{T}} \mathscr{I}_{2,2} \eta > 0$, see equation (3.5). Now, equation (3.4) is given by

$$(\beta^{\mathrm{T}} \mathscr{I}_{1,2} \eta)^2 = \beta^{\mathrm{T}} \mathscr{I}_{2,2}^{\mathrm{can}} \beta \cdot \eta^{\mathrm{T}} \mathscr{I}_{2,2} \eta > 0.$$

Lemma 3.2.5 yields

$$c \sum_{v=1}^{q} \eta^{(v)} \widetilde{\gamma}^{(v)} = \sum_{u=1}^{r} \beta^{(u)} \gamma^{(\dot{u},\ddot{u})} \frac{\mu_1^{(\dot{u})}}{\mu_0} \quad \widetilde{\Lambda}_0^{(\tau')}\text{-almost surely.} \tag{3.7}$$

In particular, equation (3.7) also holds $\Lambda_0^{(\tau')}$-almost surely. Because of (3.7) and the properties of the generalized inverse, see Definition B.1.1, it holds that

$$\beta^{\mathrm{T}} \mathscr{I}_{1,2} \mathscr{I}_{2,2}^{-} \mathscr{I}_{2,1} \beta = c^2\, \eta^{\mathrm{T}} \mathscr{I}_{2,2} \mathscr{I}_{2,2}^{-} \mathscr{I}_{2,2} \eta = c^2\, \eta^{\mathrm{T}} \mathscr{I}_{2,2} \eta = \beta^{\mathrm{T}} \mathscr{I}_{2,2}^{\mathrm{can}} \beta.$$

Putting the previous results together for all $\beta \in \ker(\mathscr{I}_{1,2}^{\mathrm{T}})^{\perp}$ and $\widetilde{\beta} \in \ker(\mathscr{I}_{1,2}^{\mathrm{T}})$ it holds that

$$\begin{aligned}(\beta + \widetilde{\beta})^{\mathrm{T}} \mathscr{I}_{1,2} \mathscr{I}_{2,2}^{-} \mathscr{I}_{2,1} (\beta + \widetilde{\beta}) &= \beta^{\mathrm{T}} \mathscr{I}_{1,2} \mathscr{I}_{2,2}^{-} \mathscr{I}_{2,1} \beta \\ &= \beta^{\mathrm{T}} \mathscr{I}_{2,2}^{\mathrm{can}} \beta = (\beta + \widetilde{\beta})^{\mathrm{T}} \mathscr{I}_{2,2}^{\mathrm{can}} (\beta + \widetilde{\beta}).\end{aligned}$$

Thus, $\beta^{\mathrm{T}} \mathscr{I}_{1,2} \mathscr{I}_{2,2}^{-} \mathscr{I}_{2,1} \beta = \beta^{\mathrm{T}} \mathscr{I}_{2,2}^{\mathrm{can}} \beta$ for all $\beta \in \mathbb{R}^r$, since the matrices are symmetric, the result is our assertion.

Proof of c). Lemma 3.1.2 and b) give the assertion concerning the convergence in distribution. It remains to show (3.6). It holds that

$$\gamma^{(\dot{w},\ddot{w})} \frac{\mu_1^{(\dot{w})}}{\mu_0} = \sum_{u=1}^{r} \beta_w^{(u)} \gamma^{(\dot{u},\ddot{u})} \frac{\mu_1^{(\dot{u})}}{\mu_0} = c^{(w)} \sum_{v=1}^{q} \eta_w^{(v)} \widetilde{\gamma}^{(v)} \tag{3.8}$$

$\widetilde{\Lambda}_0^{(\tau')}$-almost surely, where $c^{(w)} \neq 0$, $w = 1, \ldots, r$. This equality in an immediate consequence of Definition 3.1.5. Either, we have $0 = (\beta_w^{\mathrm{T}} \mathscr{I}_{1,2} \eta_w)^2 = \beta_w^{\mathrm{T}} \mathscr{I}_{2,2}^{\mathrm{can}} \beta_w \cdot \eta_w^{\mathrm{T}} \mathscr{I}_{2,2} \eta_w$, which implies $\beta_w^{\mathrm{T}} \mathscr{I}_{2,2}^{\mathrm{can}} \beta_w = 0$ and $\eta_w^{\mathrm{T}} \mathscr{I}_{2,2} \eta_w = 0$ and consequently

$$\sum_{u=1}^{r} \beta_w^{(u)} \gamma^{(\dot{u},\ddot{u})} \frac{\mu_1^{(\dot{u})}}{\mu_0} = 0 = \sum_{v=1}^{q} \eta_w^{(v)} \widetilde{\gamma}^{(v)} \quad \widetilde{\Lambda}_0^{(\tau')}\text{-almost surely}$$

(in this case set $c^{(w)} = 1$), or we have

$$0 < (\beta_w^{\mathrm{T}} \mathscr{I}_{1,2} \eta_w)^2 = \beta_w^{\mathrm{T}} \mathscr{I}_{2,2}^{\mathrm{can}} \beta_w \cdot \eta_w^{\mathrm{T}} \mathscr{I}_{2,2} \eta_w$$

then Lemma 3.2.5 gives (3.8), see also (3.7). Equation (3.8) implies that $\mathscr{I}_{1,2}(t) = c^{\mathrm{T}} \mathscr{T}^{\mathrm{T}} \mathscr{I}_{2,2}(t)$, $t \leq \tau'$, where $c = (c^{(w)} \mid w = 1, \dots, r)^{\mathrm{T}}$ and $\mathscr{T} = \big(\eta_1, \dots, \eta_r\big) \in \mathbb{R}^{q \times r}$. Because of

$$\mathscr{I}_{1,2}(t) \mathscr{I}_{2,2}^{-}(t) S_{n,2}(t) = c^{\mathrm{T}} \mathscr{T}^{\mathrm{T}} \mathscr{I}_{2,2}(t) \mathscr{I}_{2,2}^{-}(t) S_{n,2}(t) \tag{3.9}$$

we show as a first step that $\mathscr{I}_{2,2}(t)\big(\mathscr{I}_{2,2}(t)\big)^{-} S_{n,2}(t) = S_{n,2}(t)$ $P_{n,0}$-almost surely for all $t \leq \tau'$. We consider three cases. If $\mathscr{I}_{2,2}(t)$ has full rank, the assertion is trivial. Assume that $\operatorname{rank}\big(\mathscr{I}_{2,2}(t)\big) = 0$ implying $\mathscr{I}_{2,2}(t) = 0$ which gives $\widetilde{\gamma}^{(v)} = 0$ $\widetilde{\Lambda}_0^{(t \wedge \tau')}$-almost surely and also almost surely with respect to the measure $\Lambda_0^{(t \wedge \tau')}$, where Assumption 3.2.1 is used. We get that $S_{n,2}(t) = S_{n,2}(t \wedge \tau') = 0$ $P_{n,0}$-almost surely, where once more Assumption 3.2.1 and Lemma 3.2.6 are used.

Last but not least suppose that $\operatorname{rank}(\mathscr{I}_{2,2}(t)) = k$, $0 < k < q$. Therefore, we can assume that we have $\mathcal{M}_1 = \{v_1, \dots, v_k\} \subset \{1, \dots, q\}$, such that the vectors $\big(\mathscr{I}_{2,2}^{(u,v_l)}(t) \mid u = 1, \dots, q\big)^{\mathrm{T}}$, $l = 1, \dots, k$ are generating the column space of $\mathscr{I}_{2,2}(t)$. Hence, for every $u \in \mathcal{M}_2 = \{1, \dots, q\} \backslash \mathcal{M}_1$ there exists a vector $c_u \in \mathbb{R}^k$, such that

$$\begin{aligned} \int_{I(t)} \big(\widetilde{\gamma}^{(u)}(s)\big)^2 \mu_0(s)\, \alpha_0(s)\, \mathrm{d}s &= \int_{I(t)} \widetilde{\gamma}^{(u)}(s) \left(\sum_{l=1}^{k} c_u^{(l)} \widetilde{\gamma}^{(v_l)}(s) \right) \mu_0(s)\, \alpha_0(s)\, \mathrm{d}s \\ &= \int_{I(t)} \left(\sum_{l=1}^{k} c_u^{(l)} \widetilde{\gamma}^{(v_l)}(s) \right)^2 \mu_0(s)\, \alpha_0(s)\, \mathrm{d}s. \end{aligned} \tag{3.10}$$

Because of (3.10) we have for all $u \in \mathcal{M}_2$

$$\int_{I(t)} \left(\widetilde{\gamma}^{(u)}(s) - \sum_{l=1}^{k} c_u^{(l)} \widetilde{\gamma}^{(v_l)}(s) \right)^2 \mu_0(s)\, \alpha_0(s)\, \mathrm{d}s = 0$$

implying $\widetilde{\gamma}^{(u)} = \sum_{l=1}^{k} c_u^{(l)} \widetilde{\gamma}^{(v_l)}$ $\Lambda_0^{(t\wedge\tau')}$-almost surely, where once again the Assumption 3.2.1 is used. Using Assumption 3.2.1 and applying Lemma 3.2.6 yields that

$$S_{n,2}^{(u)}(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{I(t)} \sum_{l=1}^{k} c_u^{(l)} \widetilde{\gamma}^{(v_l)}(s)\,\mathrm{d}M_{n,0}^{(i)}(s) = \sum_{l=1}^{k} c_u^{(l)} S_{n,2}^{(v_l)}(t),$$

$P_{n,0}$-almost surely, so we can conclude that the rank of the extended matrix $\big(\mathscr{I}_{2,2}(t) \mid S_{n,2}(t)\big)$ is k $P_{n,0}$-almost surely, where we use that the matrix $\mathscr{I}_{2,2}(t)$ is symmetric. Proposition B.1.5 implies

$$\mathscr{I}_{2,2}(t)\big(\mathscr{I}_{2,2}(t)\big)^{-} S_{n,2}(t) = S_{n,2}(t) \qquad P_{n,0}\text{-almost surely,}$$

completing the assertion.

By now, we have shown that $\mathscr{I}_{1,2}(t)\mathscr{I}_{2,2}^{-}(t)S_{n,2}(t) = c^{\mathrm{T}}\mathscr{T}^{\mathrm{T}}S_{n,2}(t)$ $P_{n,0}$-almost surely, see equation (4.18). Applying (3.8), Assumption 3.2.1 and Lemma 3.2.6 gives

$$c^{\mathrm{T}}\mathscr{T}^{\mathrm{T}}S_{n,2}(t) = \frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} \int_{I(t)} \gamma^{(\dot{u},\ddot{u})}(s)\,\frac{\mu_1^{(\dot{u})}(s)}{\mu_0(s)}\,\mathrm{d}M_{n,0}^{(i)}(s)\;\middle|\; u = 1,\dots,r\right)^{\mathrm{T}}$$

$P_{n,0}$-almost surely, all in all equation (3.6) holds for all $t \le \tau'$. □

The notation $\mathscr{I}_{2,2}^{\mathrm{can}}$ is supposed to remind that this matrix coincides with the matrix $\mathscr{I}_{2,2}$ of the canonical hardest model. The previous result gives that the statistics $U_n(\tau')$, $n \in \mathbb{N}$, are independent of the underlying SHPSM. However, the statistics $U_n(\tau')$, $n \in \mathbb{N}$, still depend on some asymptotic quantities and the foot-point α_0. In the next step we will replace the asymptotic quantities by some consistent estimators and show asymptotic equivalence of the sequence of the new statistics with $U_n(\tau)$, $n \in \mathbb{N}$.

3.2.7 Lemma. Under Assumption 3.2.1 and Assumption 2.3.9.i – Assumption 2.3.9.v, it holds that

$$\sup_{t\in[0,\tau]} \left| \frac{\widehat{\mu}_{n,1}^{(u)}(t)}{\widehat{\mu}_{n,0}(t)} - \frac{\mu_1^{(u)}(t)}{\mu_0(t)} \right| \longrightarrow_{P_{n,0}} 0, \qquad u = 1,\dots,p$$

for all $\tau < \tau_0^c$.

Proof. First, note that one can choose $0 < 2\delta < \inf_{t\in[0,\tau]} \mu_0(t)$ exploiting Assumption 3.2.1. For all $\varepsilon > 0$, we have $A_n \subset (A_n \cap B_n) \cup B_n^{\complement}$ where we set

$$A_n = \Big\{ \sup_{t\in[0,\tau]} \Big| \frac{\widehat{\mu}_{n,1}^{(u)}(t)}{\widehat{\mu}_{n,0}(t)} - \frac{\mu_1^{(u)}(t)}{\mu_0(t)} \Big| \geq \varepsilon \Big\} \text{ and } B_n = \Big\{ \sup_{t\in[0,\tau]} \big| \widehat{\mu}_{n,0}(t) - \mu_0(t) \big| \leq \delta \Big\}.$$

Because of Assumption 2.3.9.iv we have $P_{n,0}(B_n^{\complement}) \to 0$, so there remains to be shown that $P_{n,0}(A_n \cap B_n) \to 0$. On the set B_n it holds the estimate

$$\begin{aligned} \sup_{t\in[0,\tau]} \Big| \frac{\widehat{\mu}_{n,1}^{(u)}(t)}{\widehat{\mu}_{n,0}(t)} - \frac{\mu_1^{(u)}(t)}{\mu_0(t)} \Big| &\leq \frac{1}{\delta^2} \sup_{t\in[0,\tau]} \big| \widehat{\mu}_{n,1}^{(u)}(t)\, \mu_0(t) - \mu_1^{(u)}(t)\, \widehat{\mu}_{n,0}(t) \big| \\ &\leq \frac{c_0}{\delta^2} \Big(\sup_{t\in[0,\tau]} \big| \widehat{\mu}_{n,1}^{(u)}(t) - \mu_1^{(u)}(t) \big| + \sup_{t\in[0,\tau]} \big| \widehat{\mu}_{n,0}(t) - \mu_0(t) \big| \Big) \longrightarrow_{P_{n,0}} 0, \end{aligned}$$

where we set $c_0 = \max\big\{ \sup_{t\in[0,\tau]} |\mu_1^{(u)}(t)|, \sup_{t\in[0,\tau]} \mu_0(t) \big\}$ and use Assumption 2.3.9.iv and 2.3.9.v. Consequently, $P_{n,0}(A_n \cap B_n) \to 0$. $\square$

We assumed that the weight functions are of the form $\gamma^{(\dot{u},\ddot{u})} = \gamma_0^{(\dot{u},\ddot{u})} \circ H$, where $\gamma_0^{(\dot{u},\ddot{u})} : [0,1] \longrightarrow \mathbb{R}$, $u = 1, \ldots, r$, are measurable functions and H is some cumulative distribution function on $\mathbb{R}_+$. Possible choices for the cumulative distribution function are $H = 1 - \mu_0$ or

$$H(\cdot) = \int_{[0,\cdot]} \alpha_0(t)\, \exp\Big(- \int_{[0,t]} \alpha_0(s)\, \mathrm{d}s \Big)\, \mathrm{d}t.$$

As already mentioned, see Remark 1.3.3.c, one can easily determine early or late differences on the interval $[0,1]$ and choose an appropriate weight function $\gamma_0^{(\dot{u},\ddot{u})}$. The distribution function H guarantees the right scaling with respect to the underlying probability measure $P_{n,0}$. However, H has to be estimated cf. Remark 4.4.3. This is reflected by the following premises, where we consider a more general situation. In Chapter 5 $\widehat{\gamma}_n^{(\dot{u},\ddot{u})} = \gamma_0^{(\dot{u},\ddot{u})} \circ \widehat{H}_n$ is used as an estimator, where $\widehat{H}_n$ is either $1 - \widehat{\mu}_{n,0}$ or the left continuous version of the Kaplan-Meyer estimator.

3.2.8 Assumption. Let $\big\{ \widehat{\gamma}_n^{(\dot{u},\ddot{u})}(t) \mid t \in [0,\infty) \big\}$, $u = 1, \ldots, r$, be predictable and locally bounded processes. Moreover, assume that these processes satisfy the condition

$$\left(\int_{I(\tau)} \big(\widehat{\gamma}_n^{(\dot{u},\ddot{u})}(s) - \gamma^{(\dot{u},\ddot{u})}(s)\big)^2 \mu_0(s)\,\alpha_0(s)\,\mathrm{d}s\right) \longrightarrow_{P_{n,0}} 0$$

for all $\tau < \tau_0^c$.

The statistics defined in the following Theorem merely depend on observable quantities. This property is very important for the application. Especially, we can replace $M_{n,0}^{(i)}$ by $N_n^{(i)}$, see Remark 3.2.10.a.

3.2.9 Theorem. Define the statistic $\overline{U}_n(\tau) = \big(\overline{U}_n^{(u)}(\tau) \mid u = 1, \dots, r\big)^{\mathrm{T}}$, where

$$\overline{U}_n^{(u)}(\tau) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{I(\tau)} \gamma^{(\dot{u},\ddot{u})}(s) \cdot \left(Z_{n,i}^{(\dot{u})}(s) - \frac{\widehat{\mu}_{n,1}^{(\dot{u})}(s)}{\widehat{\mu}_{n,0}(s)} \right) \mathrm{d}M_{n,0}^{(i)}(s)$$

and the statistic $\widehat{U}_n(\tau) = \big(\widehat{U}_n^{(u)}(\tau) \mid u = 1, \dots, r\big)^{\mathrm{T}}$, where

$$\widehat{U}_n^{(u)}(\tau) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{I(\tau)} \widehat{\gamma}_n^{(\dot{u},\ddot{u})}(s) \cdot \left(Z_{n,i}^{(\dot{u})}(s) - \frac{\widehat{\mu}_{n,1}^{(\dot{u})}(s)}{\widehat{\mu}_{n,0}(s)} \right) \mathrm{d}M_{n,0}^{(i)}(s).$$

a) In the situation of Theorem 3.2.4, it holds that $\overline{U}_n(\tau) - U_n(\tau) \longrightarrow_{P_{n,0}} 0$, for all $\tau \in \mathbb{R}_+$, such that $\tau \le \tau'$ and $\tau < \tau_0^c$. If $\tau' = \tau_0^c$ and additionally the condition

$$\lim_{t \to \tau_0^c} \limsup_{n\to\infty} P_{n,0}\left(\int_{(t,\tau_0^c)} \big(\gamma^{(\dot{u},\ddot{u})}\big)^2 \left(\frac{\mu_1^{(\dot{u})}}{\mu_0} - \frac{\widehat{\mu}_{n,1}^{(\dot{u})}}{\widehat{\mu}_{n,0}} \right)^2 \widehat{\mu}_{n,0}\,\alpha_0 \mathrm{d}s \ge \varepsilon \right) = 0 \tag{3.11}$$

for all $\varepsilon > 0$ and $u = 1, \dots, r$ is satisfied, the convergence in probability also holds for the case $\tau = \tau_0^c$.

b) In the situation of Theorem 3.2.4 and under Assumptions 3.2.8, it holds that $\widehat{U}_n(\tau) - U_n(\tau) \longrightarrow_{P_{n,0}} 0$ for all $\tau \le \tau'$ and $\tau < \tau_0^c$. If $\tau' = \tau_0^c$ and additionally the conditions (3.11),

$$\lim_{t \to \tau_0^c} \limsup_{n\to\infty} P_{n,0}\left(\int_{(t,\tau_0^c)} \big(\gamma^{(\dot{u},\ddot{u})}\big)^2 \left(\widehat{\mu}_{n,2}^{(\dot{u},\dot{u})} - \frac{(\widehat{\mu}_{n,1}^{(\dot{u})})^2}{\widehat{\mu}_{n,0}} \right) \alpha_0 \,\mathrm{d}s \ge \varepsilon \right) = 0 \tag{3.12}$$

for all $\varepsilon > 0$, $u = 1, \ldots, r$, and

$$\lim_{t \to \tau_0^c} \limsup_{n \to \infty} P_{n,0}\left(\int_{(t,\tau_0^c)} \big(\widehat{\gamma}^{(\dot{u},\ddot{u})}\big)^2 \left(\widehat{\mu}_{n,2}^{(\dot{u},\dot{u})} - \frac{(\widehat{\mu}_{n,1}^{(\dot{u})})^2}{\widehat{\mu}_{n,0}} \right) \alpha_0 \, \mathrm{d}s \geq \varepsilon \right) = 0 \tag{3.13}$$

for all $\varepsilon > 0$, $u = 1, \ldots, r$, are satisfied then the convergence in probability also holds for the case $\tau = \tau_0^c$.

Proof. We start with a). Assume that $\tau \leq \tau'$ and $\tau < \tau_0^c$. According to Proposition B.4.5 and Lemma 2.1.8 it suffices to show

$$R_n^{(u)}(\tau) = \int_{I(\tau)} \big(\gamma^{(\dot{u},\ddot{u})}(s)\big)^2 \left(\frac{\mu_1^{(\dot{u})}(s)}{\mu_0(s)} - \frac{\widehat{\mu}_{n,1}^{(\dot{u})}(s)}{\widehat{\mu}_{n,0}(s)} \right)^2 \widehat{\mu}_{n,0}(s)\, \alpha_0(s) \, \mathrm{d}s \longrightarrow_{P_{n,0}} 0$$

for $u = 1, \ldots, r$, where we use Theorem 3.2.4.c. As we have the estimate

$$\begin{aligned} R_n^{(u)}(\tau) \leq \left(\sup_{s \in I(\tau)} \left| \frac{\mu_1^{(\dot{u})}(s)}{\mu_0(s)} - \frac{\widehat{\mu}_{n,1}^{(\dot{u})}(s)}{\widehat{\mu}_{n,0}(s)} \right| \right)^2 \Bigg(& \int_{I(\tau)} \big(\gamma^{(\dot{u},\ddot{u})}(s)\big)^2 \mu_0(s)\, \alpha_0(s) \, \mathrm{d}s \\ & + \sup_{s \in I(\tau)} \big| \mu_0(s) - \widehat{\mu}_{n,0}(s) \big| \int_{I(\tau)} \big(\gamma^{(\dot{u},\ddot{u})}(s)\big)^2 \alpha_0(s) \, \mathrm{d}s \Bigg), \end{aligned}$$

Lemma 3.2.7 as well as Assumption 2.3.9.i, Assumption 2.3.9.iv and Assumption 2.3.9.viii yield the result that $R_n^{(u)}(\tau) \longrightarrow_{P_{n,0}} 0$.

Now, we show the extension to $\tau' = \tau_0^c$. Let τ_k, $k \in \mathbb{N}$, be a sequence of real numbers, such that $\tau_k < \tau_0^c$ and $\tau_k \uparrow \tau_0^c$, as $k \to \infty$. Set

$$\begin{aligned} X_{n,k}^{(u)} &= \overline{U}_n^{(u)}(\tau_k) - U_n^{(u)}(\tau_k) \\ V_n^{(u)} &= \overline{U}_n^{(u)}(\tau_0^c) - U_n^{(u)}(\tau_0^c). \end{aligned}$$

As we have $X_{n,k}^{(u)} \longrightarrow_{P_{n,0}} 0$ for all $k \in \mathbb{N}$, Theorem 2.1.1 gives that $V_n^{(u)} \longrightarrow_{P_{n,0}} 0$ is implied by

$$\lim_{k \to \infty} \limsup_{n \to \infty} P_{n,0}\big(\big| V_n^{(u)} - X_{n,k}^{(u)} \big| \geq \varepsilon \big) = 0, \quad u = 1, \ldots, r, \tag{3.14}$$

for all $\varepsilon > 0$. Note that convergence in distribution to some constant implies convergence in probability to that constant, cf. Witting and Müller-Funk [72, Hilfssatz 5.82]. For any $\delta > 0$ choose η, such that $\eta/\varepsilon^2 < \delta$. Corollary 2.1.7 gives that

$$P_{n,0}\big(\big|V_n^{(u)} - X_{n,k}^{(u)}\big| \geq \varepsilon\big) \leq \frac{\eta}{\varepsilon^2} + \\ P_{n,0}\left(\int_{(\tau_k,\tau_0^c)} \big(\gamma^{(\dot u,\ddot u)}(s)\big)^2 \left(\frac{\mu_1^{(\dot u)}(s)}{\mu_0(s)} - \frac{\widehat{\mu}_{n,1}^{(\dot u)}(s)}{\widehat{\mu}_{n,0}(s)}\right)^2 \widehat{\mu}_{n,0}(s)\,\alpha_0(s)\,\mathrm{d}s \geq \eta\right).$$

Because of (3.11) we get, for all sufficiently large k, that

$$\limsup_{n\to\infty} P_{n,0}\left(\int_{(\tau_k,\tau_0^c)} \big(\gamma^{(\dot u,\ddot u)}\big)^2 \left(\frac{\mu_1^{(\dot u)}}{\mu_0} - \frac{\widehat{\mu}_{n,1}^{(\dot u)}}{\widehat{\mu}_{n,0}}\right)^2 \widehat{\mu}_{n,0}\,\alpha_0\,\mathrm{d}s \geq \eta\right) < \delta - \frac{\eta}{\varepsilon^2}.$$

Consequently, $\limsup_{n\to\infty} P_{n,0}\big(|V_n^{(u)} - X_{n,k}^{(u)}| \geq \varepsilon\big) \leq \delta$ for all sufficiently large k, *i.e.* (3.14) holds.

Proof of b). Assume that $\tau \leq \tau'$ and $\tau < \tau_0^c$. We show $\widehat{U}_n^{(u)}(\tau) - \overline{U}_n^{(u)}(\tau) \longrightarrow_{P_{n,0}} 0$. According to Proposition B.4.5 and Lemma 2.1.8 this is implied by

$$\int_{I(\tau)} \big(\widehat{\gamma}_n^{(\dot u,\ddot u)} - \gamma^{(\dot u,\ddot u)}\big)^2 \left(\widehat{\mu}_{n,2}^{(\dot u,\dot u)} - \frac{(\widehat{\mu}_{n,1}^{(\dot u)})^2}{\widehat{\mu}_{n,0}}\right) \alpha_0\,\mathrm{d}s \longrightarrow_{P_{n,0}} 0. \tag{3.15}$$

First, we show that

$$\sup_{s\in I(\tau)} \left| \widehat{\mu}_{n,2}^{(\dot u,\ddot u)} - \frac{(\widehat{\mu}_{n,1}^{(\dot u)})^2}{\widehat{\mu}_{n,0}\mu_0} - \left(\mu_2^{(\dot u,\dot u)} - \frac{(\mu_1^{(\dot u)})^2}{\mu_0}\right)\right| \longrightarrow_{P_{n,0}} 0.$$

Set $\delta_0 = \inf_{s\in I(\tau)} \mu_0(s)$ and $\kappa_0 = \sup_{s\in I(\tau)} \big|\mu_1^{(\dot u)}(s)\big|$. Because of Assumption 2.3.9 we have $\delta_0 > 0$, $\kappa_0 < \infty$ and $\sup_{s\in I(\tau)} \big|\widehat{\mu}_{n,2}^{(\dot u,\ddot u)}(s) - \mu_2^{(\dot u,\dot u)}(s)\big| \longrightarrow_{P_{n,0}} 0$. Moreover, one easily sees that

$$\sup_{s\in I(\tau)} \left| \frac{\big(\widehat{\mu}_{n,1}^{(\dot u)}(s)\big)^2}{\widehat{\mu}_{n,0}(s)} - \frac{\big(\mu_1^{(\dot u)}(s)\big)^2}{\mu_0(s)}\right| \leq \frac{\kappa_0}{\delta_0} \sup_{s\in I(\tau)} \big|\widehat{\mu}_{n,1}^{(\dot u)}(s) - \mu_1^{(\dot u)}(s)\big| \\ + \sup_{s\in I(\tau)} \left|\frac{\widehat{\mu}_{n,1}^{(\dot u)}(s)}{\widehat{\mu}_{n,0}(s)} - \frac{\mu_1^{(\dot u)}(s)}{\mu_0(s)}\right| \left(\sup_{s\in I(\tau)} \big|\widehat{\mu}_{n,1}^{(\dot u)}(s) - \mu_1^{(\dot u)}(s)\big| + \kappa_0\right) \longrightarrow_{P_{n,0}} 0,$$

where Lemma 3.2.7 and Assumption 2.3.9.v are used. This result and Assumption 3.2.8 yield (3.15), since

$$\begin{aligned}
&\left|\int_{I(\tau)} \left(\widehat{\gamma}_n^{(\dot u,\ddot u)} - \gamma^{(\dot u,\ddot u)}\right)^2 \left(\widehat{\mu}_{n,2}^{(\dot u,\dot u)} - \frac{(\widehat{\mu}_{n,1}^{(\dot u)})^2}{\widehat{\mu}_{n,0}}\right) \alpha_0 \,\mathrm{d}s\right| \leq \\
&\frac{\kappa_1}{\delta_0}\int_{I(\tau)} \left(\widehat{\gamma}_n^{(\dot u,\ddot u)} - \gamma^{(\dot u,\ddot u)}\right)^2 \mu_0\,\alpha_0 \,\mathrm{d}s + \frac{1}{\delta_0}\int_{I(\tau)} \left(\widehat{\gamma}_n^{(\dot u,\ddot u)} - \gamma^{(\dot u,\ddot u)}\right)^2 \mu_0\,\alpha_0 \,\mathrm{d}s \\
&\qquad\qquad \times \sup_{s\in I(\tau)}\left|\frac{\widehat{\mu}_{n,2}^{(\dot u,\ddot u)}}{\mu_0} - \frac{(\widehat{\mu}_{n,1}^{(\dot u)})^2}{\widehat{\mu}_{n,0}\mu_0} - \left(\frac{\mu_2^{(\dot u,\dot u)}}{\mu_0} - \frac{(\mu_1^{(\dot u)})^2}{\mu_0^2}\right)\right| \longrightarrow_{P_{n,0}} 0,
\end{aligned}$$

where $\kappa_1 = \sup_{s\in I(\tau)}\left|\mu_2^{(\dot u,\dot u)}(s)\right| + \kappa_0^2/\delta_0$.

The proof of the case $\tau = \tau' = \tau_0^c$ is based on the same idea as the proof of a). Let τ_k, $k \in \mathbb{N}$, be a sequence of real numbers, such that $\tau_k < \tau_0^c$ and $\tau_k \uparrow \tau_0^c$, as $k \to \infty$. Set

$$\begin{aligned}
X_{n,k}^{(u)} &= \widehat{U}_n^{(u)}(\tau_k) - \overline{U}_n^{(u)}(\tau_k)\\
V_n^{(u)} &= \widehat{U}_n^{(u)}(\tau_0^c) - \overline{U}_n^{(u)}(\tau_0^c).
\end{aligned}$$

As we have $X_{n,k}^{(u)} \longrightarrow_{P_{n,0}} 0$ for all $k \in \mathbb{N}$, Theorem 2.1.1 gives that $V_n^{(u)} \longrightarrow_{P_{n,0}} 0$ is implied by equation (3.14).

$$\begin{aligned}
P_{n,0}\big(|V_n^{(u)} - X_{n,k}^{(u)}| \geq \varepsilon\big) \leq P_{n,0}&\Big(\big|\widehat{U}_n^{(u)}(\tau_0^c) - \widehat{U}_n^{(u)}(\tau_k)\big| \geq \varepsilon/2\Big)\\
&+ P_{n,0}\Big(\big|\overline{U}_n^{(u)}(\tau_0^c) - \overline{U}_n^{(u)}(\tau_k)\big| \geq \varepsilon/2\Big).
\end{aligned}$$

Completely analogously to the proof of a), one shows that for all $\delta > 0$ the conditions (3.12) and (3.13) yield that

$$P_{n,0}\Big(\big|\overline{U}_n^{(u)}(\tau_0^c) - \overline{U}_n^{(u)}(\tau_k)\big| \geq \varepsilon/2\Big) < \delta$$

and

$$P_{n,0}\Big(\big|\widehat{U}_n^{(u)}(\tau_0^c) - \widehat{U}_n^{(u)}(\tau_k)\big| \geq \varepsilon/2\Big) < \delta$$

for all sufficiently large k. Thus, equation (3.14) holds. □

3.2.10 Remark. a) One easily sees that condition (3.11) can be replaced by condition (3.12) and

$$\lim_{t\to\tau_0^c}\limsup_{n\to\infty} P_{n,0}\left(\int_{(t,\tau_0^c)}\left(\gamma^{(\dot{u},\ddot{u})}\right)^2\frac{1}{n}\sum_{i=1}^n\left(Z_{n,i}^{(\dot{u})}-\frac{\mu_1^{(\dot{u})}}{\mu_0}\right)^2 Y_n^{(i)}\,\alpha_0\,\mathrm{d}s\geq\varepsilon\right)=0 \tag{3.16}$$

for all $\varepsilon>0$ and $u=1,\dots,r$.

b) Consider the test statistics $\overline{U}_n(\tau)$. It holds that

$$\overline{U}_n^{(u)}(\tau)=\frac{1}{\sqrt{n}}\sum_{i=1}^n\int_{I(\tau)}\gamma^{(\dot{u},\ddot{u})}(s)\left(Z_{n,i}^{(\dot{u})}(s)-\frac{\widehat{\mu}_{n,1}^{(\dot{u})}(s)}{\widehat{\mu}_{n,0}(s)}\right)\mathrm{d}N_n^{(i)}(s),$$

$u=1,\dots,r$, because of

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^n\int_{I(\tau)}&\gamma^{(\dot{u},\ddot{u})}(s)\left(Z_{n,i}^{(\dot{u})}(s)-\frac{\widehat{\mu}_{n,1}^{(\dot{u})}(s)}{\widehat{\mu}_{n,0}(s)}\right)Y_n^{(i)}\,\alpha_0(s)\,\mathrm{d}s\\&=\sqrt{n}\int_{I(\tau)}\gamma^{(\dot{u},\ddot{u})}(s)\left(\widehat{\mu}_{n,1}^{(\dot{u})}(s)-\frac{\widehat{\mu}_{n,1}^{(\dot{u})}(s)}{\widehat{\mu}_{n,0}(s)}\widehat{\mu}_{n,0}(s)\right)\alpha_0(s)\,\mathrm{d}s=0.\end{aligned}$$

Analogously one shows that

$$\widehat{U}_n^{(u)}(\tau)=\frac{1}{\sqrt{n}}\sum_{i=1}^n\int_{I(\tau)}\widehat{\gamma}_n^{(\dot{u},\ddot{u})}(s)\left(Z_{n,i}^{(\dot{u})}(s)-\frac{\widehat{\mu}_{n,1}^{(\dot{u})}(s)}{\widehat{\mu}_{n,0}(s)}\right)\mathrm{d}N_n^{(i)}(s),$$

$u=1,\dots,r$. The statistics $\overline{U}_n$ and $\widehat{U}_n$ do not depend on the foot-point α_0 and nuisance directions $\widetilde{\gamma}$ of our parametric sub-model. This means that the statistics $\overline{U}_n$ and $\widehat{U}_n$ are independent of the underlying sequence of parametric sub-models as long as it is a SHPSM. Obviously, these statistics are promising candidates for the derivation of a testing procedure for the MCRM.

c) Calculating Cox partial likelihood for the MCRM (without localization), cf. Definition 1.3.2, gives

$$L_n(\beta)=\prod_{i=1}^n p_{n,i}(X_{n,i},\beta)^{\Delta_{n,i}},$$

where

$$p_{n,i}\bigl(t,\beta\bigr) = \frac{Y_n^{(i)}(t)\ \exp\Bigl(\beta^{\mathrm{T}} Z_{n,i} \circledcirc \gamma(t)\Bigr)}{\sum_{i=1}^{n} Y_n^{(i)}(t)\ \exp\Bigl(\beta^{\mathrm{T}} Z_{n,i} \circledcirc \gamma(t)\Bigr)},$$

$\Delta_{n,i} = \sup\bigl\{t \mid N_n^{(i)}(t) = 0\bigr\}$ and $X_{n,i} = \sup\bigl\{t \mid N_n^{(i)}(t) + \widetilde{N}_n^{(i)}(t) = 0\bigr\}$, see Andersen *et al.* [4, Example VII.2.1]. Normally, one uses the partial likelihood for inference on β by solving the score equations

$$\frac{\partial L_n\bigl(\beta\bigr)}{\partial \beta^{(u)}} = 0, \qquad u = 1, \dots, r.$$

The Wald, the likelihood ratio and score statistic depend on the solution of the score equations $\widehat{\beta}_n$, see Andersen *et al.* [4, pp. 486] or Klein and Moeschberger [43, Section 8.5]. In this thesis a different approach was used. However, the statistic $\overline{U}_n$ is also connected with Cox partial likelihood. More precisely, it holds that

$$\frac{\partial L_n\bigl(\beta/\sqrt{n}\bigr)}{\partial \beta^{(u)}}\Big|\,\beta = 0 = \overline{U}_n^{(u)}(\tau_0), \qquad u = 1, \dots, r,$$

cf. Andersen *et al.* [4, Equation 7.2.16]. This is not too surprising as Peto remarks in the Discussion on a paper of Cox [13] that in certain cases efficient rank test procedures depend on the statistic $\partial L_n\bigl(\beta\bigr)/\partial\beta^{(u)}{}_{|\,\beta = 0}$. This will also be seen in Chapter 4.

However, in order to construct tests we need an estimator for the asymptotic covariance matrix. Unfortunately, we need stricter assumptions on the weight functions to prove the consistency of the variance estimator.

3.2.11 Assumption. Let $\bigl\{\widehat{\gamma}_n^{(\dot{u},\ddot{u})}(t) \mid t \in [0,\infty)\bigr\}$, $u = 1, \dots, r$, be predictable and locally bounded processes. Moreover, assume that these processes satisfy the condition

$$\sup_{s \in I(\tau)} \Bigl|\widehat{\gamma}_n^{(\dot{u},\ddot{u})}(s) - \gamma^{(\dot{u},\ddot{u})}(s)\Bigr| \longrightarrow_{P_{n,0}} 0.$$

for all $\tau < \tau_0^c$.

3.2.12 Remark. Assumption 3.2.11 implies Assumption 3.2.8, since

$$\left(\int_{I(\tau)} \left(\widehat{\gamma}_n^{(\dot{u},\ddot{u})}(s) - \gamma^{(\dot{u},\ddot{u})}(s)\right)^2 \mu_0(s)\,\alpha_0(s)\,\mathrm{d}s\right) \leq \left(\sup_{s\in I(\tau)} \left|\widehat{\gamma}_n^{(\dot{u},\ddot{u})}(s) - \gamma^{(\dot{u},\ddot{u})}(s)\right|\right)^2 \cdot \int_{I(\tau)} \mu_0(s)\,\alpha_0(s)\,\mathrm{d}s \longrightarrow_{P_{n,0}} 0.$$

3.2.13 Theorem. Let us agree that

$$\widehat{f}_n^{(u,v)}(s) = \widehat{\gamma}_n^{(\dot{u},\ddot{u})}(s)\,\widehat{\gamma}_n^{(\dot{v},\ddot{v})}(s)\left(\widehat{\mu}_{n,2}^{(\dot{u},\dot{v})}(s) - \frac{\widehat{\mu}_{n,1}^{(\dot{u})}(s)\,\widehat{\mu}_{n,1}^{(\dot{v})}(s)}{\widehat{\mu}_{n,0}(s)}\right), \quad s \in \mathbb{R}_+,$$

$u, v = 1, \ldots, r$. Under Assumption 2.2.1, Assumption 2.3.9 and Assumption 3.2.11, $\widehat{V}_n(\tau) = \left(\widehat{V}_n^{(u,v)}(\tau) \mid u, v = 1, \ldots, r\right)$, where

$$\widehat{V}_n^{(u,v)}(\tau) = \frac{1}{n}\sum_{i=1}^{n} \int_{I(\tau)} \frac{\widehat{f}_n^{(u,v)}(s)}{\widehat{\mu}_{n,0}(s)}\,\mathrm{d}N_n^{(i)}(s),$$

is a consistent estimator of the asymptotic information matrix $\mathscr{J}^{*,\mathrm{can}}(\tau)$ for all $\tau < \tau_0^c$. If additionally $\lim_{t\to\tau_0^c} \mathscr{J}^{*,\mathrm{can}}(t) = \mathscr{J}^{*,\mathrm{can}}(\tau_0^c)$ exists and the condition

$$\lim_{t\to\tau_0^c} \limsup_{n\to\infty} P_{n,0}\left(\sum_{u=1}^{r}\sum_{v=1}^{r} c^{(u)}c^{(v)} \int_{(t,\tau_0^c)} \widehat{f}_n^{(u,v)}(s)\,\alpha_0(s)\,\mathrm{d}s > \varepsilon\right) = 0 \qquad (3.17)$$

for all $\varepsilon > 0$ and $c \in \mathbb{R}^r$ holds, then the estimator is also consistent for $\tau = \tau_0^c$.

Proof. We merely need to show that

$$\widehat{V}_n^{(u,v)} \longrightarrow_{P_{n,0}} \int_{I(\tau)} \gamma^{(\dot{u},\ddot{u})}(s)\,\gamma^{(\dot{v},\ddot{v})}(s)\left(\mu_2^{(\dot{u},\dot{v})}(s) - \frac{\mu_1^{(\dot{u})}(s)\,\mu_1^{(\dot{v})}(s)}{\mu_0(s)}\right)\alpha_0(s)\,\mathrm{d}s$$

for all $u, v = 1, \ldots, r$, where we observe that

$$\gamma^{(\dot{u},\ddot{u})}(s)\,\gamma^{(\dot{v},\ddot{v})}(s)\left(\mu_2^{(\dot{u},\dot{v})}(s) - \frac{\mu_1^{(\dot{u})}(s)\,\mu_1^{(\dot{v})}(s)}{\mu_0(s)}\right), \quad s \in I(\tau),$$

$u = 1, \dots, p$, are bounded functions. Clearly, $\widehat{\mu}_{n,0}(s) = 0$ implies $\widehat{f}_n^{(u,v)}(s) = 0$. It holds that

$$\begin{aligned} \widehat{V}_n^{(u,v)}(\tau) - \int_{I(\tau)} \widehat{f}_n^{(u,v)}(s) \frac{\widehat{\mu}_{n,0}(s)}{\widehat{\mu}_{n,0}(s)} \alpha_0(s) \, \mathrm{d}s \\ = \widehat{V}_n^{(u,v)}(\tau) - \frac{1}{n} \sum_{i=1}^n \int_{I(\tau)} \frac{\widehat{f}_n^{(u,v)}(s)}{\widehat{\mu}_{n,0}(s)} Y_n^{(i)}(s) \, \alpha_0(s) \, \mathrm{d}s \\ = \frac{1}{n} \sum_{i=1}^n \int_{I(\tau)} \frac{\widehat{f}_n^{(u,v)}(s)}{\widehat{\mu}_{n,0}(s)} \, \mathrm{d}M_{n,0}^{(i)}(s). \end{aligned} \tag{3.18}$$

Note that the process $\{\widehat{f}_n^{(u,v)}(s \wedge \tau)/\widehat{\mu}_{n,0}(s \wedge \tau) \mid s \in \mathbb{R}_+\}$ is predictable and locally bounded. We show that the right hand side of (3.18) converges to 0 in $P_{n,0}$-probability, as $n \to \infty$. According to Lemma 2.1.8 this is implied by

$$\frac{1}{n^2} \sum_{i=1}^n \int_{I(\tau)} \left(\frac{\widehat{f}_n^{(u,v)}(s)}{\widehat{\mu}_{n,0}(s)} \right)^2 Y_n^{(i)}(s) \, \alpha_0(s) \, \mathrm{d}s \longrightarrow_{P_{n,0}} 0.$$

It holds the estimate

$$0 \leq \frac{1}{n^2} \sum_{i=1}^n \int_{I(\tau)} \left(\frac{\widehat{f}_n^{(u,v)}}{\widehat{\mu}_{n,0}} \right)^2 Y_n^{(i)} \, \alpha_0 \, \mathrm{d}s \leq \frac{2}{n} \int_{I(\tau)} g_n \, \alpha_0 \, \mathrm{d}s,$$

where we set

$$g_n(s) = \left(\widehat{\gamma}_n^{(\dot{u},\ddot{u})}(s) \, \widehat{\gamma}_n^{(\dot{v},\ddot{v})}(s) \right)^2 \left(\frac{\left(\widehat{\mu}_{n,2}^{(\dot{u},\dot{v})}(s)\right)^2}{\widehat{\mu}_{n,0}(s)} + \frac{\left(\widehat{\mu}_{n,1}^{(\dot{u})}(s) \, \widehat{\mu}_{n,1}^{(\dot{v})}(s)\right)^2}{\left(\widehat{\mu}_{n,0}(s)\right)^3} \right).$$

Exploiting Assumption 3.2.1 one can choose $0 < 2\delta < \{\inf_{t \in I(\tau)} \mu_0(t)\}$. Hence, for all $\varepsilon > 0$ we have

$$\begin{aligned} &\left\{ \frac{2}{n} \int_{I(\tau)} g_n(s) \, \alpha_0(s) \, \mathrm{d}s \geq \varepsilon \right\} \\ &\subset \left(\left\{ \frac{2}{n} \int_{I(\tau)} g_n(s) \, \alpha_0(s) \, \mathrm{d}s \geq \varepsilon \right\} \cap \left\{ \sup_{s \in I(\tau)} \left| \widehat{\mu}_{n,0}(s) - \mu_0(s) \right| \leq \delta \right\} \right) \\ &\qquad \cup \left\{ \sup_{s \in I(\tau)} \left| \widehat{\mu}_{n,0}(s) - \mu_0(s) \right| \geq \delta \right\} = A_n \cup B_n. \end{aligned}$$

It remains to be proved that $P_{n,0}(A_n) \longrightarrow 0$ and $P_{n,0}(B_n) \longrightarrow 0$, as $n \to \infty$. The latter is obvious because of Assumption 2.3.9.iv.

Since $\widehat{\mu}_{n,0} \geq \delta$ is implied by $\sup_{s \in I(\tau)} \left|\widehat{\mu}_{n,0}(s) - \mu_0(s)\right| \leq \delta$ on the set $B_n^{\complement}$, we have the estimate

$$\begin{aligned} &\frac{2}{n} \int_{I(\tau)} g_n \, \alpha_0 \, \mathrm{d}s, \leq \\ &\qquad \frac{2}{n\delta^3} \int_{I(\tau)} \left(\widehat{\gamma}_n^{(\dot{u},\ddot{u})} \, \widehat{\gamma}_n^{(\dot{v},\ddot{v})}\right)^2 \left(\left(\widehat{\mu}_{n,2}^{(\dot{u},\dot{v})}\right)^2 + \left(\widehat{\mu}_{n,1}^{(\dot{u})} \, \widehat{\mu}_{n,1}^{(\dot{v})}\right)^2\right) \alpha_0 \, \mathrm{d}s \end{aligned} \tag{3.19}$$

on the set $B_n^{\complement}$. One easily shows that

$$\begin{aligned} &\int_{I(\tau)} \left(\widehat{\gamma}_n^{(\dot{u},\ddot{u})} \, \widehat{\gamma}_n^{(\dot{v},\ddot{v})}\right)^2 \left(\left(\widehat{\mu}_{n,2}^{(\dot{u},\dot{v})}\right)^2 + \left(\widehat{\mu}_{n,1}^{(\dot{u})} \, \widehat{\mu}_{n,1}^{(\dot{v})}\right)^2\right) \alpha_0 \, \mathrm{d}s \\ &\qquad - \int_{I(\tau)} \left(\gamma^{(\dot{u},\ddot{u})} \, \gamma^{(\dot{v},\ddot{v})}\right)^2 \left(\left(\mu_2^{(\dot{u},\dot{v})}\right)^2 + \left(\mu_1^{(\dot{u})} \, \mu_1^{(\dot{v})}\right)^2\right) \alpha_0 \, \mathrm{d}s \longrightarrow_{P_{n,0}} 0, \end{aligned}$$

where one uses Assumption 2.3.9 and Assumption 3.2.11. Consequently, the right hand side of (3.19) converges to 0 in $P_{n,0}$-probability. By the same considerations, one proves that

$$\int_{I(\tau)} \widehat{f}_n^{(u,v)} \, \frac{\widehat{\mu}_{n,0}}{\widehat{\mu}_{n,0}} \, \alpha_0 \, \mathrm{d}s - \int_{I(\tau)} \gamma^{(\dot{u},\ddot{u})} \, \gamma^{(\dot{v},\ddot{v})} \left(\mu_2^{(\dot{u},\dot{v})} - \frac{\mu_1^{(\dot{u})} \, \mu_1^{(\dot{v})}}{\mu_0}\right) \alpha_0 \, \mathrm{d}s \longrightarrow_{P_{n,0}} 0.$$

Thus, consistency holds for $\tau < \tau_0^c$.

The proof of consistency for $\tau = \tau_0^c$ is a bit trickier. Since $\widehat{V}_n(t)$ and $\mathscr{J}^{*,\mathrm{can}}(t)$ are both symmetric, it holds that $\widehat{V}_n(t) \longrightarrow_{P_{n,0}} \mathscr{J}^{*,\mathrm{can}}(t)$ is equivalent to $c^{\mathrm{T}} \widehat{V}_n(t) c \longrightarrow_{P_{n,0}} c^{\mathrm{T}} \mathscr{J}^{*,\mathrm{can}}(t) c$ for all $c \in \mathbb{R}^r$. Let us introduce some abbreviations

$$\begin{aligned} \widehat{X}_n(t) &= c^{\mathrm{T}} \widehat{V}_n(t) \, c \\ X_n(t) &= \sum_{u=1}^{r} \sum_{v=1}^{r} c^{(u)} c^{(v)} \int_{I(t)} \widehat{f}_n^{(u,v)}(s) \, \alpha_0(s) \, \mathrm{d}s \\ X(t) &= c^{\mathrm{T}} \mathscr{J}^{*,\mathrm{can}}(t) c. \end{aligned}$$

We have already shown that $\widehat{X}_n(\tau) - X_n(\tau) \longrightarrow_{P_{n,0}} 0$ for all $\tau < \tau_0^c$. As a first step we extend this result to $\tau = \tau_0^c$ using Theorem 2.1.1. We now merely have to prove that

$$\lim_{t\to\tau_0^c}\limsup_{n\to\infty} P_{n,0}\Big(\big|\widehat{X}_n(\tau_0^c)-\widehat{X}_n(t)-X_n(\tau_0^c)+X_n(t)\big|\geq\varepsilon\Big)=0 \tag{3.20}$$

for all $\varepsilon>0$. We show that

$$\lim_{t\to\tau_0^c}\limsup_{n\to\infty} P_{n,0}\Big(\big|\widehat{X}_n(\tau_0^c)-\widehat{X}_n(t)\big|\geq\varepsilon\Big)=0. \tag{3.21}$$

Since $\lim_{t\to\tau_0^c}\limsup_{n\to\infty} P_{n,0}\big(\big|X_n(\tau_0^c)-X_n(t)\big|\geq\varepsilon\big)=0$ is exactly condition (3.17), the condition (3.20) holds, if condition (3.21) holds.

From Remark 3.2.14 we know that $\big\{\widehat{X}_n(s\wedge\tau_0^c)\mid s\in\mathbb{R}_+\big\}$ is an increasing and non-negative process. As

$$\big\{\widehat{X}_n(\tau_0^c\wedge\tau\wedge s)-\widehat{X}_n(\tau_0^c\wedge t\wedge s)-X_n(\tau_0^c\wedge\tau\wedge s)+X_n(\tau_0^c\wedge t\wedge s)\mid s\in\mathbb{R}_+\big\},$$

$t<\tau<\tau_0^c$, is a local martingale, cf. Jacod and Shiryaev [32, Theorem I.3.18], the process

$$\big\{\widehat{X}_n(\tau_0^c\wedge\tau\wedge s)-\widehat{X}_n(\tau_0^c\wedge t\wedge s)\mid s\in\mathbb{R}_+\big\}$$

is Lenglart-dominated by the process

$$\big\{X_n(\tau_0^c\wedge\tau\wedge s)-X_n(\tau_0^c\wedge t\wedge s)\mid s\in\mathbb{R}_+\big\}.$$

Therefore, we can apply Theorem 2.1.5

$$P_{n,0}\bigg(\sup_{0\leq s\leq\tau}\widehat{X}_n(\tau_0^c\wedge\tau\wedge s)-\widehat{X}_n(t\wedge s)\geq\varepsilon\bigg)\leq\frac{\eta}{\varepsilon}+P_{n,0}\big(X_n(\tau)-X_n(t)\geq\eta\big)$$

Applying the Monotone Convergence Theorem, it results that

$$P_{n,0}\bigg(\sup_{0\leq s\leq\tau_0^c}\widehat{X}_n(\tau_0^c\wedge s)-\widehat{X}_n(t\wedge s)\geq\varepsilon\bigg)\leq\frac{\eta}{\varepsilon}+P_{n,0}\big(X_n(\tau_0^c)-X_n(t)\geq\eta\big),$$

where we also use the fact that the process $\widehat{X}_n$ does not jump at the point τ_0^c and that the paths of X_n are continuous. Thus, we proved

$$P_{n,0}\big(\widehat{X}_n(\tau_0^c)-\widehat{X}_n(t)\geq\varepsilon\big)\leq\frac{\eta}{\varepsilon}+P_{n,0}\big(X_n(\tau_0^c)-X_n(t)\geq\eta\big),$$

The condition (3.17) yields $\lim_{t\to\tau_0^c}\limsup_{n\to\infty} P_{n,0}\big(\widehat{X}_n(\tau_0^c)-\widehat{X}_n(t)\geq\varepsilon\big)\leq \eta/\varepsilon$. As $\eta>0$ was arbitrary, equation (3.20) holds, *i.e.* $\widehat{X}_n(\tau_0^c)-X_n(\tau_0^c)\longrightarrow_{P_{n,0}} 0$, as $n\to\infty$.

As the last step we show that $X_n(\tau_0^c)-X(\tau_0^c)\longrightarrow_{P_{n,0}} 0$. We have already shown that $X_n(t)-X(t)\longrightarrow_{P_{n,0}} 0$ for all $t<\tau_0^c$. Again, Theorem 2.1.1 gives that we merely have to show that

$$\lim_{t\to\tau_0^c}\limsup_{n\to\infty} P_{n,0}\Big(\big|X_n(\tau_0^c)-X_n(t)-X(\tau_0^c)+X(t)\big|\geq\varepsilon\Big)=0.$$

However, this condition holds trivially, because of condition (3.17) and the fact that $\lim_{t\to\tau_0^c} X(t)=X(\tau_0^c)$. □

3.2.14 Remark. a) The matrix $\widehat{V}_n$ is positive semi-definite. As the sum of positive semi-definite matrices is again a positive semi-definite matrix, *e.g.* cf. Brunner and Munzel [12, Satz B.40], it is sufficient to show that

$$\sum_{u=1}^{r}\sum_{v=1}^{r} w^{(u)}\, w^{(v)}\int_{I(\tau)} \widehat{f}_n^{(u,v)}(s)\,\frac{1}{\widehat{\mu}_{n,0}(s)}\,\mathrm{d}N_n^{(i)}(s)\geq 0$$

for all $w\in\mathbb{R}^r$ and $i\in\{1,\dots,n\}$. Using the abbreviation

$$c^{(s)}=\sum_{u=1}^{r_s} w^{(\sum_{l=1}^{s-1} r_l+u)}\ \widehat{\gamma}_n^{(s,u)}$$

we have

$$\begin{aligned}
&\sum_{u=1}^{r}\sum_{v=1}^{r} w^{(u)}\, w^{(v)}\,\widehat{f}_n^{(u,v)}\\
&\quad=\sum_{u_1=1}^{p}\sum_{u_2=1}^{r_{u_1}}\sum_{v_1=1}^{p}\sum_{v_2=1}^{r_{v_1}} w^{(\sum_{l=1}^{u_1-1} r_l+u_2)}\ w^{(\sum_{l=1}^{v_1-1} r_l+v_2)}\ \widehat{\gamma}_n^{(u_1,u_2)}\ \widehat{\gamma}_n^{(v_1,v_2)}\\
&\qquad\qquad\times\left(\widehat{\mu}_2^{(u_1,v_1)}-\frac{\widehat{\mu}_1^{(u_1)}\widehat{\mu}_1^{(v_1)}}{\widehat{\mu}_0}\right)\\
&\quad=\sum_{u_1=1}^{p}\sum_{v_1=1}^{p} c^{(u_1)}\ c^{(v_1)}\cdot\left(\widehat{\mu}_2^{(u_1,v_1)}-\frac{\widehat{\mu}_1^{(u_1)}\widehat{\mu}_1^{(v_1)}}{\widehat{\mu}_0}\right)\\
&\quad=\frac{1}{n}\sum_{j=1}^{n}(c^{\mathrm{T}}Z_{n,j})^2\,Y_n^{(j)}-\frac{1}{n\sum_{j=1}^{n} Y_n^{(j)}}\left(\sum_{j=1}^{n} c^{\mathrm{T}}Z_{n,j}\,Y_n^{(j)}\right)^2.
\end{aligned}$$

Set $\delta^{-1} = \sum_{j=1}^n Y_n^{(j)} > 0$. Multiplying the right hand side of the previous equation with δ gives

$$\frac{1}{n}\sum_{j=1}^{n}(c^{\mathrm{T}} Z_{n,j})^2\, Y_n^{(j)}\,\delta - \frac{1}{n}\left(\sum_{j=1}^{n} c^{\mathrm{T}} Z_{n,j}\, Y_n^{(j)}\,\delta\right)^2 .$$

Using the Jensen-Inequality gives the estimate

$$\sum_{j=1}^{n}(c^{\mathrm{T}} Z_{n,j})^2\, Y_n^{(j)}\,\delta \geq \left(\sum_{j=1}^{n} c^{\mathrm{T}} Z_{n,j}\, Y_n^{(j)}\delta\right)^2 ,$$

i.e the assertion.

b) The condition (3.17) is implied by

$$\lim_{t\to\tau_0^c}\limsup_{n\to\infty} P_{n,0}\left(\left|\int_{(t,\tau_0^c)} \widehat{\gamma}_n^{(\dot{u},\ddot{u})}\,\widehat{\gamma}_n^{(\dot{v},\ddot{v})}\left(\widehat{\mu}_{n,2}^{(\dot{u},\dot{v})} - \frac{\widehat{\mu}_{n,1}^{(\dot{u})}\,\widehat{\mu}_{n,1}^{(\dot{v})}}{\widehat{\mu}_{n,0}}\right)\alpha_0\,\mathrm{d}s\right| \geq \varepsilon\right) = 0$$

for all $u, v = 1, \ldots, r$ and $\varepsilon > 0$.

In the next chapter tests for linear and multivariate one-sided hypotheses are rigorously developed and their asymptotic properties are investigated for sequences of hardest parametric sub-models. Even though we start with some semi-parametric model, these tests turn out to be non-parametric.

4 Deriving Testing Procedures

In Chapter 2 we established asymptotic normality for sequences of parametric sub-models and in Chapter 3 the notion of sequences of hardest parametric sub-models (SHPSM). In Section 4.1 and Section 4.2 multivariate one-sided testing problems and linear testing problems are considered. The results are applied to SHPSM in Section 4.3. Moreover, it is shown that the resulting tests are non-parametric procedures and that they are generalizations of the projective-type tests of Mayer [53] and the general class of tests introduced by Jones and Crowley [39], see Section 4.4.

4.1 Multivariate One-Sided Testing Problems

First of all, let us introduce the premises for this section. Analog to previous chapters we consider sequences of experiments that are asymptotically normal, again.

4.1.1 Assumption. Let $(\Omega_n, \mathcal{A}_n, \mathfrak{P}_n)$, $n \in \mathbb{N}$, where $\mathfrak{P}_n = \{P_{n,\xi} \mid \xi^{\mathrm{T}} = (\beta^{\mathrm{T}}, \eta^{\mathrm{T}}),\ \beta \in \mathbb{R}^r,\ \eta \in \mathbb{R}^q\}$, be a sequence of experiments that is asymptotically normal with central sequence S_n, $n \in \mathbb{N}$, and asymptotic information matrix $\mathscr{I}$ that is partitioned as follows

$$\mathscr{I} = \begin{pmatrix} \mathscr{I}_{1,1} & \mathscr{I}_{1,2} \\ \mathscr{I}_{2,1} & \mathscr{I}_{2,2} \end{pmatrix},$$

where $\mathscr{I}_{1,1}$ is some $(r \times r)$ matrix and $\mathscr{I}_{2,2}$ is some $(q \times q)$ matrix. These premises mean that

$$\frac{\mathrm{d}P_{n,\xi}}{\mathrm{d}P_{n,0}} - \xi^{\mathrm{T}} S_n + \frac{1}{2}\xi^{\mathrm{T}} \mathscr{I}\xi \longrightarrow_{P_{n,0}} 0 \qquad \text{as } n \to \infty,$$

and $S_n \xrightarrow{\mathfrak{D}}_{P_{n,0}} \mathcal{N}(0, \mathscr{I})$, as $n \to \infty$.

4.1.2 Remark. Using the First Le Cam Lemma, cf. Witting and Müller-Funk [72, Korollar 6.124], one sees that the sequences of probability measures $\{P_{n,\xi} \mid n \in \mathbb{N}\}$ and $\{P_{n,0} \mid n \in \mathbb{N}\}$ are mutual contiguous. Especially, convergence in $P_{n,0}$-probability implies convergence in $P_{n,\xi}$-probability. The First Le Cam Lemma, Slutsky's Lemma and the Cramér-Wold device, cf. Witting and Müller-Funk [72, Korollar 6.124, Korollar 5.83, Korollar 5.69], give that $S_n \xrightarrow{\mathfrak{D}}_{P_{n,\xi}} \mathcal{N}(\mathscr{I}\xi, \mathscr{I})$, as $n \to \infty$. Hence, the sequence of experiments converges weakly to some Gauss shift experiment

$$\big(\Omega, \mathcal{A}, \mathfrak{G}\big), \qquad \mathfrak{G} = \{P_\xi \mid \xi^{\mathrm{T}} = (\beta^{\mathrm{T}}, \eta^{\mathrm{T}}) \in \mathbb{R}^{r+q}\},$$

where

$$S : (\Omega, \mathcal{A}) \longrightarrow (\mathbb{R}^{r+q}, \mathbb{B}^{r+q}), \qquad \mathfrak{L}\big(S \mid P_\xi\big) = \mathcal{N}\big(\mathscr{I}\xi, \mathscr{I}\big), \tag{4.1}$$

and

$$\frac{\mathrm{d}P_\xi}{\mathrm{d}P_0} = \exp\Big(S^{\mathrm{T}}\xi - \frac{1}{2}\xi^{\mathrm{T}} \mathscr{I}\xi\Big). \tag{4.2}$$

Before we can state the multivariate one-sided testing problem we have to introduce some more notation.

4.1.3 Definition. Let $m \geq 1$ be some integer and $\emptyset \neq \mathcal{K} \subset \{1, \ldots, m\}$. Then we define the mappings

$$\begin{aligned} \pi^m_{\mathcal{K}} &: \mathbb{R}^m \longrightarrow \mathbb{R}^{|\mathcal{K}|}, & \pi^m_{\mathcal{K}}(x) &= (\mathscr{T}^m_{\mathcal{K}})^{\mathrm{T}} x, \\ \rho^m_{\mathcal{K}} &: \mathbb{R}^{m\times m} \longrightarrow \mathbb{R}^{|\mathcal{K}|\times|\mathcal{K}|}, & \rho^m_{\mathcal{K}}(\mathscr{M}) &= (\mathscr{T}^m_{\mathcal{K}})^{\mathrm{T}} \mathscr{M} (\mathscr{T}^m_{\mathcal{K}}), \end{aligned}$$

where $\mathscr{T}^m_{\mathcal{K}} = \big(e_k \mid k \in \mathcal{K}\big)$, $e_k = \big(\delta_{1,k}, \ldots, \delta_{m,k}\big)^{\mathrm{T}}$ and $\delta_{u,v}$ denotes the Kronecker symbol.

The functions $\pi_{\mathcal{K}}^m$ and $\rho_{\mathcal{K}}^m$ are obviously projections. Let $\mathcal{J} \subset \{1, \ldots, r\}$ denote the components of the parameter β, that we are interested in, *i.e.* we want to test one-sided hypotheses that only depend on these components. The remaining components of β are regarded as nuisance parameters. Consequently, the vector $\beta_{\mathcal{J}} = \pi_{\mathcal{J}}^r(\beta)$ contains the interesting parameter and $\beta_{\mathcal{J}^\complement} = \pi_{\mathcal{J}^\complement}^r(\beta)$ contains the nuisance parameter. More precisely, the multivariate one-sided testing problem $\mathcal{H}_1^{\mathcal{J}}$ versus $\mathcal{K}_1^{\mathcal{J}}$, where

$$\mathcal{H}_1^{\mathcal{J}} : \beta_{\mathcal{J}} = 0, \, \beta_{\mathcal{J}^\complement} \in \mathbb{R}^{r-|\mathcal{J}|}, \, \eta \in \mathbb{R}^q$$

and

$$\mathcal{K}_1^{\mathcal{J}} : \beta_{\mathcal{J}} \geq 0, \; \beta_{\mathcal{J}} \neq 0, \, \beta_{\mathcal{J}^\complement} \in \mathbb{R}^{r-|\mathcal{J}|}, \, \eta \in \mathbb{R}^q,$$

is the subject of this section. Examples for one-sided testing problems are the two-sample problem with covariate adjustment or more generally any one-sided testing problem in the presence of concomitant covariates.

In order to derive some reasonable testing procedure we study the testing problem $\mathcal{H}_{\mathcal{J}}^1$ versus $\mathcal{K}_{\mathcal{J}}^1$ under the limit model and derive some test statistic. This statistic will be the basis to propose a test statistic for finite $n \in \mathbb{N}$. As we allow the asymptotic information matrix $\mathscr{I}$ to be degenerated, the hypothesis $\mathcal{H}_1^{\mathcal{J}}$ and the alternative $\mathcal{K}_1^{\mathcal{J}}$ are not necessarily disjoint. The next result helps us to state conditions for guaranteeing that the hypothesis and the alternative are disjoint

4.1.4 Lemma. $P_\xi = P_{\xi'}$ is equivalent to $\mathscr{I}\xi = \mathscr{I}\xi'$.

Proof. As $P_0\big(S \in \mathrm{Im}(\mathscr{I})\big) = 1$, cf. Witting [71, Hilfssatz 1.90], we get that

$$\frac{\mathrm{d}P_\xi}{\mathrm{d}P_0} = \exp\left(S^{\mathrm{T}} \mathscr{I}^-(\mathscr{I}\xi) - \frac{1}{2}(\mathscr{I}\xi)^{\mathrm{T}} \mathscr{I}^-(\mathscr{I}\xi)\right) \quad P_{n,0}\text{-almost surely.}$$

Therefore $\mathscr{I}\xi = \mathscr{I}\xi' \Rightarrow P_\xi = P_{\xi'}$ is trivial. On the other hand $P_\xi = P_{\xi'}$ implies that $P_0^S\big\{s \in \mathrm{Im}(\mathscr{I}) \mid f(s) = 0\big\} = 1$, where

$$f(s) = s^{\mathrm{T}} \mathscr{I}^-(\mathscr{I}\xi - \mathscr{I}\xi') - \frac{1}{2}(\mathscr{I}\xi)^{\mathrm{T}} \mathscr{I}^-(\mathscr{I}\xi) + \frac{1}{2}(\mathscr{I}\xi')^{\mathrm{T}} \mathscr{I}^-(\mathscr{I}\xi').$$

As $\mathfrak{L}(S \mid P_0) = \mathcal{N}(0, \mathscr{I})$, it holds that

$$0 = \mathbb{E}(f(S)) = -\frac{1}{2}(\mathscr{I}\xi)^{\mathrm{T}}\mathscr{I}^{-}(\mathscr{I}\xi) + \frac{1}{2}(\mathscr{I}\xi')^{\mathrm{T}}\mathscr{I}^{-}(\mathscr{I}\xi')$$

and

$$0 = \mathbb{E}\big(f(S) - \mathbb{E}f(S)\big)^2 = (\mathscr{I}\xi - \mathscr{I}\xi')^{\mathrm{T}}\mathscr{I}^{-}(\mathscr{I}\xi - \mathscr{I}\xi').$$

As $(s_1, s_2) \mapsto s_1^{\mathrm{T}}\mathscr{I}^{-}s_2$ is an inner product on $\mathrm{Im}(\mathscr{I})$, see Proposition B.2.5, we get that $\mathscr{I}\xi - \mathscr{I}\xi' = 0$. Remember that $\xi^{\mathrm{T}} = (\beta^{\mathrm{T}}, \eta^{\mathrm{T}})$ and ${\xi'}^{\mathrm{T}} = ({\beta'}^{\mathrm{T}}, {\eta'}^{\mathrm{T}})$. □

The previous result gives that the hypothesis $\mathcal{H}_1^{\mathcal{J}}$ and the alternative $\mathcal{H}_1^{\mathcal{J}}$ are disjoint, if and only if $\Theta(\mathcal{H}_1^{\mathcal{J}}) \cap \Theta(\mathcal{K}_1^{\mathcal{J}}) = \emptyset$, where

$$\Theta(\mathcal{H}_1^{\mathcal{J}}) = \{\mathscr{I}\xi \mid \xi \in \mathcal{H}_1^{\mathcal{J}}\} \quad \text{and} \quad \Theta(\mathcal{K}_1^{\mathcal{J}}) = \{\mathscr{I}\xi \mid \xi \in \mathcal{K}_1^{\mathcal{J}}\}$$

are the induced parameter sets of the hypothesis and the alternative. Later, the matrix $\mathscr{I}$ corresponds with the asymptotic information matrix of some sequence of hardest parametric sub-models, see Section 4.3. This means that we do not know the matrix $\mathscr{I}$ in general. Therefore, we state a criterion that depends on the matrix

$$\mathscr{I}^* = \mathscr{I}_{1,1} - \mathscr{I}_{1,2}\mathscr{I}_{2,2}^{-}\mathscr{I}_{2,1}$$

which corresponds with the asymptotic covariance matrix of the MCRM, see Theorem 3.2.4.c. We will see that this matrix is known to some extend. In Discussion 4.3.2 reasons for considering models with a degenerated asymptotic information matrix are provided and it is shown that these models satisfy the condition (4.3) given in the following result.

4.1.5 Proposition. Assume that

$$\beta \in \ker(\mathscr{I}^*)\backslash\{0\} \quad \Longrightarrow \quad \beta_{\mathcal{J}} \ngeq 0 \text{ and } -\beta_{\mathcal{J}} \ngeq 0, \tag{4.3}$$

where $\kappa \ngeq 0$ means that $\kappa^{(u)} < 0$ for some u. Then the hypothesis $\mathcal{H}_1^{\mathcal{J}}$ and the alternative $\mathcal{K}_1^{\mathcal{J}}$ are disjoint in the limit experiment. Note that condition (4.3) trivially holds, if the information matrix $\mathscr{I}^*$ is non-degenerated, since $\ker(\mathscr{I}^*)\backslash\{0\} = \emptyset$.

Proof. Lemma 4.1.4 implies that $P_\xi = P_{\xi'}$, if and only if $\mathscr{J}\xi = \mathscr{J}\xi'$. Without loss of generality, we can assume that $\eta, \eta' \in \ker(\mathscr{J}_{2,2})^\perp$, as $\mathscr{J}_{2,2}\widetilde{\eta} = 0$ and $\mathscr{J}_{1,2}\widetilde{\eta} = 0$ for all $\widetilde{\eta} \in \ker(\mathscr{J}_{2,2})$, see Proposition B.3.4.a. $\mathscr{J}\xi = \mathscr{J}\xi'$ is equivalent to

$$\begin{pmatrix} \mathscr{J}_{1,1} & \mathscr{J}_{1,2} \\ \mathscr{J}_{2,1} & \mathscr{J}_{2,2} \end{pmatrix} \begin{pmatrix} \beta \\ 0 \end{pmatrix} = \begin{pmatrix} \mathscr{J}_{1,1} & \mathscr{J}_{1,2} \\ \mathscr{J}_{2,1} & \mathscr{J}_{2,2} \end{pmatrix} \begin{pmatrix} \beta' \\ \eta' - \eta \end{pmatrix}$$

and

$$\mathscr{J}_{1,1}\big(\beta - \beta'\big) = \mathscr{J}_{1,2}\big(\eta' - \eta\big) \quad \text{and} \quad \mathscr{J}_{2,1}\big(\beta - \beta'\big) = \mathscr{J}_{2,2}\big(\eta' - \eta\big).$$

and

$$\mathscr{J}_{1,1}\big(\beta - \beta'\big) - \mathscr{J}_{1,2}\big(\eta' - \eta\big) = 0 \quad \text{and} \quad \mathscr{J}_{2,2}^{-}\mathscr{J}_{2,1}\big(\beta - \beta'\big) = \big(\eta' - \eta\big).$$

Putting these equations together yields that $\mathscr{J}^*\big(\beta - \beta'\big) = 0$, *i.e.* $\big(\beta - \beta'\big) \in \ker(\mathscr{J}^*)$.

Assume that $\xi \in \mathcal{H}_1^{\mathcal{J}}$ and $\xi' \in \mathcal{K}_1^{\mathcal{J}}$ and that $\mathscr{J}\xi = \mathscr{J}\xi'$. Applying the previous considerations, we get

$$\big(\beta - \beta'\big) = \begin{pmatrix} \mathscr{T}_{\mathcal{J}}^r & \mathscr{T}_{\mathcal{J}^\complement}^r \end{pmatrix} \begin{pmatrix} 0 - \beta'_{\mathcal{J}} \\ \beta_{\mathcal{J}^\complement} - \beta'_{\mathcal{J}^\complement} \end{pmatrix} \in \ker(\mathscr{J}^*)$$

and that $\beta'_{\mathcal{J}} \geq 0$ and $\beta'_{\mathcal{J}} \neq 0$. This contradicts our assumption. $\square$

Behnen and Neuhaus [8], who consider a similar testing problem, suggest the asymptotic likelihood ratio test statistic as basis for inference on β. Following their idea, we also aim to develop an asymptotic likelihood ratio test. For the testing problem $\mathcal{H}_1^{\mathcal{J}}$ versus $\mathcal{K}_1^{\mathcal{J}}$ the likelihood ratio test statistic is given by

$$T = 2\log \frac{\sup_{\xi \in \mathcal{H}_1^{\mathcal{J}} \cup \mathcal{K}_1^{\mathcal{J}}} \frac{\mathrm{d}P_\xi}{\mathrm{d}P_0}}{\sup_{\xi \in \mathcal{H}_1^{\mathcal{J}}} \frac{\mathrm{d}P_\xi}{\mathrm{d}P_0}},$$

see also Witting and Müller-Funk [72, pp. 215] for a justification of this proceeding.

For defining the statistic T one does not need the assumption that the hypothesis $\mathcal{H}_1^{\mathfrak{J}}$ and the alternative $\mathcal{K}_1^{\mathfrak{J}}$ are disjoint. Basically, T is the likelihood ratio test statistic for the testing problem

$$\widetilde{\mathcal{H}}_1^{\mathfrak{J}} : \xi \in \Theta_{\mathfrak{J},0} \qquad \text{versus} \qquad \widetilde{\mathcal{K}}_1^{\mathfrak{J}} : \xi \in \Theta_{\mathfrak{J},1} \backslash \Theta_{\mathfrak{J},0},$$

where

$$\Theta_{\mathfrak{J},0} = \big\{ \xi \mid \mathscr{J}\xi \in \Theta(\mathcal{H}_1^{\mathfrak{J}}) \big\} \quad \text{and} \quad \Theta_{\mathfrak{J},1} = \big\{ \xi \mid \mathscr{J}\xi \in \Theta(\mathcal{K}_1^{\mathfrak{J}}) \big\}.$$

If the condition (4.3) holds this testing problem is equivalent to the original testing problem, since $\Theta_{\mathfrak{J},0}$ and $\Theta_{\mathfrak{J},1}$ are disjoint. The transformation is based on the fact that $P_\xi = P_{\xi'}$ if and only if $\mathscr{J}\xi = \mathscr{J}\xi'$, see Lemma 4.1.4. The next results help us to simplify the statistic T.

4.1.6 Lemma. Let $\mathscr{A}$ be some symmetric, positive semi-definite $(k \times k)$ matrix, $s \in \mathrm{Im}(\mathscr{A})$ and $c \in \mathbb{R}$. It holds that

$$\sup_{x \in \mathbb{R}^k} s^{\mathrm{T}} x - \frac{1}{2} x^{\mathrm{T}} \mathscr{A} x + c = \frac{1}{2} s^{\mathrm{T}} \mathscr{A}^{-} s + c.$$

Proof. As $s \in \mathrm{Im}(\mathscr{A})$, we can write $s = \mathscr{A} s_0$ for some $s_0 \in \mathbb{R}^k$. It holds that $\mathscr{A}\mathscr{A}^{-} s = \mathscr{A}\mathscr{A}^{-}\mathscr{A} s_0 = \mathscr{A} s_0 = s$ and therefore $x_0 = \mathscr{A}^{-} s$ is a solution of $\mathscr{A} x = s$, see Proposition B.1.5. A Taylor-expansion at x_0 gives that

$$s^{\mathrm{T}} x - \frac{1}{2} x^{\mathrm{T}} \mathscr{A} x + c = \frac{1}{2} s^{\mathrm{T}} \mathscr{A}^{-} s + c - \frac{1}{2} (x - x_0)^{\mathrm{T}} \mathscr{A} (x - x_0). \tag{4.4}$$

Since $\mathscr{A}$ is positive semi-definite, it follows the assertion. □

4.1.7 Lemma. Let us introduce the following abbreviations

$$\mathscr{U}(\mathscr{J}) = \mathscr{T}^{r+q}_{\{1,\ldots,r\}} - \mathscr{T}^{r+q}_{\{r+1,\ldots,r+q\}} \mathscr{J}_{2,2}^{-} \mathscr{J}_{2,1}$$

and

$$\mathscr{Y}_{\mathfrak{J}}(\mathscr{J}^*) = \mathscr{T}^r_{\mathfrak{J}} - \mathscr{T}^r_{\mathfrak{J}^{\complement}} \big(\mathscr{H}^{\mathfrak{J}}_{2,2}(\mathscr{J}^*) \big)^{-} \mathscr{H}^{\mathfrak{J}}_{2,1}(\mathscr{J}^*)$$

where we set

$$\mathscr{H}^{\mathfrak{J}}_{i,j}(\mathscr{J}^*) = \mathscr{T}^{\mathrm{T}}_{\mathfrak{J},i} \mathscr{J}^* \mathscr{T}_{\mathfrak{J},j}, \qquad \mathscr{T}_{\mathfrak{J},1} = \mathscr{T}^r_{\mathfrak{J}}, \quad \text{and} \quad \mathscr{T}_{\mathfrak{J},2} = \mathscr{T}^r_{\mathfrak{J}^{\complement}}.$$

Abbreviating

$$\mathscr{H}^*_{\mathcal{J}}(\mathscr{I}^*) = \mathscr{H}^{\mathcal{J}}_{1,1}(\mathscr{I}^*) - \mathscr{H}^{\mathcal{J}}_{1,2}(\mathscr{I}^*)\big(\mathscr{H}^{\mathcal{J}}_{2,2}(\mathscr{I}^*)\big)^{-}\mathscr{H}^{\mathcal{J}}_{2,1}(\mathscr{I}^*),$$

we define

$$L_{\mathcal{J},\mathscr{I}^*}(y) = 2\sup_{\kappa\geq 0,\,\kappa\in\mathbb{R}^{|\mathcal{J}|}}\Big(\kappa^{\mathrm{T}}y - \frac{1}{2}\kappa^{\mathrm{T}}\mathscr{H}^*_{\mathcal{J}}(\mathscr{I}^*)\kappa\Big), \qquad y\in\mathbb{R}^{|\mathcal{J}|}, \tag{4.5}$$

$$\widetilde{L}_{\mathcal{J},\mathscr{I}^*}(u) = L_{\mathcal{J},\mathscr{I}^*}\big(\mathscr{Y}_{\mathcal{J}}(\mathscr{I}^*)^{\mathrm{T}}u\big), \qquad u\in\mathbb{R}^r, \tag{4.6}$$

and

$$L_{\mathcal{J},1}(u,\mathscr{I}^*) = \max\big\{f_{\mathfrak{I},\mathcal{J}}(u,\mathscr{I}^*) \mid \emptyset\neq\mathfrak{I}\subset\{1,\ldots,|\mathcal{J}|\}\big\}, \qquad u\in\mathbb{R}^r, \tag{4.7}$$

where

$$f_{\mathfrak{I},\mathcal{J}}(u,\mathscr{I}^*) = Q_{\mathfrak{I},\mathcal{J}}(u,\mathscr{I}^*)\cdot\prod_{i\in\mathfrak{I}}\mathbb{1}\Big(\pi^{|\mathfrak{I}|}_{\{i\}}\big(R_{\mathfrak{I},\mathcal{J}}(u,\mathscr{I}^*)\big)\geq 0\Big) \tag{4.8}$$

and

$$\begin{aligned}
Q_{\mathfrak{I},\mathcal{J}}(u,\mathscr{I}^*) &= y_{\mathfrak{I},\mathcal{J}}(u,\mathscr{I}^*)^{\mathrm{T}}R_{\mathfrak{I},\mathcal{J}}(u,\mathscr{I}^*),\\
R_{\mathfrak{I},\mathcal{J}}(u,\mathscr{I}^*) &= \Big(\rho^{|\mathcal{J}|}_{\mathfrak{I}}\big(\mathscr{H}^*_{\mathcal{J}}(\mathscr{I}^*)\big)\Big)^{-}y_{\mathfrak{I},\mathcal{J}}(u,\mathscr{I}^*),\\
y_{\mathfrak{I},\mathcal{J}}(u,\mathscr{I}^*) &= \pi^{|\mathcal{J}|}_{\mathfrak{I}}\big(\mathscr{Y}_{\mathcal{J}}(\mathscr{I}^*)^{\mathrm{T}}u\big).
\end{aligned} \tag{4.9}$$

a) For any $s\in\mathrm{Im}(\mathscr{I})$, it holds that

$$2\log\frac{\sup_{\xi\in\mathcal{H}^{\mathcal{J}}_1\cup\mathcal{K}^{\mathcal{J}}_1}\exp\Big(s^{\mathrm{T}}\xi-\frac{1}{2}\xi^{\mathrm{T}}\mathscr{I}\xi\Big)}{\sup_{\xi\in\mathcal{H}^{\mathcal{J}}_1}\exp\Big(s^{\mathrm{T}}\xi-\frac{1}{2}\xi^{\mathrm{T}}\mathscr{I}\xi\Big)} = \widetilde{L}_{\mathcal{J},\mathscr{I}^*}\big(\mathscr{U}(\mathscr{I})^{\mathrm{T}}s\big). \tag{4.10}$$

b) For any $u\in\mathrm{Im}(\mathscr{I}^*)$, it holds that $\widetilde{L}_{\mathcal{J},\mathscr{I}^*}\big(u\big) = L_{\mathcal{J},1}\big(u,\mathscr{I}^*\big)$.

c) $L_{\mathcal{J},\mathscr{I}^*}:\mathrm{Im}\big(\mathscr{H}^*_{\mathcal{J}}(\mathscr{I}^*)\big)\longrightarrow\mathbb{R}$ and $\widetilde{L}_{\mathcal{J},\mathscr{I}^*}:\mathrm{Im}(\mathscr{I}^*)\longrightarrow\mathbb{R}$ are convex and continuous functions.

Proof. For the proof set $\mathscr{U}=\mathscr{U}(\mathscr{I})$, $\mathscr{H}_{i,j}=\mathscr{H}^{\mathcal{J}}_{i,j}(\mathscr{I}^*)$, $\mathscr{H}^*=\mathscr{H}^*_{\mathcal{J}}(\mathscr{I}^*)$ and $\mathscr{Y}_{\mathcal{J}}=\mathscr{Y}_{\mathcal{J}}(\mathscr{I}^*)$. Obviously, it holds that

$$\begin{aligned}
\mathscr{U}^{\mathrm{T}}\mathscr{I}\mathscr{U} &= \begin{pmatrix}\mathscr{I}_{1,1}-\mathscr{I}_{1,2}\mathscr{I}^{-}_{2,2}\mathscr{I}_{2,1} & \mathscr{I}_{1,2}-\mathscr{I}_{1,2}\mathscr{I}^{-}_{2,2}\mathscr{I}_{2,2}\end{pmatrix}\mathscr{U}\\
&= \begin{pmatrix}\mathscr{I}^* & 0\end{pmatrix}\mathscr{U} = \mathscr{I}^*,
\end{aligned}$$

where we use Proposition B.3.4.b and the properties of a generalized inverse. Analogously, it is proved that $\mathscr{Y}_{\mathcal{J}}^{\mathrm{T}} \mathscr{J}^* \mathscr{Y}_{\mathcal{J}} = \mathscr{H}^*$. Note that these matrices are also symmetric and positive semi-definite. Abbreviate $s_1 = \pi^{r+q}_{\{1,\dots,r\}}(s)$ and $s_2 = \pi^{r+q}_{\{r+1,\dots,q\}}(s)$ and set

$$\begin{aligned} p(\beta_{\mathcal{J}}, \beta_{\mathcal{J}^{\complement}}, \eta) &= \beta_{\mathcal{J}}^{\mathrm{T}} \pi^r_{\mathcal{J}}(s_1) + \beta_{\mathcal{J}^{\complement}}^{\mathrm{T}} \pi^r_{\mathcal{J}^{\complement}}(s_1) + \eta^{\mathrm{T}} s_2 \\ &\quad - \frac{1}{2} \beta_{\mathcal{J}^{\complement}}^{\mathrm{T}} \rho^r_{\mathcal{J}^{\complement}}(\mathscr{J}_{1,1}) \beta_{\mathcal{J}^{\complement}} - \frac{1}{2} \beta_{\mathcal{J}}^{\mathrm{T}} \rho^r_{\mathcal{J}}(\mathscr{J}_{1,1}) \beta_{\mathcal{J}} - \beta_{\mathcal{J}^{\complement}}^{\mathrm{T}} \mathscr{T}^{r\,\mathrm{T}}_{\mathcal{J}^{\complement}} \mathscr{J}_{1,1} \mathscr{T}^r_{\mathcal{J}} \beta_{\mathcal{J}} \\ &\quad - \eta^{\mathrm{T}} \mathscr{J}_{2,1} \mathscr{T}^r_{\mathcal{J}} \beta_{\mathcal{J}} - \eta^{\mathrm{T}} \mathscr{J}_{2,1} \mathscr{T}^r_{\mathcal{J}^{\complement}} \beta_{\mathcal{J}^{\complement}} - \frac{1}{2} \eta^{\mathrm{T}} \mathscr{J}_{2,2} \eta. \end{aligned}$$

One sees that the left hand side of (4.10) is equal to

$$\begin{aligned} &2 \sup\big\{p(\beta_{\mathcal{J}}, \beta_{\mathcal{J}^{\complement}}, \eta) \mid \beta_{\mathcal{J}} \geq 0,\ \beta_{\mathcal{J}}^{\complement} \in \mathbb{R}^{r-|\mathcal{J}|},\ \eta \in \mathbb{R}^q\big\} \\ &\qquad - 2 \sup\big\{p(0, \beta_{\mathcal{J}^{\complement}}, \eta) \mid \beta_{\mathcal{J}}^{\complement} \in \mathbb{R}^{r-|\mathcal{J}|},\ \eta \in \mathbb{R}^q\big\} \\ &= 2 \sup\Big\{\sup\Big\{\sup\big\{p(\beta_{\mathcal{J}}, \beta_{\mathcal{J}^{\complement}}, \eta) \mid \eta \in \mathbb{R}^q\big\} \Bigm| \beta_{\mathcal{J}^{\complement}} \in \mathbb{R}^{r-|\mathcal{J}|}\Big\} \Bigm| \beta_{\mathcal{J}} \geq 0\Big\} \\ &\qquad - 2 \sup\Big\{\sup\big\{p(0, \beta_{\mathcal{J}^{\complement}}, \eta) \mid \eta \in \mathbb{R}^q\big\} \Bigm| \beta_{\mathcal{J}^{\complement}} \in \mathbb{R}^{r-|\mathcal{J}|}\Big\} \\ &= 2 \sup\big\{M(\beta_{\mathcal{J}}, \beta_{\mathcal{J}^{\complement}}, \eta) \mid \beta_{\mathcal{J}} \geq 0\big\} - 2\, M(0, \beta_{\mathcal{J}^{\complement}}, \eta) \end{aligned}$$

where

$$M(\beta_{\mathcal{J}}, \beta_{\mathcal{J}^{\complement}}, \eta) = \sup\Big\{\sup\big\{p(\beta_{\mathcal{J}}, \beta_{\mathcal{J}^{\complement}}, \eta) \mid \eta \in \mathbb{R}^q\big\} \Bigm| \beta_{\mathcal{J}^{\complement}} \in \mathbb{R}^{r-|\mathcal{J}|}\Big\}.$$

Proposition B.3.4.b and Proposition B.3.4.c give that

$$s_2 - \mathscr{J}_{2,1} \mathscr{T}_{\mathcal{J}} \beta_{\mathcal{J}} - \mathscr{J}_{2,1} \mathscr{T}_{\mathcal{J}^{\complement}} \beta_{\mathcal{J}^{\complement}} \in \operatorname{Im}(\mathscr{J}_{2,2}).$$

Applying Lemma 4.1.6 yields

$$\begin{aligned} \sup\big\{p(\beta_{\mathcal{J}}, \beta_{\mathcal{J}^{\complement}}, \eta) \mid \eta \in \mathbb{R}^q\big\} &= \beta_{\mathcal{J}}^{\mathrm{T}} \pi^r_{\mathcal{J}}\big(\mathscr{U}^{\mathrm{T}} s\big) - \frac{1}{2} \beta_{\mathcal{J}}^{\mathrm{T}} \mathscr{H}_{1,1} \beta_{\mathcal{J}} \\ &\quad + \frac{1}{2} s_2 \mathscr{J}_{2,2}^{-} s_2 + \beta_{\mathcal{J}^{\complement}}^{\mathrm{T}} \pi^r_{\mathcal{J}^{\complement}}\big(\mathscr{U}^{\mathrm{T}} s\big) - \frac{1}{2} \beta_{\mathcal{J}^{\complement}}^{\mathrm{T}} \mathscr{H}_{2,2} \beta_{\mathcal{J}^{\complement}} - \beta_{\mathcal{J}^{\complement}}^{\mathrm{T}} \mathscr{H}_{2,1} \beta_{\mathcal{J}} \end{aligned}$$

after some tedious computations. As

$$\begin{pmatrix} \mathscr{H}_{1,1} & \mathscr{H}_{1,2} \\ \mathscr{H}_{2,1} & \mathscr{H}_{2,2} \end{pmatrix} = \begin{pmatrix} \mathscr{T}_{\mathcal{J},1}^{\mathrm{T}} \\ \mathscr{T}_{\mathcal{J},2}^{\mathrm{T}} \end{pmatrix} \mathscr{J}^* \begin{pmatrix} \mathscr{T}_{\mathcal{J},1} & \mathscr{T}_{\mathcal{J},2} \end{pmatrix},$$

Proposition B.3.4.b and Proposition B.3.4.c give that $\pi^r_{\mathcal{J}^\complement}(\mathscr{U}^{\mathrm{T}} s) - \mathscr{H}_{2,1}\beta_{\mathcal{J}} \in \mathrm{Im}(\mathscr{H}_{2,2})$. Applying Lemma 4.1.6 yields that

$$\begin{aligned} M(\beta_{\mathcal{J}}, \beta_{\mathcal{J}^\complement}, \eta) &= \beta_{\mathcal{J}}^{\mathrm{T}} \mathscr{Y}_{\mathcal{J}}^{\mathrm{T}} \mathscr{U}^{\mathrm{T}} s - \frac{1}{2}\beta_{\mathcal{J}}^{\mathrm{T}} \mathscr{H}^* \beta_{\mathcal{J}} \\ &\quad + \frac{1}{2}\Big(\pi^r_{\mathcal{J}^\complement}\big(\mathscr{U}^{\mathrm{T}} s\big)\Big)^{\mathrm{T}} \mathscr{H}_{2,2}^{-}\Big(\pi^r_{\mathcal{J}^\complement}\big(\mathscr{U}^{\mathrm{T}} s\big)\Big) + \frac{1}{2} s_2^{\mathrm{T}} \mathscr{I}_{2,2}^{-} s_2. \end{aligned}$$

As $2\sup\big\{M(\beta_{\mathcal{J}}, \beta_{\mathcal{J}^\complement}, \eta) \mid \beta_{\mathcal{J}} \geq 0\big\} - 2\,M(0, \beta_{\mathcal{J}^\complement}, \eta) = \widetilde{L}_{\mathcal{J}, \mathscr{I}^*}\big(\mathscr{U}^{\mathrm{T}} s\big)$, the proof of a) is complete.

Proof of b). As $\mathscr{Y}_{\mathcal{J}}^{\mathrm{T}} u \in \mathrm{Im}(\mathscr{H}^*)$, cf. Proposition B.3.4.c, it holds that

$$\widetilde{L}_{\mathcal{J}, \mathscr{I}^*}\big(u\big) = 2\sup_{\kappa \geq 0}\Big(\big(\mathscr{H}^*\kappa\big)^{\mathrm{T}}\big(\mathscr{H}^*\big)^{-}\mathscr{Y}_{\mathcal{J}}^{\mathrm{T}} u - \frac{1}{2}\big(\mathscr{H}^*\kappa\big)^{\mathrm{T}}\big(\mathscr{H}^*\big)^{-}\big(\mathscr{H}^*\big)\kappa\Big).$$

Proposition B.2.5 yields the assertion.

Proof of c). First we note that the set $\mathrm{Im}(\mathscr{I}^*)$ is convex and that $0 \leq \widetilde{L}_{\mathcal{J}, \mathscr{I}^*}(u) < \infty$ for all $u \in \mathrm{Im}(\mathscr{I}^*)$, because of b).

$$\begin{aligned} \widetilde{L}_{\mathcal{J}, \mathscr{I}^*}\big(\lambda\, u_1 + (1-\lambda)\, u_2\big) &= L_{\mathcal{J}, \mathscr{I}^*}\Big(\mathscr{Y}_{\mathcal{J}}^{\mathrm{T}}\big(\lambda\, u_1 + (1-\lambda)\, u_2\big)\Big) \\ &\leq \lambda\, L_{\mathcal{J}, \mathscr{I}^*}\big(\mathscr{Y}_{\mathcal{J}}^{\mathrm{T}} u_1\big) + (1-\lambda)\, L_{\mathcal{J}, \mathscr{I}^*}\big(\mathscr{Y}_{\mathcal{J}}^{\mathrm{T}} u_2\big) \\ &= \lambda\, \widetilde{L}_{\mathcal{J}, \mathscr{I}^*}(u_1) + (1-\lambda)\, \widetilde{L}_{\mathcal{J}, \mathscr{I}^*}(u_2) \end{aligned}$$

for all $u_i \in \mathrm{Im}(\mathscr{I}^*)$ and $\lambda \in (0,1)$, *i.e.* $\widetilde{L}_{\mathcal{J}, \mathscr{I}^*}$ is convex and therefore continuous, cf. Borwein and Lewis [10, Theorem 4.1.3]. Since any $y_i \in \mathrm{Im}(\mathscr{H}^*)$ can be represented as $y_i = \mathscr{H}^* \widetilde{y}_i$, one gets that $y_i = \mathscr{Y}_{\mathcal{J}}^{\mathrm{T}} u_i$, where $u_i = \mathscr{I}^* \mathscr{Y}_{\mathcal{J}} \widetilde{y}_i$. Since $\mathrm{Im}(\mathscr{H}^*)$ is convex, we have

$$\begin{aligned} L_{\mathcal{J}, \mathscr{I}^*}\big(\lambda\, y_1 + (1-\lambda)\, y_2\big) &= \widetilde{L}_{\mathcal{J}, \mathscr{I}^*}\big(\lambda\, u_1 (1-\lambda)\, u_2\big) \\ &\leq \lambda\, \widetilde{L}_{\mathcal{J}, \mathscr{I}^*}(u_1) + (1-\lambda)\, \widetilde{L}_{\mathcal{J}, \mathscr{I}^*}\,(u_2) \\ &= \lambda\, L_{\mathcal{J}, \mathscr{I}^*}(y_1) + (1-\lambda)\, L_{\mathcal{J}, \mathscr{I}^*}(y_2), \end{aligned}$$

where we use the convexity of $\widetilde{L}_{\mathcal{J}, \mathscr{I}^*}$. Thus, $L_{\mathcal{J}, \mathscr{I}^*}$ is convex and continuous, cf. Borwein and Lewis [10, Theorem 4.1.3]. $\square$

Using the previous result we can simplify our test statistic T, whose distribution is calculated in Theorem 4.1.14.

4.1.8 Corollary. It holds that

$$T = L_{\mathfrak{J},1}(U, \mathscr{I}^*) \quad P_\xi\text{-almost surely},$$

where $U = \mathscr{U}(\mathscr{I})^{\mathrm{T}} S$.

Proof. As $P_\xi \ll P_0$, it suffices to show the assertion for P_0. Witting [71, Hilfssatz 1.90] gives that $P_0\big(S \in \operatorname{Im}(\mathscr{I})\big) = 1$ and $P_0\big(U \in \operatorname{Im}(\mathscr{I}^*)\big) = 1$. Consequently, Lemma 4.1.7.a and Lemma 4.1.7.b imply the assertion. □

4.1.9 Remark. In Discussion 3.1.1 we showed that the statistic

$$T_{\beta_0} = \beta_0^{\mathrm{T}}\big(S_1 - \mathscr{I}_{1,2}\mathscr{I}_{2,2}^{-1}S_2\big) - \frac{1}{2}\beta_0^{\mathrm{T}}\mathscr{I}^*\beta_0 = \beta_0^{\mathrm{T}}\mathscr{U}(\mathscr{I})^{\mathrm{T}}S - \frac{1}{2}\beta_0^{\mathrm{T}}\mathscr{I}^*\beta_0,$$

where we set $S_1 = \pi^{r+q}_{\{1,\dots,r\}}(S)$ and $S_2 = \pi^{r+q}_{\{r+1,\dots,r+q\}}(S)$, is efficient for the testing problem $\beta = 0$, $\eta \in \mathbb{R}^q$ versus $\beta = \beta_0$, $\eta \in \mathbb{R}^q$. Remember that the critical values can be chosen independently of η. Now, let us assume that β is one-dimensional, *i.e.* $r = 1$. Clearly,

$$\widetilde{T} = \frac{\sqrt{\mathscr{I}^*}}{\beta_0}\Big(T_{\beta_0} + \frac{1}{2}\beta_0^{\mathrm{T}}\mathscr{I}^*\beta_0\Big) = \sqrt{\mathscr{I}^*}\,\mathscr{U}^{\mathrm{T}}(\mathscr{I})S.$$

is an efficient test statistic for $\beta = 0$, $\eta \in \mathbb{R}^q$ versus $\beta = \beta_0$, $\eta \in \mathbb{R}^q$. More precisely, the test $\psi = \mathbb{1}\big(\widetilde{T} > c(\alpha)\big)$, where $c(\alpha)$ is chosen, such that $P_0\big(\widetilde{T} > c(\alpha)\big) = \alpha$, is the most powerful α-test. This test is independent of β_0, therefore it is even the most powerful α-test for the testing problem $\beta = 0$, $\eta \in \mathbb{R}^q$ versus $\beta > 0$, $\eta \in \mathbb{R}^q$. As $\psi = \mathbb{1}\big(T > \big(c(\alpha)\big)^2\big)$, our likelihood ratio test statistic is optimal in the case of a one-dimensional β. So, we can expect to obtain reasonable testing procedures by using the likelihood ratio test statistic. The efficiency of our tests will be discussed in greater detail later in this section.

Since the statistic U corresponds with the statistic $U_n = \mathscr{U}(\mathscr{J})^{\mathrm{T}} S_n$ for finite n, the previous result suggests that $L_{\mathcal{J},1}\big(U_n, \mathscr{J}^*\big)$ is a reasonable test statistic for the testing problem $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathcal{J}}$. However, this statistic still depends on asymptotic quantities. In order to get some applicable test we have to replace these quantities by suitable estimators. In the next few steps we provide the results needed for proving that if $\widehat{V}_n - \mathscr{J}^* \longrightarrow_{P_{n,0}} 0$ and $U_n - \widehat{U}_n \longrightarrow_{P_{n.0}} 0$, as $n \to \infty$, it holds that

$$L_{\mathcal{J},1}\big(U_n, \mathscr{J}^*\big) - L_{\mathcal{J},1}\big(\widehat{U}_n, \widehat{V}_n\big) \longrightarrow_{P_{n,\xi}} 0, \qquad \text{as } n \to \infty,$$

for all $\xi \in \mathbb{R}^{r+q}$. Hence, one can use $L_{\mathcal{J},1}\big(\widehat{U}_n, \widehat{V}_n\big)$ as test statistic that does not depend anymore on asymptotic quantities for suitable $\widehat{U}_n$ and $\widehat{V}_n$. The following Lemma is a generalization of a result that can be found in Janssen [34, p. 151].

4.1.10 Lemma. Assume that $\mathscr{A}_n$, $n \in \mathbb{N}$, is a sequence of real, symmetric, positive semi-definite $(k \times k)$ random matrices, such that

(i) $\mathscr{A}_n - \mathscr{A} \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$, where $\mathscr{A}$ is a real, symmetric, positive semi-definite $(k \times k)$ matrix, and

(ii) $\ker(\mathscr{A}) \subset \ker(\mathscr{A}_n)$ $\mathbb{P}_n$-almost surely for all sufficiently large $n \in \mathbb{N}$.

The following assertions hold true.

a) $\mathscr{A}_n^- - \mathscr{A}^- \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$.

b) $(\mathscr{V}^{\mathrm{T}} \mathscr{A}_n \mathscr{V})^- - (\mathscr{V}^{\mathrm{T}} \mathscr{A} \mathscr{V})^- \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$ for any $(k \times m)$-matrix $\mathscr{V}$.

c) In the case that $k = r$ and using the notation of Lemma 4.1.7, where we replace $\mathscr{J}^*$ by $\mathscr{A}$ and $\mathscr{A}_n$, it holds that

   (i) $\mathscr{Y}_{\mathcal{J}}(\mathscr{A}_n) - \mathscr{Y}_{\mathcal{J}}(\mathscr{A}) \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$.

   (ii) $\ker\big(\mathscr{H}_{\mathcal{J}}^*(\mathscr{A})\big) \subset \ker\big(\mathscr{H}_{\mathcal{J}}^*(\mathscr{A}_n)\big)$ $\mathbb{P}_n$-almost surely for all sufficiently large $n \in \mathbb{N}$ and $\mathscr{H}_{\mathcal{J}}^*(\mathscr{A}_n) - \mathscr{H}_{\mathcal{J}}^*(\mathscr{A}) \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$.

Proof. Without loss of generality we can assume that all random matrices are defined on the same probability space. By applying the sub-sub-sequence principle for random variables that converge in probability, cf. Proposition B.4.8, we merely have to show the assertion for sub-sequences of fixed sequences of non-random matrices.

$\ker(\mathscr{A}) \subset \ker(\mathscr{A}_n)$ yields $\dim \operatorname{Im}(\mathscr{A}) \geq \dim \operatorname{Im}(\mathscr{A}_n)$, *i.e.* $\operatorname{rank}(\mathscr{A}_n) \leq \operatorname{rank}(\mathscr{A})$. Abbreviating $l = \operatorname{rank}(\mathscr{A})$, we note that the set of all real $(k \times k)$ matrices with rank greater or equal to l is an open set. Thus, we can assume that $\operatorname{rank}(\mathscr{A}_n) = l$ for all sufficiently large n. Using the representation $\mathscr{A}_n = \mathscr{F}_n^{\mathrm{T}} \mathscr{D}_n \mathscr{F}_n$, where $\mathscr{F}_n$ is some orthogonal matrix and $\mathscr{D}_n = \operatorname{diag}\big(\lambda_{n,1}, \ldots, \lambda_{n,l}, 0, \ldots, 0\big)$ we get that $\mathscr{A}_n^- = \mathscr{F}_n^{\mathrm{T}} \mathscr{D}_n^- \mathscr{F}_n$, see Proposition B.1.6. One immediately sees that $\liminf_{n\to\infty} \min_{1 \leq i \leq l}\big(\lambda_{n,i}\big) > 0$. Consequently, it holds that $0 < \operatorname{trace}\big(\mathscr{A}_n^-\big) = \sum_{i=1}^{l} \lambda_{n,i}^{-1} \leq K < \infty$. Thus, the sequence $\mathscr{A}_n^-$, $n \in \mathbb{N}$, is relative compact and therefore contains an accumulation point $\mathscr{A}_0$.

We see that $\mathscr{A}_n \mathscr{A}_n^-$, $\mathscr{A}_n^- \mathscr{A}_n$, $\mathscr{A}_n \mathscr{A}_n^- \mathscr{A}_n$ and $\mathscr{A}_n^- \mathscr{A}_n \mathscr{A}_n^-$ have the accumulation points $\mathscr{A}\mathscr{A}_0$, $\mathscr{A}_0\mathscr{A}$, $\mathscr{A}\mathscr{A}_0\mathscr{A}$ and $\mathscr{A}_0\mathscr{A}\mathscr{A}_0$ respectively. Now, one easily checks that $\mathscr{A}_0$ satisfies the conditions of Definition B.1.1. As the generalized inverse is uniquely determined, see Proposition B.1.2, the proof of a) is complete.

Proof of b). Assume that $\kappa \in \ker(\mathscr{V}^{\mathrm{T}} \mathscr{A} \mathscr{V})$. Using Proposition B.3.2.b gives that $\mathscr{V}\kappa \in \ker(\mathscr{A})$ and therefore $\mathscr{V}\kappa \in \ker(\mathscr{A}_n)$ $\mathbb{P}$-almost surely. Again, Proposition B.3.2.b yields that $\kappa \in \ker(\mathscr{V}^{\mathrm{T}} \mathscr{A}_n \mathscr{V})$ $\mathbb{P}$-almost surely. The assertion is implied by a).

Proof of the first part of c). Applying b) yields that

$$\big(\mathscr{H}_{2,2}^{\mathfrak{d}}(\mathscr{A}_n)\big)^- - \big(\mathscr{H}_{2,2}^{\mathfrak{d}}(\mathscr{A})\big)^- \longrightarrow_{\mathbb{P}_n} 0, \qquad \text{as } n \to \infty.$$

Proposition B.4.6 gives that $\mathscr{H}_{2,1}^{\mathfrak{d}}(\mathscr{A}_n) - \mathscr{H}_{2,1}^{\mathfrak{d}}(\mathscr{A}) \longrightarrow_{\mathbb{P}_n} 0$. Finally, Proposition B.4.6 yields the assertion.

Proof of the second part of c). Proposition B.4.6 gives that $\mathscr{H}_{i,j}^{\mathfrak{d}}(\mathscr{A}_n) - \mathscr{H}_{i,j}^{\mathfrak{d}}(\mathscr{A}) \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$. b) yields that

$$\big(\mathscr{H}_{2,2}^{\mathfrak{d}}(\mathscr{A}_n)\big)^- - \big(\mathscr{H}_{2,2}^{\mathfrak{d}}(\mathscr{A})\big)^- \longrightarrow_{\mathbb{P}_n} 0, \qquad \text{as } n \to \infty.$$

Combing these results and Proposition B.4.6 imply the first assertion.

We show that $\ker\bigl(\mathscr{H}_{\mathcal{J}}^*(\mathscr{A})\bigr) = \pi_{\mathcal{J}}^r\bigl(\ker(\mathscr{A})\bigr)$. Assume that $\kappa \in \ker\bigl(\mathscr{H}_{\mathcal{J}}^*(\mathscr{A})\bigr)$ Using Proposition B.3.2.b one gets that $\mathscr{Y}_{\mathcal{J}}(\mathscr{A})\kappa \in \ker(\mathscr{A})$, since $\mathscr{H}_{\mathcal{J}}^*(\mathscr{A}) = \mathscr{Y}_{\mathcal{J}}(\mathscr{A})^{\mathrm{T}}\mathscr{A}\mathscr{Y}_{\mathcal{J}}(\mathscr{A})$. As $\pi_{\mathcal{J}}^r\bigl(\mathscr{Y}_{\mathcal{J}}(\mathscr{A})\kappa\bigr) = \kappa$, it results directly that $\ker\bigl(\mathscr{H}_{\mathcal{J}}^*(\mathscr{A})\bigr) \subset \pi_{\mathcal{J}}^r\bigl(\ker(\mathscr{A})\bigr)$. Assume that $\kappa \in \ker(\mathscr{A})$. Proposition B.3.2.b gives that

$$0 = \kappa^{\mathrm{T}}\mathscr{A}\kappa = \begin{pmatrix} \pi_{\mathcal{J}}^r(\kappa) \\ \pi_{\mathcal{J}^{\complement}}^r(\kappa) \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \mathscr{H}_{1,1}^{\mathcal{J}}(\mathscr{A}) & \mathscr{H}_{1,2}^{\mathcal{J}}(\mathscr{A}) \\ \mathscr{H}_{2,1}^{\mathcal{J}}(\mathscr{A}) & \mathscr{H}_{2,2}^{\mathcal{J}}(\mathscr{A}) \end{pmatrix} \begin{pmatrix} \pi_{\mathcal{J}}^r(\kappa) \\ \pi_{\mathcal{J}^{\complement}}^r(\kappa) \end{pmatrix}$$

and that

$$\begin{aligned} \mathscr{H}_{1,1}^{\mathcal{J}}(\mathscr{A})\pi_{\mathcal{J}}^r(\kappa) + \mathscr{H}_{1,2}^{\mathcal{J}}(\mathscr{A})\pi_{\mathcal{J}^{\complement}}^r(\kappa) = 0. \\ \mathscr{H}_{2,1}^{\mathcal{J}}(\mathscr{A})\pi_{\mathcal{J}}^r(\kappa) + \mathscr{H}_{2,2}^{\mathcal{J}}(\mathscr{A})\pi_{\mathcal{J}^{\complement}}^r(\kappa) = 0. \end{aligned}$$

Using these equations and Proposition B.3.4.b we finally get that

$$\begin{aligned} \mathscr{H}_{\mathcal{J}}^*(\mathscr{A})\pi_{\mathcal{J}}^r(\kappa) &= \mathscr{H}_{1,1}^{\mathcal{J}}(\mathscr{A})\pi_{\mathcal{J}}^r(\kappa) - \mathscr{C}^{\mathrm{T}}\mathscr{H}_{2,2}^{\mathcal{J}}(\mathscr{A})\bigl(\mathscr{H}_{2,2}^{\mathcal{J}}(\mathscr{A})\bigr)^{-}\mathscr{H}_{2,2}^{\mathcal{J}}(\mathscr{A})\mathscr{C}\pi_{\mathcal{J}}^r(\kappa) \\ &= \mathscr{H}_{1,1}^{\mathcal{J}}(\mathscr{A})\pi_{\mathcal{J}}^r(\kappa) - \mathscr{C}^{\mathrm{T}}\mathscr{H}_{2,1}^{\mathcal{J}}(\mathscr{A})\pi_{\mathcal{J}}^r(\kappa) \\ &= \mathscr{H}_{1,1}^{\mathcal{J}}(\mathscr{A})\pi_{\mathcal{J}}^r(\kappa) + \mathscr{C}^{\mathrm{T}}\mathscr{H}_{2,2}^{\mathcal{J}}(\mathscr{A})\pi_{\mathcal{J}^{\complement}}^r(\kappa) \\ &= \mathscr{H}_{1,1}^{\mathcal{J}}(\mathscr{A})\pi_{\mathcal{J}}^r(\kappa) + \mathscr{H}_{1,2}^{\mathcal{J}}(\mathscr{A})\pi_{\mathcal{J}^{\complement}}^r(\kappa) = 0, \end{aligned}$$

i.e. $\ker\bigl(\mathscr{H}_{\mathcal{J}}^*(\mathscr{A})\bigr) \supset \pi_{\mathcal{J}}^r\bigl(\ker(\mathscr{A})\bigr)$. Analogously, one shows that $\pi_{\mathcal{J}}^r\bigl(\ker(\mathscr{A}_n)\bigr) = \ker\bigl(\mathscr{H}_{\mathcal{J}}^*(\mathscr{A}_n)\bigr)$. Consequently, it holds that

$$\ker\bigl(\mathscr{H}_{\mathcal{J}}^*(\mathscr{A})\bigr) = \pi_{\mathcal{J}}^r\bigl(\ker(\mathscr{A})\bigr) \subset \pi_{\mathcal{J}}^r\bigl(\ker(\mathscr{A}_n)\bigr) = \ker\bigl(\mathscr{H}_{\mathcal{J}}^*(\mathscr{A}_n)\bigr)$$

$\mathbb{P}_n$-almost surely for all sufficiently large $n \in \mathbb{N}$. □

4.1.11 Lemma. Let $\mathscr{A}$ be some $(k \times k)$ positive semi-definite, symmetric matrix, $\kappa \in \mathrm{Im}(\mathscr{A})$ and define

$$g_{k,\mathcal{I}}(x) = \prod_{i \in \mathcal{I}} \mathbb{1}\bigl(\pi_{\{i\}}^k(x) \geq 0\bigr), \qquad x \in \mathbb{R}^k, \quad \emptyset \neq \mathcal{I} \subset \{1, \ldots, k\}.$$

Assume that $X_n \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} X$, as $n \to \infty$, where $X \sim \mathcal{N}(\kappa, \mathscr{A})$.

a) $g_{k,\mathcal{I}}(X_n)$ converges in distribution.

b) If additionally $X_n - \widehat{X}_n \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$, and $\rho^k_{\{i\}}(\mathscr{A}) > 0$, $i \in \mathcal{J}$, then $g_{k,\mathcal{J}}(X_n) - g_{k,\mathcal{J}}(\widehat{X}_n) \longrightarrow_{\mathbb{P}_n} 0$.

Proof. Note that $\mathbb{P}(X \in C) = 1$, where $C = \mathrm{Im}(\mathscr{A})$, cf. Witting [71, Hilfssatz 1.90], and define the $(k-1)$-dimensional hyperplanes $\{x \in \mathbb{R}^k \mid \pi^k_{\{i\}}(x) = 0\}$, $i = 1, \dots, k$. Set

$$\mathcal{M} = \Big\{ i \in \mathcal{J} \,\Big|\, C \subset \{x \in \mathbb{R}^k \mid \pi^k_{\{i\}}(x) = 0\} \Big\} \quad \text{and} \quad \widetilde{\mathcal{M}} = \mathcal{J} \backslash \mathcal{M},$$

where $\mathcal{M}$ contains the indices of the hyperplane that completely cover C. Clearly, it holds that $g_{k,\mathcal{J}} = g_{k,\widetilde{\mathcal{M}}}$ $\mathbb{P}^X$-almost surely, where $g_{k,\emptyset} = 1$.

Let $D(g_{k,\widetilde{\mathcal{M}}})$ denote the set of points where $g_{k,\widetilde{\mathcal{M}}}$ is not continuous. Obviously, we have the inclusion $D(g_{k,\widetilde{\mathcal{M}}}) \subset \bigcup_{i \in \widetilde{\mathcal{M}}} \{x \in \mathbb{R}^k \mid \pi^k_{\{i\}}(x) = 0\}$. $C \cap \{x \in \mathbb{R}^k \mid \pi^k_{\{i\}}(x) = 0\}$, $i \in \widetilde{\mathcal{M}}$, are linear sub-spaces of C whose dimension is strictly smaller than the dimension of C. Hence, we have

$$\mathbb{P}^X\Big(\{x \in \mathbb{R}^k \mid \pi^k_{\{i\}}(x) = 0\} \cap C \Big) = 0, \qquad i \in \widetilde{\mathcal{M}},$$

the Continuous Mapping Theorem, cf. Witting and Müller-Funk [72, Satz 5.43], yields the assertion.

Proof of b) by induction with respect to $|\mathcal{J}|$. For $|\mathcal{J}| = 1$ one gets that $\pi^k_{\mathcal{J}}(X) \sim \mathcal{N}\big(\pi^k_{\mathcal{J}}(\kappa), \rho^k_{\{i\}}(\mathscr{A})\big)$, and therefore $\mathbb{P}\big(\pi^k_{\mathcal{J}}(X) = 0\big) = 0$. Proposition B.4.7 implies the assertion. Assume that the assertion holds for all $|\mathcal{J}| \leq l$, $l \geq 1$. It is shown that the assertion holds for $|\mathcal{J}| = l + 1$. Choose non-empty sets $\mathcal{M}_i$, $i = 1, 2$, such that $\mathcal{M}_1 \cup \mathcal{M}_2 = \mathcal{J}$ and $\mathcal{M}_1 \cap \mathcal{M}_2 = \emptyset$. Clearly, it holds that $|\mathcal{M}_i| \leq l$. As

$$g_{k,\mathcal{J}}(X_n) - g_{k,\mathcal{J}}(\widehat{X}_n) = g_{k,\mathcal{M}_1}(X_n) \cdot g_{k,\mathcal{M}_2}(X_n) - g_{k,\mathcal{M}_1}(\widehat{X}_n) \cdot g_{k,\mathcal{M}_2}(\widehat{X}_n)$$

a), the induction assumption and Proposition B.4.6 yield the assertion. □

Let us summarize the premises of the previous results.

4.1.12 Assumption. Assume that the following conditions hold.

(i) $\mathscr{A}$ is some $(r \times r)$ positive semi-definite, symmetric matrix, $\rho^r_{\{i\}}(\mathscr{A}) > 0$, $i = 1, \ldots, r$.

(ii) $\kappa \in \operatorname{Im}(\mathscr{A})$.

(iii) $X_n \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} X$, as $n \to \infty$, where $X \sim \mathcal{N}(\kappa, \mathscr{A})$.

(iv) $\widehat{X}_n - X_n \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$.

(v) $\widehat{\mathscr{A}_n} - \mathscr{A} \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$.

(vi) $\ker(\mathscr{A}) \subset \ker(\widehat{\mathscr{A}_n})$ $\mathbb{P}_n$-almost surely for all sufficiently large $n \in \mathbb{N}$.

4.1.13 Theorem. Let Assumption 4.1.12 be satisfied. Then it holds that

a) $L_{\mathcal{J},1}(X_n, \mathscr{A}) - L_{\mathcal{J},1}(\widehat{X}_n, \widehat{\mathscr{A}_n}) \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$, and

b) $L_{\mathcal{J},1}(\widehat{X}_n, \widehat{\mathscr{A}_n}) \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} L_{\mathcal{J},1}(X, \mathscr{A})$, as $n \to \infty$,

where use the notation provided in Lemma 4.1.7.

Proof. First we show the estimate

$$\bigl|\max\{a_1, \ldots, a_l\} - \max\{b_1, \ldots, b_l\}\bigr| \le \max\bigl\{|a_1 - b_1|, \ldots, |a_l - b_l|\bigr\}. \quad (4.11)$$

Without loss of generality one can assume that $\max_i\{a_i\} = a_{i_0} \ge \max_i\{b_i\} = b_{i_1}$. The estimate

$$\bigl|\max_i\{a_i\} - \max_i\{b_i\}\bigr| = a_{i_0} - b_{i_0} + b_{i_0} - b_{i_1} \le a_{i_0} - b_{i_0} \le \max_i\bigl\{|a_i - b_i|\bigr\},$$

where we use $a_{i_0} - b_{i_0} \ge 0$ and $b_{i_0} - b_{i_1} \le 0$, gives the assertion. Because of the estimate (4.11) and Proposition B.4.5, it suffices to show that

$$f_{\mathcal{I},\mathcal{J}}(X_n, \mathscr{A}) - f_{\mathcal{I},\mathcal{J}}(X_n, \widehat{\mathscr{A}_n}) \longrightarrow_{\mathbb{P}_n} 0, \quad \text{as } n \to \infty, \qquad \emptyset \ne \mathcal{I} \subset \bigl\{1, \ldots, |\mathcal{J}|\bigr\}.$$

Using Lemma 4.1.10.c.i and Proposition B.4.6, one gets that $y_{\mathcal{I},\mathcal{J}}(X_n, \mathscr{A}) - y_{\mathcal{I},\mathcal{J}}(\widehat{X}_n, \widehat{\mathscr{A}_n}) \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$. Applying Lemma 4.1.10 and Proposition B.4.6 gives that $R_{\mathcal{I},\mathcal{J}}(X_n, \mathscr{A}) - R_{\mathcal{I},\mathcal{J}}(\widehat{X}_n, \widehat{\mathscr{A}_n}) \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$.

Now, Proposition B.4.6 yields that $Q_{\mathcal{I},\mathcal{J}}(X_n, \mathscr{A}) - Q_{\mathcal{I},\mathcal{J}}(\widehat{X}_n, \widehat{\mathscr{A}_n}) \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$. The Continuous Mapping Theorem, cf. Witting and Müller-Funk [72, Satz 5.43], yields that $Q_{\mathcal{I},\mathcal{J}}(X_n, \mathscr{A})$ converges in distribution as $n \to \infty$.

The Continuous Mapping Theorem gives that $R_{\mathcal{I},\mathcal{J}}(X_n, \mathscr{A}) \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} R_{\mathcal{I},\mathcal{J}}(X, \mathscr{A})$, as $n \to \infty$. One readily checks that

$$R_{\mathcal{I},\mathcal{J}}(X, \mathscr{A}) \sim \mathcal{N}\Bigg(\Big(\rho_{\mathcal{I}}^{|\mathcal{J}|}\big(\mathscr{H}_{\mathcal{J}}^*(\mathscr{A})\big)\Big)^{-} \pi_{\mathcal{I}}^{|\mathcal{J}|}\big(\mathscr{Y}_{\mathcal{J}}(\mathscr{A})^{\mathrm{T}}\kappa\big), \Big(\rho_{\mathcal{I}}^{|\mathcal{J}|}\big(\mathscr{H}_{\mathcal{J}}^*(\mathscr{A})\big)\Big)^{-}\Bigg)$$

and that

$$\Big(\rho_{\mathcal{I}}^{|\mathcal{J}|}\big(\mathscr{H}_{\mathcal{J}}^*(\mathscr{A})\big)\Big)^{-} \pi_{\mathcal{I}}^{|\mathcal{J}|}\big(\mathscr{Y}_{\mathcal{J}}(\mathscr{A})^{\mathrm{T}}\kappa\big) \in \mathrm{Im}\Bigg(\Big(\rho_{\mathcal{I}}^{|\mathcal{J}|}\big(\mathscr{H}_{\mathcal{J}}^*(\mathscr{A})\big)\Big)^{-}\Bigg),$$

where one uses Proposition B.3.4.c. Furthermore, Lemma 4.1.11 yields that $g_{|\mathcal{I}|,\{1,\dots,|\mathcal{I}|\}}\big(R_{\mathcal{I},\mathcal{J}}(X_n, \mathscr{A})\big)$ converges in distribution, as $n \to \infty$, and that

$$g_{|\mathcal{I}|,\{1,\dots,|\mathcal{I}|\}}\big(R_{\mathcal{I},\mathcal{J}}(X_n, \mathscr{A})\big) - g_{|\mathcal{I}|,\{1,\dots,|\mathcal{I}|\}}\big(R_{\mathcal{I},\mathcal{J}}(\widehat{X}_n, \widehat{\mathscr{A}_n})\big) \longrightarrow_{\mathbb{P}_n} 0, \quad \text{as } n \to \infty.$$

All in all, one gets that

$$\begin{aligned} f_{\mathcal{I},\mathcal{J}}(X_n, \mathscr{A}) - f_{\mathcal{I},\mathcal{J}}(X_n, \widehat{\mathscr{A}_n}) &= Q_{\mathcal{I},\mathcal{J}}(X_n, \mathscr{A}) \cdot g_{|\mathcal{I}|,\{1,\dots,|\mathcal{I}|\}}\big(R_{\mathcal{I},\mathcal{J}}(X_n, \mathscr{A})\big) \\ &\quad - Q_{\mathcal{I},\mathcal{J}}(\widehat{X}_n, \widehat{\mathscr{A}_n}) \cdot g_{|\mathcal{I}|,\{1,\dots,|\mathcal{I}|\}}\big(R_{\mathcal{I},\mathcal{J}}(\widehat{X}_n, \widehat{\mathscr{A}_n})\big) \longrightarrow_{\mathbb{P}_n} 0, \end{aligned}$$

as $n \to \infty$. The proof of a) is complete.

Proof of b). Because of Slutsky's Lemma, cf. Witting and Müller-Funk [72, Korollar 5.84], and a) we merely have to show that $L_{\mathcal{J},1}(X_n, \mathscr{A}) \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} L_{\mathcal{J},1}(X, \mathscr{A})$, as $n \to \infty$. Witting [71, Hilfssatz 1.90] yields that $\mathbb{P}\big(X \in \mathrm{Im}(\mathscr{A})\big) = 1$. Because of Lemma 4.1.7.b and Lemma 4.1.7.c the Continuous Mapping Theorem can be applied and yields the assertion. $\square$

The following result contains as a special case Theorem 3.2.7 of Behnen and Neuhaus [7]. The proof presented here relies on ideas that can be found in Behnen and Neuhaus [7] and an unpublished paper of Mayer on his dissertation [53].

4.1.14 Theorem. Assume that $X \sim \mathcal{N}\big(0, \mathscr{A}\big)$, where $\mathscr{A}$ is some $(r \times r)$ positive semi-definite, symmetric matrix. Set $\mathscr{H}_{\mathcal{J}}^* = \mathscr{H}_{\mathcal{J}}^*(\mathscr{A})$ and

$$F_{\mathcal{J},\mathscr{A}}(t) = \mathbb{P}\big(L_{\mathcal{J},1}(X, \mathscr{A}) \le t\big), \qquad t \in \mathbb{R}.$$

Furthermore, suppose that $\mathrm{rank}(\mathscr{H}_{\mathcal{J}}^*) > 0$.

a) Assume that $Y \sim \mathcal{N}(0, \mathscr{H}_{\mathcal{J}}^*)$ and that $Z_k \sim \chi_k^2$, where χ_k^2 denotes a central chi-squared distribution with k degrees of freedom. For all $t > 0$ it holds that

$$1 - F_{\mathcal{J},\mathscr{A}}(t) = \sum_{\emptyset \neq \mathcal{I} \subset \left\{1,\ldots,|\mathcal{J}|\right\}} \mathbb{P}\Big(Z_{\operatorname{rank}\left(\rho_{\mathcal{I}}^{\mathcal{J}}(\mathscr{H}_{\mathcal{J}}^*)\right)} > t\Big) \cdot \mathbb{P}\big(Y_{\mathcal{I}}^* \in \mathcal{V}_{\mathcal{I}}^+\big)$$
$$\times \mathbb{P}\Big(\bigcap\nolimits_{i \in \mathcal{I}^{\complement}} \big\{\pi_{\{i\}}^{|\mathcal{J}|}(Y - Y_{\mathcal{I}}^*) < 0\big\}\Big),$$

where we set

$$Y_{\mathcal{I}}^* = \mathscr{H}_{\mathcal{J}}^* \mathscr{T}_{\mathcal{I}}^{|\mathcal{J}|} \big(\rho_{\mathcal{I}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^*)\big)^- \pi_{\mathcal{I}}^{|\mathcal{J}|}(Y), \qquad \mathcal{I} \subset \big\{1, \ldots, |\mathcal{J}|\big\},$$

and

$$\mathcal{V}_{\mathcal{I}}^+ = \Big\{\mathscr{H}_{\mathcal{J}}^* \mathscr{T}_{\mathcal{I}}^{|\mathcal{J}|} \kappa \mid \kappa^{(u)} \geq 0,\, u = 1, \ldots, |\mathcal{I}|\Big\}, \qquad \mathcal{I} \subset \big\{1, \ldots, |\mathcal{J}|\big\},$$

and define $\mathbb{P}\Big(\bigcap_{i \in \emptyset} \{\ldots\}\Big) = 1$.

b) Assume that

$$Y_{\mathcal{I}} \sim \mathcal{N}\Big(0, \big(\rho_{\mathcal{I}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^*)\big)^{-1}\Big) \quad \text{and} \quad \widetilde{Y}_{\mathcal{I}} \sim \mathcal{N}\Big(0, \big(\rho_{\mathcal{I}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^{*-1})\big)^{-1}\Big),$$

then for all $t > 0$ it holds that

$$1 - F_{\mathcal{J},\mathscr{A}}(t) = \sum_{\emptyset \neq \mathcal{I} \subset \left\{1,\ldots,|\mathcal{J}|\right\}} \mathbb{P}\big(Z_{|\mathcal{I}|} > t\big) \cdot \mathbb{P}\Big(\bigcap\nolimits_{i \in \mathcal{I}} \big\{\pi_{\{i\}}^{|\mathcal{J}|}(Y_{\mathcal{I}}) \geq 0\big\}\Big)$$
$$\times \mathbb{P}\Big(\bigcap\nolimits_{i \in \mathcal{I}^{\complement}} \big\{\pi_{\{i\}}^{|\mathcal{J}|}(\widetilde{Y}_{\mathcal{I}}) < 0\big\}\Big),$$

where $\mathbb{P}\big(\bigcap_{i \in \emptyset} \{\ldots\}\big) = 1$.

c) It holds that $F_{\mathcal{J},\mathscr{A}}(0) \leq \frac{1}{2}$.

d) $F_{\mathcal{J},\mathscr{A}}$ is continuous and strictly increasing on the interval $(0, \infty)$.

Proof. For the proof we use the concepts and notation provided in Section B.2, especially Proposition B.2.5. Define

$$\mathcal{V}_{\mathcal{I}} = \Big\{\mathscr{H}_{\mathcal{J}}^* \mathscr{T}_{\mathcal{I}}^{|\mathcal{J}|} \kappa \mid \kappa \in \mathbb{R}^{|\mathcal{I}|}\Big\}$$

and

$$\mathcal{V}_{\mathcal{I}}^{\perp} = \Big\{ x \in \operatorname{Im}(\mathscr{H}_{\mathcal{J}}^{*}) \mid <x, y>_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}} = 0 \text{ for all } y \in \mathcal{V}_{\mathcal{I}} \Big\},$$

where $\mathcal{I} \subset \big\{1, \dots, |\mathcal{J}|\big\}$. Especially, we define $\mathcal{V}_{\emptyset} = \mathcal{V}_{\emptyset}^{+} = \{0\}$.

Set $Y = \mathscr{Y}_{\mathcal{J}}(\mathscr{A})^{\mathrm{T}} X$. Using Witting [71, Hilfssatz 1.90] and Proposition B.3.4.c gives that $\mathbb{P}\big(Y \in \operatorname{Im}(\mathscr{H}_{\mathcal{J}}^{*})\big) = 1$. Consequently, one gets that $L_{\mathcal{J},1}(X, \mathscr{A}) = \big|\big|\Pi_{\mathcal{V}_{\{1,\dots,|\mathcal{J}|\}}^{+}}(Y)\big|\big|_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}}^{2}$ $\mathbb{P}$-almost surely by applying Proposition B.2.5.e. Define

$$\Omega_{i,1} = \Big\{ \big\langle Y - \Pi_{\mathcal{V}_{\{1,\dots,|\mathcal{J}|\}}^{+}}(Y), \mathscr{H}_{\mathcal{J}}^{*} \mathscr{T}_{\{i\}}^{|\mathcal{J}|} \big\rangle_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}} = 0 \Big\}$$

and

$$\Omega_{i,2} = \Big\{ \big\langle Y - \Pi_{\mathcal{V}_{\{1,\dots,|\mathcal{J}|\}}^{+}}(Y), \mathscr{H}_{\mathcal{J}}^{*} \mathscr{T}_{\{i\}}^{|\mathcal{J}|} \big\rangle_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}} < 0 \Big\},$$

Proposition B.2.3.b implies that the sets

$$\begin{aligned}
\Omega_{\mathcal{I}} &= \Big(\bigcap\nolimits_{i \in \mathcal{I}} \Omega_{i,1}\Big) \cap \Big(\bigcap\nolimits_{i \in \mathcal{I}^{\complement}} \Omega_{i,2}\Big) \\
&= \Big\{ \Pi_{\mathcal{V}_{\{1,\dots,|\mathcal{J}|\}}^{+}}(Y) = \Pi_{\mathcal{V}_{\mathcal{I}}}(Y),\ \big\langle Y - \Pi_{\mathcal{V}_{\mathcal{I}}}(Y), \mathscr{H}_{\mathcal{J}}^{*} \mathscr{T}_{\{i\}}^{|\mathcal{J}|} \big\rangle_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}} < 0,\ i \in \mathcal{I}^{\complement} \Big\} \\
&= \Big\{ \Pi_{\mathcal{V}_{\{1,\dots,|\mathcal{J}|\}}^{+}}(Y) = \Pi_{\mathcal{V}_{\mathcal{I}}}(Y) \text{ and } \big\langle \Pi_{\mathcal{V}_{\mathcal{I}}^{\perp}}(Y), \mathscr{H}_{\mathcal{J}}^{*} \mathscr{T}_{\{i\}}^{|\mathcal{J}|} \big\rangle_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}} < 0,\ i \in \mathcal{I}^{\complement} \Big\},
\end{aligned}$$

$\mathcal{I} \subset \big\{1, \dots, |\mathcal{J}|\big\}$, are a disjoint decomposition of the sample space.

For $t > 0$ we get

$$\Big\{ \big|\big|\Pi_{\mathcal{V}_{\{1,\dots,|\mathcal{J}|\}}^{+}}(Y)\big|\big|_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}}^{2} > t \Big\} \cap \Omega_{\mathcal{I}} = A_{\mathcal{I}} \cap B_{\mathcal{I}} \cap C_{\mathcal{I}}.$$

where

$$A_{\mathcal{I}} = \Big\{ \big|\big|\Pi_{\mathcal{V}_{\mathcal{I}}}(Y)\big|\big|_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}}^{2} > t \Big\}, \qquad B_{\mathcal{I}} = \Big\{ \Pi_{\mathcal{V}_{\mathcal{I}}}(Y) \in \mathcal{V}_{\mathcal{I}}^{+} \Big\}$$

and

$$C_{\mathcal{I}} = \Big\{ \big\langle \Pi_{\mathcal{V}_{\mathcal{I}}^{\perp}}(Y), \mathscr{H}_{\mathcal{J}}^{*} \mathscr{T}_{\{i\}}^{|\mathcal{J}|} \big\rangle_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}} < 0,\ i \in \mathcal{I}^{\complement} \Big\}.$$

Noting that $Y \sim \mathcal{N}(0, \mathscr{H}_{\mathcal{J}}^{*})$ and using the representation given in Proposition B.2.5.c, one sees that Eaton [18, Proposition 3.4] is applicable and that

$\Pi_{\mathcal{V}_\mathrm{J}}(Y)$ and $\Pi^{\perp}_{\mathcal{V}_\mathrm{J}}(Y)$ are stochastically independent. Therefore, the events $A_\mathrm{J} \cap B_\mathrm{J}$ and C_J are stochastically independent.

In the next step it is proved that the events A_J and B_J are stochastically independent. Using Proposition B.2.5.c one gets that $\Pi_{\mathcal{V}_\mathrm{J}}(Y) = Y^*_\mathrm{J}$ and

$$\big|\big|\Pi_{\mathcal{V}_\mathrm{J}}(Y)\big|\big|^2_{(\mathscr{H}^*_{\mathcal{J}})^-} = \pi^{|\mathcal{J}|}_\mathrm{J}(Y)^\mathrm{T}\big(\rho^{|\mathcal{J}|}_\mathrm{J}(\mathscr{H}^*_{\mathcal{J}})\big)^-\rho^{|\mathcal{J}|}_\mathrm{J}(\mathscr{H}^*_{\mathcal{J}})\big(\rho^{|\mathcal{J}|}_\mathrm{J}(\mathscr{H}^*_{\mathcal{J}})\big)^-\pi^{|\mathcal{J}|}_\mathrm{J}(Y).$$

Note that $\big(\rho^{|\mathcal{J}|}_\mathrm{J}(\mathscr{H}^*_{\mathcal{J}})\big)^-\pi^{|\mathcal{J}|}_\mathrm{J}(Y) \sim \mathcal{N}\big(0, \big(\rho^{|\mathcal{J}|}_\mathrm{J}(\mathscr{H}^*_{\mathcal{J}})\big)^-\big)$. Consider the distribution family

$$\mathfrak{P} = \{P_c \mid c > 0\}, \qquad \text{where } P_c = \mathcal{N}\Big(0, c\cdot\big(\rho^{|\mathcal{J}|}_\mathrm{J}(\mathscr{H}^*_{\mathcal{J}})\big)^-\Big).$$

Clearly, it holds that

$$\frac{\mathrm{d}P_c}{\mathrm{d}P_1}(z) = c^{-\mathrm{rank}(\rho^{|\mathcal{J}|}_\mathrm{J}(\mathscr{H}^*_{\mathcal{J}}))/2}\cdot\exp\Big(-\frac{1}{2}\big(1/c-1\big)z^\mathrm{T}\rho^{|\mathcal{J}|}_\mathrm{J}(\mathscr{H}^*_{\mathcal{J}})z\Big).$$

One sees that $z^\mathrm{T}\rho^{|\mathcal{J}|}_\mathrm{J}(\mathscr{H}^*_{\mathcal{J}})z$ is a boundedly complete and sufficient statistic for the exponential family $\mathfrak{P}$, cf. Witting [71, Korollar 3.20, Satz 3.39]. As $\mathcal{V}^+_\mathrm{J}$ is a closed convex cone, see Definition B.2.2, it holds that

$$\begin{aligned}P_c\Big(\big\{z \mid \mathscr{H}^*_{\mathcal{J}}\mathscr{T}^{|\mathcal{J}|}_\mathrm{J}z \in \mathcal{V}^+_\mathrm{J}\big\}\Big) &= P_c\Big(\big\{z \mid 1/\sqrt{c}\cdot\mathscr{H}^*_{\mathcal{J}}\mathscr{T}^{|\mathcal{J}|}_\mathrm{J}z \in \mathcal{V}^+_\mathrm{J}\big\}\Big)\\ &= P_c\Big(\big\{z \mid \mathscr{H}^*_{\mathcal{J}}\mathscr{T}^{|\mathcal{J}|}_\mathrm{J}(1/\sqrt{c}\cdot z) \in \mathcal{V}^+_\mathrm{J}\big\}\Big) = P_1\Big(\big\{z \mid \mathscr{H}^*_{\mathcal{J}}\mathscr{T}^{|\mathcal{J}|}_\mathrm{J}z \in \mathcal{V}^+_\mathrm{J}\big\}\Big).\end{aligned}$$

Thus, the distribution of the auxiliary statistic $\mathbb{1}\big(\mathscr{H}^*_{\mathcal{J}}\mathscr{T}^{|\mathcal{J}|}_\mathrm{J}z \in \mathcal{V}^+_\mathrm{J}\big)$ is independent of the parameter c. Basu's Theorem, cf. Lehmann [47, Theorem 2, p. 191], gives that $z^\mathrm{T}\rho^{|\mathcal{J}|}_\mathrm{J}(\mathscr{H}^*_{\mathcal{J}})z$ and $\mathbb{1}\big(\mathscr{H}^*_{\mathcal{J}}\mathscr{T}^{|\mathcal{J}|}_\mathrm{J}z \in \mathcal{V}^+_\mathrm{J}\big)$ are stochastically independent. Consequently, the events A_J and B_J are stochastically independent.

So far, we have proved that

$$\mathbb{P}\big(L_{\mathcal{J},1}(X,\mathscr{A}) > t\big) = \sum_{\emptyset\neq\mathrm{J}\subset\{1,\dots,|\mathcal{J}|\}}\mathbb{P}(A_\mathrm{J})\cdot\mathbb{P}(B_\mathrm{J})\cdot\mathbb{P}(C_\mathrm{J}), \qquad t > 0,$$

where we set $\mathbb{P}(C_{\{1,\dots,|\mathcal{J}|\}}) = 1$.

According to Proposition B.3.3.b we can find $(|\mathbb{J}| \times l)$ matrices $\mathscr{B}$ and $\mathscr{C}$, where $l = \operatorname{rank}\big(\rho_{\mathbb{J}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^{*})\big)$, such that $\mathscr{B}\mathscr{B}^{\mathrm{T}} = \rho_{\mathbb{J}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^{*})$, $\mathscr{C}\mathscr{C}^{\mathrm{T}} = \big(\rho_{\mathbb{J}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^{*})\big)^{-}$ and $\mathscr{C}^{\mathrm{T}}\mathscr{B} = \mathscr{B}^{\mathrm{T}}\mathscr{C} = \mathscr{E}_l$, where $\mathscr{E}_l$ denoted the $(l \times l)$ unity matrix. As

$$\big|\big|\Pi_{\mathcal{V}_{\mathbb{J}}}(Y)\big|\big|^2_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}} = \pi_{\mathbb{J}}^{|\mathcal{J}|}(Y)^{\mathrm{T}}\big(\rho_{\mathbb{J}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^{*})\big)^{-}\pi_{\mathbb{J}}^{|\mathcal{J}|}(Y) = \pi_{\mathbb{J}}^{|\mathcal{J}|}(Y)^{\mathrm{T}}\mathscr{C}\mathscr{C}^{\mathrm{T}}\pi_{\mathbb{J}}^{|\mathcal{J}|}(Y)$$

and $\mathscr{C}^{\mathrm{T}}\pi_{\mathbb{J}}^{|\mathcal{J}|}(Y) \sim \mathcal{N}(0, \mathscr{E}_l)$, it follows that $\big|\big|\Pi_{\mathcal{V}_{\mathbb{J}}}(Y)\big|\big|^2_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}}$ is χ_l^2-distributed.

Using $\Pi_{\mathcal{V}_{\mathbb{J}}}(Y) = Y_{\mathbb{J}}^{*}$, $\Pi_{\mathcal{V}_{\mathbb{J}}^{\perp}}(Y) = Y - \Pi_{\mathcal{V}_{\mathbb{J}}}(Y)$ and $\mathbb{P}\big(Y \in \operatorname{Im}(\mathscr{H}_{\mathcal{J}}^{*})\big) = 1$, one receives the representation of the sets $B_{\mathbb{J}}$ and $C_{\mathbb{J}}$.

Proof of b). In the case that $\operatorname{rank}\big(\mathscr{H}_{\mathcal{J}}^{*}\mathscr{T}_{\mathbb{J}}^{|\mathcal{J}|}\big) = |\mathbb{J}|$, the condition $Y_{\mathbb{J}}^{*} \in \mathcal{V}_{\mathbb{J}}^{+}$ is equivalent to $\pi_{\{i\}}^{|\mathbb{J}|}\big(\big(\rho_{\mathbb{J}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^{*})\big)^{-}\pi_{\mathbb{J}}^{|\mathcal{J}|}(Y)\big) \geq 0$, $i \in \mathbb{J}$. As

$$\big(\rho_{\mathbb{J}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^{*})\big)^{-}\pi_{\mathbb{J}}^{|\mathcal{J}|}(Y) \sim \mathcal{N}\Big(0, \big(\rho_{\mathbb{J}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^{*})\big)^{-1}\Big),$$

we get the new representation of the set $B_{\mathbb{J}}$.

Note that $\mathscr{H}_{\mathcal{J}}^{*}$ is not degenerated, therefore one readily checks that $\mathcal{V}_{\mathbb{J}}^{\perp} = \{\mathscr{T}_{\mathbb{J}^{\complement}}^{|\mathcal{J}|}\kappa \mid \kappa \in \mathbb{R}^{|\mathbb{J}|}\}$. We know that $\Pi_{\mathcal{V}_{\mathbb{J}}^{\perp}}(Y) = \big(\rho_{\mathbb{J}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^{*-1})\big)^{-1}\pi_{\mathbb{J}^{\complement}}^{|\mathcal{J}|}(Y)$, see Proposition B.2.5.d. Because of

$$\Pi_{\mathcal{V}_{\mathbb{J}}^{\perp}}(Y) \sim \mathcal{N}\Big(0, \big(\rho_{\mathbb{J}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^{*-1})\big)^{-1}\Big),$$

it results the representation of the set $C_{\mathbb{J}}$.

Proof of c). As rank $(\mathscr{H}_{\mathcal{J}}^{*}) > 0$, one can choose i_0, such that $\rho_{\{i\}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^{*}) > 0$. The estimate

$$\begin{aligned}\mathbb{P}\big(L_{\mathcal{J},1}(X,\mathscr{A}) = 0\big) &= \mathbb{P}\Big(\bigcap\nolimits_{\emptyset \neq \mathbb{J} \subset \{1,\ldots,|\mathcal{J}|\}}\big\{f_{\mathbb{J},\mathcal{J}}(X,\mathscr{A}) = 0\big\}\Big)\\ &\leq \mathbb{P}\big(f_{\mathbb{J},\{i_0\}}(X,\mathscr{A}) = 0\big) = \mathbb{P}\Big(\big(\rho_{\{i_0\}}^{|\mathcal{J}|}(\mathscr{H}_{\mathcal{J}}^{*})\big)^{-1}\pi_{\{i_0\}}^{|\mathcal{J}|}(Y) \leq 0\Big) = \frac{1}{2}\end{aligned}$$

gives the assertion.

d) is an immediate consequence of the representation of $1 - F_{\mathcal{J},\mathscr{A}}$ given in a), $\operatorname{rank}(\mathscr{H}_{\mathcal{J}}^{*}) > 0$ and the fact that $t \mapsto \mathbb{P}(Z_k > t)$, $k > 0$, is continuous and strictly decreasing on the interval $(0, \infty)$. □

In Behnen and Neuhaus, cf. [7, pp. 158], the computation of the survival function $1 - F_{\mathfrak{J},\mathscr{A}}$ is demonstrated for the cases $|\mathfrak{J}| = 2$ and $|\mathfrak{J}| = 3$.

4.1.15 Corollary. Under Assumption 4.1.12, it holds that

$$c_{\mathfrak{J},1}(\alpha, \widehat{\mathscr{A}_n}) - F^{-1}_{\mathfrak{J},\mathscr{A}}(1-\alpha) \longrightarrow_{\mathbb{P}_n} 0, \qquad \text{as } n \to \infty,$$

where $\alpha \in (0, 1/2)$ and one sets $c_{\mathfrak{J},1}(\alpha, \widehat{\mathscr{A}_n}) = F^{-1}_{\mathfrak{J},\widehat{\mathscr{A}_n}}(1-\alpha)$, $n \in \mathbb{N}$.

Proof. Without loss of generality we can assume that all random variables are defined on the same probability space. Using the sub-sub-sequence principle for random variables that converge in probability, cf. Bauer [6, Korollar 20.8], we can also assume that $\widehat{\mathscr{A}_n} \to \mathscr{A}$ almost surely.

Assume that $\widetilde{X}_n \sim \mathcal{N}(0, \widehat{\mathscr{A}_n})$. As $\widetilde{X}_n \xrightarrow{\mathfrak{D}} X$, using Theorem 4.1.13.b yields that $L_{\mathfrak{J},1}(\widetilde{X}_n, \widehat{\mathscr{A}_n}) \xrightarrow{\mathfrak{D}} L_{\mathfrak{J},1}(X, \mathscr{A})$. Theorem 4.1.14.d and Witting and Müller-Funk, cf. [72, Satz 5.58], give that $F_{\mathfrak{J},\widehat{\mathscr{A}_n}}(t) \to F_{\mathfrak{J},\mathscr{A}}(t)$ almost surely for all $t > 0$. Theorem 4.1.14.c and Theorem 4.1.14.d imply that $F^{-1}_{\mathfrak{J},\mathscr{A}}$ is continuous on $(1/2, 1)$. Witting and Müller-Funk [72, Satz 5.76] and the sub-sub-sequence principle for random variables that converge in probability give the result. □

4.1.16 Assumption. Let the Assumption 4.1.1 hold and suppose that $\widehat{U}_n : \Omega_n \longrightarrow \mathbb{R}^r$, $n \in \mathbb{N}$, are measurable mappings and that $\widehat{V}_n : \Omega_n \longrightarrow \mathbb{R}^{r \times r}$, $n \in \mathbb{N}$, are random matrices satisfying the following conditions.

a) $\widehat{U}_n - U_n \longrightarrow_{P_{n,0}} 0$, as $n \to \infty$, where $U_n = \mathscr{U}(\mathscr{J})^{\mathrm{T}} S_n$, $n \in \mathbb{N}$.

b) $\widehat{V}_n - \mathscr{J}^* \longrightarrow_{P_{n,0}} 0$, as $n \to \infty$.

c) $\ker(\mathscr{J}^*) \subset \ker(\widehat{V}_n)$ $P_{n,0}$-almost surely for all sufficiently large $n \in \mathbb{N}$.

d) $\rho^r_{\{i\}}(\mathscr{J}^*) > 0$, $i = 1, \ldots, r$.

For the testing problem $\widetilde{\mathcal{H}}^{\mathfrak{J}}_1$ versus $\widetilde{\mathcal{K}}^{\mathfrak{J}}_1$, we propose the sequence of tests $\varphi_{n,1}$, $n \in \mathbb{N}$, where

$$\varphi_{n,1} = \begin{cases} 1, & \\ 0, & \end{cases} L_{\mathfrak{J},1}(\widehat{U}_n, \widehat{V}_n) - c_{\mathfrak{J},1}(\alpha, \widehat{V}_n) \begin{matrix} > \\ \leq \end{matrix} 0 \,.$$

For a practical application the following representation of the test is more convenient

$$\varphi_{n,1} = \begin{cases} 1, \\ 0, \end{cases} F_{\mathcal{J},\widehat{V}_n}\big(L_{\mathcal{J},1}(\widehat{U}_n, \widehat{V}_n)\big) - (1-\alpha) \begin{array}{c} > \\ \leq \end{array} 0 \, .$$

4.1.17 Corollary. If $\alpha \in (0, 1/2)$ and Assumption 4.1.16 holds, it holds that

$$\mathbb{E}_{n,\xi}\big(\varphi_{n,1}\big) \to P_\xi\big(L_{\mathcal{J},1}(U, \mathscr{J}^*) - c_{\mathcal{J},1}(\alpha, \mathscr{J}^*) > 0\big),$$

where $U = \mathscr{U}(\mathscr{J})^{\mathrm{T}} S$ and $\mathfrak{L}\big(U \mid P_\xi\big) = \mathcal{N}(\mathscr{J}^*\beta, \mathscr{J}^*)$. Especially, the sequence of tests $\varphi_{n,1}$, $n \in \mathbb{N}$, keeps asymptotically the level on the hypothesis $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$. Moreover, if

$$\kappa_1^{\mathrm{T}} \mathscr{H}_{\mathcal{J}}^*(\mathscr{J}^*)\kappa_2 \geq 0, \qquad \text{for all } \kappa_i \in \mathbb{R}^{|\mathcal{J}|},\ \kappa_i \geq 0,\ i = 1, 2, \tag{4.12}$$

then the sequence of tests $\varphi_{n,1}$, $n \in \mathbb{N}$, is asymptotically unbiased.

Proof. The Continuous Mapping Theorem, cf. Witting and Müller-Funk [72, Satz 5.43] gives that $U_n \xrightarrow{\mathfrak{D}}_{P_{n,\xi}} U$, where one checks that $U \sim \mathcal{N}(\mathscr{J}^*\beta, \mathscr{J}^*)$, since

$$\mathscr{U}(\mathscr{J})^{\mathrm{T}} \mathscr{J} \xi = \mathscr{J}^*\beta + \big(\mathscr{J}_{1,2} - \mathscr{J}_{1,2}\mathscr{J}_{2,2}^{-}\mathscr{J}_{2,1}\big)\eta = \mathscr{J}^*\beta,$$

see Proposition B.3.4.b, and $\mathscr{U}(\mathscr{J})^{\mathrm{T}} \mathscr{J} \mathscr{U}(\mathscr{J}) = \mathscr{J}^*$. Setting $\mathscr{A} = \mathscr{J}^*$, $\widehat{\mathscr{A}_n} = \widehat{V}_n$, $X_n = U_n$, $\widehat{X}_n = \widehat{U}_n$ and $\mathbb{P}_n = P_{n,0}$, $n \in \mathbb{N}$, one sees that Assumption 4.1.12 holds. Theorem 4.1.13.a, Corollary 4.1.15 and Remark 4.1.2 give that

$$L_{\mathcal{J},1}(\widehat{U}_n, \widehat{V}_n) - c_{\mathcal{J},1}(\alpha, \widehat{V}_n) - \big(L_{\mathcal{J},1}(U_n, \mathscr{J}^*) - c_{\mathcal{J},1}(\alpha, \mathscr{J}^*)\big) \longrightarrow_{P_{n,\xi}} 0.$$

Setting $\mathscr{A} = \widehat{\mathscr{A}_n} = \mathscr{J}^*$, $X_n = \widehat{X}_n = U_n$, and $\mathbb{P}_n = P_{n,\xi}$, $n \in \mathbb{N}$, one sees that Assumption 4.1.12 holds. Theorem 4.1.13.b and Slutsky's Lemma, cf. Witting and Müller-Funk [72, Korollar 5.84], yield that

$$L_{\mathcal{J},1}\big(\widehat{U}_n, \widehat{V}_n\big) - c_{\mathcal{J},1}\big(\alpha, \widehat{V}_n\big) \xrightarrow{\mathfrak{D}}_{P_{n,\xi}} L_{\mathcal{J},1}\big(U, \mathscr{J}^*\big) - c_{\mathcal{J},1}\big(\alpha, \mathscr{J}^*\big).$$

Theorem 4.1.14.c and Theorem 4.1.14.d imply that

$$P_0\Big(L_{\mathfrak{J},1}(U, \mathscr{J}^*) - c_{\mathfrak{J},1}(\alpha, \mathscr{J}^*) = 0\Big) = 0.$$

As $P_\xi \ll P_0$, the Portmanteau Theorem, cf. Witting and Müller-Funk [72, Satz 5.40], gives the first assertion.

Abbreviating $\mathscr{H}^{\mathfrak{J}}_{i,j} = \mathscr{H}^{\mathfrak{J}}_{i,j}(\mathscr{J}^*)$ and $\mathscr{H}^*_{\mathfrak{J}} = \mathscr{H}^*_{\mathfrak{J}}(\mathscr{J}^*)$, one readily shows that

$$\mathscr{Y}_{\mathfrak{J}}(\mathscr{J}^*)^{\mathrm{T}} \mathscr{J}^* \beta = \mathscr{H}^*_{\mathfrak{J}} \beta_{\mathfrak{J}} + \Big(\mathscr{H}^{\mathfrak{J}}_{1,2} - \mathscr{H}^{\mathfrak{J}}_{1,2}\big(\mathscr{H}^{\mathfrak{J}}_{2,2}\big)^- \mathscr{H}^{\mathfrak{J}}_{2,2}\Big)\beta_{\mathfrak{J}^\complement} = \mathscr{H}^*_{\mathfrak{J}} \beta_{\mathfrak{J}}$$

where one uses Proposition B.3.4.b. Thus, it holds that $\mathfrak{L}(\mathscr{Y}_{\mathfrak{J}}(\mathscr{J}^*)^{\mathrm{T}} U \mid P_\xi) = \mathcal{N}(\mathscr{H}^*_{\mathfrak{J}} \beta_{\mathfrak{J}}, \mathscr{H}^*_{\mathfrak{J}})$.

Assume that $\xi \in \widetilde{\mathcal{H}}^{\mathfrak{J}}_1$. As $\mathscr{J}\xi = \mathscr{J}\xi'$ is equivalent to $P_\xi = P_{\xi'}$, cf. Lemma 4.1.4, without loss of generality it can be assumed that $\xi \in \mathcal{H}^{\mathfrak{J}}_1$. Using the previous considerations and Lemma 4.1.7.b, we get that

$$\begin{aligned} \mathfrak{L}\big(L_{\mathfrak{J},1}(U, \mathscr{J}^*) \mid P_\xi\big) &= \mathfrak{L}\Big(L_{\mathfrak{J},\mathscr{J}^*}\big(\mathscr{Y}_{\mathfrak{J}}(\mathscr{J}^*)^{\mathrm{T}} U\big) \mid P_\xi\Big) \\ &= \mathfrak{L}\Big(L_{\mathfrak{J},\mathscr{J}^*}\big(\mathscr{Y}_{\mathfrak{J}}(\mathscr{J}^*)^{\mathrm{T}} U\big) \mid P_0\Big) = \mathfrak{L}\big(L_{\mathfrak{J},1}(U, \mathscr{J}^*) \mid P_0\big). \end{aligned}$$

Now, Theorem 4.1.14 yields that

$$\begin{aligned} P_\xi\big(L_{\mathfrak{J},1}(U, \mathscr{J}^*) > c_{\mathfrak{J},1}(\alpha, \mathscr{J}^*)\big) &= P_0\big(L_{\mathfrak{J},1}(U, \mathscr{J}^*) > c_{\mathfrak{J},1}(\alpha, \mathscr{J}^*)\big) \\ &= 1 - F_{\mathfrak{J},\mathscr{J}^*}\big(F^{-1}_{\mathfrak{J},\mathscr{J}^*}(1-\alpha)\big) = \alpha. \end{aligned} \tag{4.13}$$

So, it remains to be proved that the sequence of tests is asymptotically unbiased. First, we note that

$$\begin{aligned} &P_\xi\big(L_{\mathfrak{J},1}(U, \mathscr{J}^*) > c_{\mathfrak{J},1}(\alpha, \mathscr{J}^*)\big) \\ &\qquad = P_\xi\Big(L_{\mathfrak{J},\mathscr{J}^*}\big(\mathscr{Y}_{\mathfrak{J}}(\mathscr{J}^*)^{\mathrm{T}} U\big) > c_{\mathfrak{J},1}(\alpha, \mathscr{J}^*)\Big) \\ &\qquad = P_0\Big(L_{\mathfrak{J},\mathscr{J}^*}\big(\mathscr{Y}_{\mathfrak{J}}(\mathscr{J}^*)^{\mathrm{T}} U + \mathscr{H}^*_{\mathfrak{J}} \beta_{\mathfrak{J}}\big) > c_{\mathfrak{J},1}(\alpha, \mathscr{J}^*)\Big), \end{aligned} \tag{4.14}$$

cf. Lemma 4.1.7.b. Furthermore, it holds that

$$
\begin{aligned}
L_{\mathcal{J},\mathscr{J}^*}\big(\mathscr{Y}_{\mathcal{J}}(\mathscr{J}^*)^{\mathrm{T}}u + \mathscr{H}^*_{\mathcal{J}}\beta_{\mathcal{J}}\big)&\\
&= \sup_{\kappa\geq 0,\, \kappa\in\mathbb{R}^{|\mathcal{J}|}} \Big(\kappa^{\mathrm{T}}\big(\mathscr{Y}_{\mathcal{J}}(\mathscr{J}^*)^{\mathrm{T}}u + \mathscr{H}^*_{\mathcal{J}}\beta_{\mathcal{J}}\big) - \frac{1}{2}\kappa^{\mathrm{T}}\mathscr{H}^*_{\mathcal{J}}\kappa\Big)\\
&\geq \sup_{\kappa\geq 0,\, \kappa\in\mathbb{R}^{|\mathcal{J}|}} \Big(\kappa^{\mathrm{T}}\mathscr{Y}_{\mathcal{J}}(\mathscr{J}^*)^{\mathrm{T}}u - \frac{1}{2}\kappa^{\mathrm{T}}\mathscr{H}^*_{\mathcal{J}}\kappa\Big),
\end{aligned}
\tag{4.15}
$$

where we use the condition stated in (4.12). Using again Lemma 4.1.7.b as well as combining (4.13), (4.14), (4.15), we finally get that

$$P_\xi\big(L_{\mathcal{J},1}(U, \mathscr{J}^*) > c_{\mathcal{J},1}(\alpha, \mathscr{J}^*)\big) \geq \alpha$$

for all $\xi \in \mathcal{H}_1^{\mathcal{J}}$. □

In the last part of this section, it is shown that the sequence of tests $\varphi_{n,1}$, $n \in \mathbb{N}$, is asymptotically admissible for the testing problem $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathcal{J}}$. But first let us remember the notion of admissibility.

4.1.18 Definition. a) In the limit model, a test ϕ' is said to be admissible for the testing problem $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathcal{J}}$, if for any other test ϕ satisfying

$$\mathbb{E}_\xi(\phi) \leq \mathbb{E}_\xi(\phi'), \quad \xi \in \widetilde{\mathcal{H}}_1^{\mathcal{J}}, \qquad \text{and} \qquad \mathbb{E}_\xi(\phi) \geq \mathbb{E}_\xi(\phi'), \quad \xi \in \widetilde{\mathcal{K}}_1^{\mathcal{J}},$$

it follows that $\phi = \phi'$ P_ξ-almost everywhere for all $\xi \in \mathbb{R}^{r+q}$.

b) A sequence of tests ϕ'_n, $n \in \mathbb{N}$, is said to be asymptotically admissible for the testing problem $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathcal{J}}$, if for any other sequence of tests ϕ_n, $n \in \mathbb{N}$, satisfying

$$\limsup_{n\to\infty} \mathbb{E}_{n,\xi}(\phi_n) \leq \liminf_{n\to\infty} \mathbb{E}_{n,\xi}(\phi'_n), \qquad \xi \in \widetilde{\mathcal{H}}_1^{\mathcal{J}},$$

and

$$\liminf_{n\to\infty} \mathbb{E}_{n,\xi}(\phi_n) \geq \limsup_{n\to\infty} \mathbb{E}_{n,\xi}(\phi'_n), \qquad \xi \in \widetilde{\mathcal{K}}_1^{\mathcal{J}},$$

holds that $\phi_n - \phi'_n \longrightarrow_{P_{n,\xi}} 0$.

The following result will be essential for proving that a sequence of tests for multivariate one-sided testing problems is admissible.

4.1.19 Theorem. Let $\xi_k \in \mathbb{R}^{r+q}$, $k \in \mathbb{N}$, and $a_k \in \mathbb{R}$, $k \in \mathbb{N}$, be arbitrary. If $\xi_k \in \widetilde{\mathcal{K}}_1^{\mathfrak{J}}$, $k \in \mathbb{N}$, then the test $\phi'(S)$, where

$$\phi'(s) = \begin{cases} 1, & s^{\mathrm{T}}\xi_k > a_k \text{ for some } k \in \mathbb{N}, \\ 0, & s^{\mathrm{T}}\xi_k \leq a_k \text{ for all } k \in \mathbb{N}, \end{cases}$$

is admissible for the testing problem $\widetilde{\mathcal{H}}_1^{\mathfrak{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathfrak{J}}$ and uniquely determined by its distribution.

Proof. The proof of Theorem 71.14 in Strasser [68] is also applicable for this Theorem. The crucial point is the fact that the ξ_k, $k \in \mathbb{N}$, belong to the alternative $\widetilde{\mathcal{K}}_1^{\mathfrak{J}}$. □

4.1.20 Proposition. Under Assumption 4.1.16, the test $\phi_1'(S)$, where

$$\phi_1'(s) = \begin{cases} 1, \\ 0, \end{cases} L_{\mathfrak{J},1}\big(\mathscr{U}(\mathscr{J})^{\mathrm{T}}s, \mathscr{J}^*\big) - c_{\mathfrak{J},1}(\alpha, \mathscr{J}^*) \begin{array}{c} > \\ \leq \end{array} 0,$$

and $\alpha \in (0, 1/2)$, is admissible for $\widetilde{\mathcal{H}}_1^{\mathfrak{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathfrak{J}}$ and uniquely determined by its distribution.

Proof. We show that ϕ_1' has a representation as the test considered in Theorem 4.1.19. Set $\mathscr{H}_{\mathfrak{J}}^* = \mathscr{H}_{\mathfrak{J}}^*(\mathscr{J}^*)$. The function $L_{\mathfrak{J},\mathscr{J}^*} : \mathrm{Im}(\mathscr{H}_{\mathfrak{J}}^*) \longrightarrow \mathbb{R}$,

$$L_{\mathfrak{J},\mathscr{J}^*}(y) = 2 \sup_{\kappa \geq 0} \left(\kappa^{\mathrm{T}} y - \frac{1}{2} \kappa^{\mathrm{T}} \mathscr{H}_{\mathfrak{J}}^* \kappa \right), \qquad y \in \mathrm{Im}(\mathscr{H}_{\mathfrak{J}}^*),$$

is a continuous and convex function according to Lemma 4.1.7.c. Thus, the set $C^* = \big\{L_{\mathfrak{J},\mathscr{J}^*}(\cdot) \leq c_{\mathfrak{J},1}(\alpha, \mathscr{J}^*)\big\}$ is convex and closed. By basic separation theorems we know that there exists $\kappa_k \in \mathrm{Im}(\mathscr{H}_{\mathfrak{J}}^*)\backslash\{0\}$ and $a_k \in \mathbb{R}$, $k \in \mathbb{N}$, such that

$$C^* = \bigcap_{k=1}^{\infty} \big\{ y \in \mathrm{Im}(\mathscr{H}_{\mathfrak{J}}^*) \mid \kappa_k^{\mathrm{T}} y \leq a_k \big\}.$$

In the following we use the concept provided in Section B.2. Let us introduce some notation. Set $\mathcal{V} = \mathrm{Im}(\mathscr{H}_{\mathcal{J}}^{*})$ and $<y_1, y_2>_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}} = y_1^{\mathrm{T}}(\mathscr{H}_{\mathcal{J}}^{*})^{-}y_2$, $y_i \in \mathrm{Im}(\mathscr{H}_{\mathcal{J}}^{*})$. $\big(\mathcal{V}, <\cdot,\cdot>_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}}\big)$, is a Hilbert space and $\mathcal{V}_0^{+} = \big\{\mathscr{H}_{\mathcal{J}}^{*}\kappa \mid \kappa \geq 0\big\}$ a closed convex cone, see Proposition B.2.5. It holds that

$$\big\{y \in \mathrm{Im}(\mathscr{H}_{\mathcal{J}}^{*}) \mid y^{\mathrm{T}}\kappa_k \leq a_k\big\} = \big\{y \in \mathcal{V} \mid <y, \mathscr{H}_{\mathcal{J}}^{*}\kappa_k>_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}} \leq a_k\big\}. \tag{4.16}$$

In the next step we show that $\mathscr{H}_{\mathcal{J}}^{*}\kappa_k \in \mathcal{V}_0^{+}$, which is clearly equivalent to $\Pi_{\mathcal{V}_0^{+}}(\mathscr{H}_{\mathcal{J}}^{*}\kappa_k) = \mathscr{H}_{\mathcal{J}}^{*}\kappa_k$, where $\Pi_{\mathcal{V}_0^{+}}$ denotes the projection on $\mathcal{V}_0^{+}$ in the sense of Proposition B.2.3.a. Assume that $\widetilde{y} = \lambda \cdot \big(\mathscr{H}_{\mathcal{J}}^{*}\kappa_k - \Pi_{\mathcal{V}_0^{+}}(\mathscr{H}_{\mathcal{J}}^{*}\kappa_k)\big) \in \mathcal{V}\backslash\{0\}$. It holds that

$$<\widetilde{y}, \mathscr{H}_{\mathcal{J}}^{*}\kappa_k>_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}} = \lambda\big|\big|\mathscr{H}_{\mathcal{J}}^{*}\kappa_k - \Pi_{\mathcal{V}_0^{+}}(\mathscr{H}_{\mathcal{J}}^{*}\kappa_k)\big|\big|^2_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}},$$

where we use Proposition B.2.3.b. Using equation (4.16), for all sufficiently large $\lambda > 0$ it holds that $\widetilde{y}^{\mathrm{T}}\kappa_k > a_k$, *i.e.* $\widetilde{y} \notin C^{*}$. On the other hand, it holds that

$$L_{\mathcal{J},\mathscr{J}^{*}}(\widetilde{y}) = \big|\big|\Pi_{\mathcal{V}_0^{+}}(\widetilde{y})\big|\big|^2_{(\mathscr{H}_{\mathcal{J}}^{*})^{-}} = 0,$$

where we use Proposition B.2.4 and Proposition B.2.5.b. $L_{\mathcal{J},\mathscr{J}^{*}}(\widetilde{y}) = 0$ means $\widetilde{y} \in C^{*}$, this is a contradiction. Consequently, there exists $\widetilde{\kappa}_k \in \mathbb{R}^{|\mathcal{J}|}$, such that $\widetilde{\kappa}_k^{(u)} \geq 0$, and $\mathscr{H}_{\mathcal{J}}^{*}\kappa_k = \mathscr{H}_{\mathcal{J}}^{*}\widetilde{\kappa}_k$, where we use the Definition of $\mathcal{V}_0^{+}$. Moreover, it holds that

$$\big\{y \in \mathrm{Im}(\mathscr{H}_{\mathcal{J}}^{*}) \mid y^{\mathrm{T}}\kappa_k \leq a_k\big\} = \big\{y \in \mathrm{Im}(\mathscr{H}_{\mathcal{J}}^{*}) \mid y^{\mathrm{T}}\widetilde{\kappa}_k \leq a_k\big\},$$

which can be derived by applying equation (4.16) twice. Remember that we want to show that ϕ_1' has a representation as the test considered in Theorem 4.1.19. Therefore, we set

$$\xi_k = \begin{pmatrix} \mathscr{Y}_{\mathcal{J}}(\mathscr{J}^{*})\widetilde{\kappa}_k \\ \mathscr{J}_{2,2}^{-}\mathscr{J}_{2,1}\mathscr{Y}_{\mathcal{J}}(\mathscr{J}^{*})\widetilde{\kappa}_k \end{pmatrix}, \qquad k \in \mathbb{N},$$

and note that $\pi_{\mathcal{J}}^{r+q}(\xi_k) = \widetilde{\kappa}_k \geq 0$. Therefore, it holds that $\xi_k \in \mathcal{K}_1^{\mathcal{J}}$.

Thus, we have that $\mathscr{Y}_{\mathcal{J}}(\mathscr{J}^{*})^{\mathrm{T}}\mathscr{U}(\mathscr{J})^{\mathrm{T}}\mathscr{J}\xi = \mathscr{H}_{\mathcal{J}}^{*}\beta_{\mathcal{J}} = 0$ for all $\xi \in \widetilde{\mathcal{H}}_1^{\mathcal{J}}$, where we use the Definition of $\mathcal{H}_1^{\mathcal{J}}$ and $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$. We also have that $\mathscr{H}_{\mathcal{J}}^{*}\widetilde{\kappa}_k, = \mathscr{H}_{\mathcal{J}}^{*}\kappa_k \neq 0$, where we use that $\kappa_k \in \mathrm{Im}(\mathscr{H}_{\mathcal{J}}^*)\backslash\{0\}$. Thus, we get that $\xi_k \in \widetilde{\mathcal{K}}_1^{\mathcal{J}}$. Finally, it holds that

$$\begin{aligned}
&\bigcap_{k=1}^{\infty} \bigl\{ s \in \mathrm{Im}(\mathscr{J}) \mid \xi_k^{\mathrm{T}} s \leq a_k \bigr\} \\
&\qquad = \bigcap_{k=1}^{\infty} \bigl\{ s \in \mathrm{Im}(\mathscr{J}) \mid s^{\mathrm{T}} \mathscr{U}(\mathscr{J}) \mathscr{Y}_{\mathcal{J}}(\mathscr{J}^*) \widetilde{\kappa}_k \leq a_k \bigr\} \\
&\qquad = \Bigl\{ s \in \mathrm{Im}(\mathscr{J}) \mid L_{\mathcal{J},\mathscr{J}^*}\bigl(\mathscr{Y}_{\mathcal{J}}(\mathscr{J}^*)^{\mathrm{T}} \mathscr{U}(\mathscr{J})^{\mathrm{T}} s \bigr) \leq c_{\mathcal{J},1}(\alpha, \mathscr{J}^*) \Bigr\} \\
&\qquad = \Bigl\{ s \in \mathrm{Im}(\mathscr{J}) \mid L_{\mathcal{J},1}\bigl(\mathscr{U}(\mathscr{J})^{\mathrm{T}} s, \mathscr{J}^* \bigr) \leq c_{\mathcal{J},1}(\alpha, \mathscr{J}^*) \Bigr\},
\end{aligned}$$

where we use Proposition B.3.4.c and Lemma 4.1.7.b. As $P_\xi\bigl(S \in \mathrm{Im}(\mathscr{J})\bigr) = 1$, cf. Witting [71, Hilfssatz 1.90], we have that

$$\phi_1'(S) = \begin{cases} 1, & S^{\mathrm{T}} \xi_k > a_k \text{ for some } k \in \mathbb{N}, \\ 0, & S^{\mathrm{T}} \xi_k \leq a_k \text{ for all } k \in \mathbb{N}. \end{cases}$$

Theorem 4.1.19 yields the assertion. □

4.1.21 Theorem. Suppose that Assumption 4.1.16 holds and that $\alpha \in (0, 1/2)$. The sequence of tests $\varphi_{n,1}$, $n \in \mathbb{N}$, is asymptotically admissible for the testing problem $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathcal{J}}$.

Proof. Corollary 4.1.17 gives

$$\lim_{n\to\infty} \mathbb{E}_{n,\xi}(\varphi_{n,1}) = \mathbb{E}_\xi(\phi_1'), \quad \xi \in \widetilde{\mathcal{H}}_1^{\mathcal{J}} \cup \widetilde{\mathcal{K}}_1^{\mathcal{J}}.$$

Let ϕ_n, $n \in \mathbb{N}$, be another sequence of tests and let n' be some infinite subsequence of the natural numbers. The Uniform Weak Compactness Lemma, cf. Witting and Müller-Funk [72, Satz 6.150], yields that there exists a subsubsequence n_k' and a test ϕ in the limit model, such that

$$\lim_{k\to\infty} \mathbb{E}_{n_k',\xi}(\phi_{n_k'}) = \mathbb{E}_\xi(\phi).$$

Assume that $\mathbb{E}_\xi(\phi) \leq \mathbb{E}_\xi(\phi_1')$, if $\xi \in \widetilde{\mathcal{H}}_1^{\mathcal{J}}$ and $\mathbb{E}_\xi(\phi) \geq \mathbb{E}_\xi(\phi_1')$, if $\xi \in \widetilde{\mathcal{K}}_1^{\mathcal{J}}$. Since ϕ_1' is admissible, cf. Proposition 4.1.20, it holds that $\mathbb{E}_\xi(\phi) = \mathbb{E}_\xi(\phi_1')$. The subsequence principle yields $\lim_{n\to\infty} \mathbb{E}_{n,\xi}(\phi_n) = \mathbb{E}_\xi(\phi_1')$. Hence, the sequences ϕ_n and $\varphi_{n,1}$ converge in distribution to ϕ_1' in the sense of Strasser [68, Definition 62.1]. ϕ_1' is uniquely determined by its distribution and non-randomized, see Proposition 4.1.20. Strasser [68, Theorem 63.6, Remark 63.2] gives the assertion. □

The last result means that there exists no sequence of tests that is uniformly better than the proposed sequence of tests. However, other sequences of admissible tests can be constructed with the help of Theorem 4.1.19.

4.2 Linear Testing Problems

Analog to the previous section we assume that Assumption 4.1.1 holds. In this section it is aimed to construct a testing procedure for linear hypotheses. More precisely, it is aimed to tackle the testing problem

$$\mathcal{H}_2^{\mathfrak{L}_0} : \beta \in \mathfrak{L}_0,\ \eta \in \mathbb{R}^q \quad \text{versus} \quad \mathcal{K}_2^{\mathfrak{L}_1} : \beta \in \mathfrak{L}_1 \backslash \mathfrak{L}_0,\ \eta \in \mathbb{R}^q,$$

where $\mathfrak{L}_0$ and $\mathfrak{L}_1$ are linear sub-spaces of $\mathbb{R}^r$, such that $\mathfrak{L}_0 \subset \mathfrak{L}_1$, and $\mathfrak{L}_0 \neq \mathfrak{L}_1$. As we allow the asymptotic information matrix $\mathscr{J}$ to be degenerated, the sub-spaces have to satisfy an additional regularity condition, which we want to discuss in the following paragraph. But first, let us introduce some more notation. Let $\mathscr{L}_i \in \mathbb{R}^{r\times l_i}$ be some matrix, such that $\mathrm{Im}(\mathscr{L}_i) = \mathfrak{L}_i$, $i = 0, 1$. Furthermore, we set

$$\mathscr{V}_i = \begin{pmatrix} \mathscr{L}_i & 0 \\ 0 & \mathscr{E}_q \end{pmatrix}, \qquad i = 0, 1,$$

where $\mathscr{E}_q$ denotes the $(q \times q)$-unity matrix. Looking at the limit experiment $\big(\Omega, \mathcal{A}, \{P_\xi \mid \xi \in \mathbb{R}^{r+q}\}\big)$ and remembering that $P_\xi = P_{\xi'}$ is equivalent to $\mathscr{J}\xi = \mathscr{J}\xi'$, see Lemma 4.1.4, one sees that hypothesis $\mathcal{H}_2^{\mathfrak{L}_0}$ and alternative $\mathcal{K}_2^{\mathfrak{L}_1}$ are disjoint, if and only if

$$\Theta\big(\mathcal{H}_2^{\mathfrak{L}_0}\big) \cap \Theta\big(\mathcal{K}_2^{\mathfrak{L}_1}\big) = \emptyset, \tag{4.17}$$

where

$$\Theta\big(\mathcal{H}_2^{\mathcal{L}_0}\big) = \big\{\mathscr{I}\xi \mid \xi \in \mathcal{H}_2^{\mathcal{L}_0}\big\} \quad \text{and} \quad \Theta\big(\mathcal{K}_2^{\mathcal{L}_1}\big) = \big\{\mathscr{I}\xi \mid \xi \in \mathcal{K}_2^{\mathcal{L}_1}\big\}$$

are the induced parameter sets of the hypothesis and the alternative. If $\mathscr{I}$ is not degenerated the condition (4.17) is trivially satisfied. Finally, if the hypothesis $\mathcal{H}_2^{\mathcal{L}_0}$ and the alternative $\mathcal{H}_2^{\mathcal{K}_1}$ are disjoint the testing problem transforms to

$$\widetilde{\mathcal{H}}_2^{\mathcal{L}_0} : \xi \in \Theta_{\mathcal{L}_0} \quad \text{versus} \quad \widetilde{\mathcal{K}}_2^{\mathcal{L}_1} : \xi \in \Theta_{\mathcal{L}_1} \backslash \Theta_{\mathcal{L}_0}$$

where

$$\Theta_{\mathcal{L}_0} = \big\{\xi \mid \mathscr{I}\xi \in \Theta(\mathcal{H}_2^{\mathcal{L}_0})\big\} \quad \text{and} \quad \Theta_{\mathcal{L}_1} = \big\{\xi \mid \mathscr{I}\xi \in \Theta(\mathcal{K}_2^{\mathcal{L}_1})\big\}.$$

Analog to Section 4.1 we study the testing problem $\widetilde{\mathcal{H}}_2^{\mathcal{L}_0}$ versus $\widetilde{\mathcal{H}}_2^{\mathcal{L}_1}$ under the limit model and derive some test statistic. This statistic will be the basis to propose some test statistic for finite $n \in \mathbb{N}$. To find a reasonable test we state a well-known result, cf. Witting and Müller-Funk [72, Satz 6.168], which we slightly modify by allowing the covariance matrix to be degenerated.

4.2.1 Theorem. Define the mappings $\Pi_{\mathcal{V}_i} : \mathrm{Im}(\mathscr{I}) \longrightarrow \mathrm{Im}(\mathscr{I}\mathscr{V}_i)$,

$$\Pi_{\mathcal{V}_i}(s) = \mathscr{I}\mathscr{V}_i\big(\mathscr{V}_i^{\mathrm{T}}\mathscr{I}\mathscr{V}_i\big)^{-}\mathscr{V}_i^{\mathrm{T}}s, \qquad i = 0, 1.$$

The mapping $\Pi_{\mathcal{V}_i}$ is obviously the orthogonal projection on the space $\mathcal{V}_i = \mathrm{Im}(\mathscr{I}\mathscr{V}_i)$ with respect to the inner product $<s_1, s_2>_{\mathscr{I}^-} = s_1^{\mathrm{T}}\mathscr{I}^{-}s_2$, see Proposition B.2.5. Moreover, set

$$T(s) = \big(\Pi_{\mathcal{V}_1}(s) - \Pi_{\mathcal{V}_0}(s)\big)^{\mathrm{T}}\mathscr{I}^{-}\big(\Pi_{\mathcal{V}_1}(s) - \Pi_{\mathcal{V}_0}(s)\big).$$

Under the limit experiment the following assertions hold true.

a) Under P_ξ, the statistic $T(S)$ is distributed according to a $\chi_l^2\big(\delta\big)$-distribution, where $l = \dim(\mathcal{V}_1) - \dim(\mathcal{V}_0)$ and $\delta = T(\mathscr{I}\xi)$. Additionally, if $\xi \in \widetilde{\mathcal{H}}_2^{\mathcal{L}_0} \cup \widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$, then $\delta = 0$ is equivalent to $\mathscr{I}\xi \in \mathrm{Im}\big(\mathscr{I}\mathscr{V}_0\big)$.

b) Let $\mathfrak{Q}_0$ denote the group of affine transformations $\pi : \mathrm{Im}(\mathscr{J}) \longrightarrow \mathrm{Im}(\mathscr{J})$, $\pi(x) = \mathscr{Q}x + u$, where $\mathscr{Q} : \mathrm{Im}(\mathscr{J}) \longrightarrow \mathrm{Im}(\mathscr{J})$ is a linear mapping, $u \in \mathrm{Im}\big(\mathscr{J}\mathscr{V}_0\big)$ and $\mathrm{Im}\big(\mathscr{Q}\mathscr{J}\mathscr{V}_0\big) = \mathrm{Im}\big(\mathscr{J}\mathscr{V}_0\big)$, $\mathrm{Im}\big(\mathscr{Q}\mathscr{J}\mathscr{V}_1\big) = \mathrm{Im}\big(\mathscr{J}\mathscr{V}_1\big)$, $\mathrm{Im}(\mathscr{Q}\mathscr{J}) = \mathrm{Im}(\mathscr{Q})$, $\mathscr{Q}\mathscr{J}\mathscr{Q}^{\mathrm{T}} = \mathscr{J}$ as well as $\mathscr{Q}^{\mathrm{T}}\mathscr{J}^{-}\mathscr{Q} = \mathscr{J}^{-}$.

The testing problem $\widetilde{\mathcal{H}}_2^{\mathcal{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$ is invariant with respect to $\mathfrak{Q}_0$. Moreover, T is a maximal invariant statistic with respect to $\mathfrak{Q}_0$ in the sense that $T(x) = T(\pi x)$ for all $\pi \in \mathfrak{Q}_0$, $x \in \mathrm{Im}(\mathscr{J})$ and that $T(x) = T(y)$ implies the existence of $\pi \in \mathfrak{Q}_0$, such that $\Pi_{\mathcal{V}_1}(y) = \Pi_{\mathcal{V}_1}(\pi x) = \pi\Pi_{\mathcal{V}_1 1}(x)$.

c) The test $\varphi = \mathbb{1}\big(T(S) > \chi^2_{l,\alpha}\big)$ is a uniformly most powerful invariant α-test for the testing problem $\widetilde{\mathcal{H}}_2^{\mathcal{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$, where $\chi^2_{l,\alpha}$ denotes the $(1-\alpha)$-quantile of a χ^2-distribution with l degrees of freedom and l is given in a).

Proof. See Appendix A.4. □

The next result helps us to simplify the statistic T.

4.2.2 Lemma. Let us define

$$\Pi_{\mathcal{L}_0,\mathcal{L}_1}(u, \mathscr{J}^*) = \Pi_{\mathcal{L}_1}(u, \mathscr{J}^*) - \Pi_{\mathcal{L}_0}(u, \mathscr{J}^*), \qquad u \in \mathbb{R}^r,$$

where

$$\Pi_{\mathcal{L}_i}(u, \mathscr{J}^*) = \mathscr{J}^*\mathscr{L}_i\big(\mathscr{L}_i^{\mathrm{T}}\mathscr{J}^*\mathscr{L}_i\big)^{-}\mathscr{L}_i^{\mathrm{T}}u, \qquad u \in \mathbb{R}^r, \quad i = 0, 1,$$

and

$$L_{\mathcal{L}_0,\mathcal{L}_1,2}(u, \mathscr{J}^*) = \Pi_{\mathcal{L}_0,\mathcal{L}_1}(u, \mathscr{J}^*)^{\mathrm{T}}\big(\mathscr{J}^*\big)^{-}\Pi_{\mathcal{L}_0,\mathcal{L}_1}(u, \mathscr{J}^*).$$

It holds that

$$T(s) = L_{\mathcal{L}_0,\mathcal{L}_1,2}\big(\mathscr{U}(\mathscr{J})^{\mathrm{T}}s, \mathscr{J}^*\big) \quad \text{for all } s \in \mathrm{Im}(\mathscr{J}),$$

where $\mathscr{U}(\mathscr{J})$ is introduced in Lemma 4.1.7.

Proof. In the proof we use the concept provided in Section B.2, especially Proposition B.2.5. As $\Pi_{\mathcal{V}_i}(\cdot)$, $i = 0,1$, are orthogonal projections on $\mathcal{V}_i$ see Proposition B.2.5.d, and as for all $w \in \ker(\mathscr{J}^*)$ it holds that

$$\begin{pmatrix} \mathscr{J}_{1,1} & \mathscr{J}_{1,2} \\ \mathscr{J}_{2,1} & \mathscr{J}_{2,2} \end{pmatrix} \begin{pmatrix} w \\ 0 \end{pmatrix} = \begin{pmatrix} \mathscr{J}_{1,1} w \\ \mathscr{J}_{2,1} w \end{pmatrix} = \begin{pmatrix} \mathscr{J}_{1,2} \mathscr{J}_{2,2}^{-} \mathscr{J}_{2,1} w \\ \mathscr{J}_{2,2} \mathscr{J}_{2,2}^{-} \mathscr{J}_{2,1} w \end{pmatrix}$$
$$= \begin{pmatrix} \mathscr{J}_{1,1} & \mathscr{J}_{1,2} \\ \mathscr{J}_{2,1} & \mathscr{J}_{2,2} \end{pmatrix} \begin{pmatrix} 0 \\ \mathscr{J}_{2,2}^{-} \mathscr{J}_{2,1} w \end{pmatrix},$$

where we use Proposition B.3.4.b, we can find matrices $\widetilde{\mathscr{L}}_i$, $i = 0,1$, whose columns are linearly independent and elements of $\mathrm{Im}(\mathscr{J}^*)$, such that $\mathcal{V}_i = \mathrm{Im}(\mathscr{J} \widetilde{\mathscr{V}}_i)$, where

$$\widetilde{\mathscr{V}}_i = \begin{pmatrix} \widetilde{\mathscr{L}}_i & 0 \\ 0 & \mathscr{E}_q \end{pmatrix}, \qquad i = 0,1.$$

Clearly, it holds that $\mathrm{Im}(\mathscr{J}^* \mathscr{L}_i) = \mathrm{Im}(\mathscr{J}^* \widetilde{\mathscr{L}}_i)$, $i = 0,1$.

Using the uniqueness of orthogonal projections, see Proposition B.2.3.a, and Proposition B.2.5.d we get that

$$\Pi_{\mathcal{V}_i}(s) = \mathscr{J} \widetilde{\mathscr{V}}_i \big(\widetilde{\mathscr{V}}_i^{\mathrm{T}} \mathscr{J} \widetilde{\mathscr{V}}_i \big)^{-} \widetilde{\mathscr{V}}_i^{\mathrm{T}} s, \quad s \in \mathrm{Im}(\mathscr{J}), \qquad i = 0,1, \tag{4.18}$$

and

$$\Pi_{\mathcal{L}_i}(u, \mathscr{J}^*) = \mathscr{J}^* \widetilde{\mathscr{L}}_i \big(\widetilde{\mathscr{L}}_i^{\mathrm{T}} \mathscr{J}^* \widetilde{\mathscr{L}}_i \big)^{-} \widetilde{\mathscr{L}}_i^{\mathrm{T}} u, \quad u \in \mathrm{Im}(\mathscr{J}^*), \quad i = 0,1. \tag{4.19}$$

As the columns of $\widetilde{\mathscr{L}}_i$ are linearly independent and elements of $\mathrm{Im}(\mathscr{J}^*)$, the matrix $\mathscr{A}_i = \widetilde{\mathscr{L}}_i^{\mathrm{T}} \mathscr{J}^* \widetilde{\mathscr{L}}_i$ is invertible. Moreover, one readily checks that

$$\big(\widetilde{\mathscr{V}}_i^{\mathrm{T}} \mathscr{J} \widetilde{\mathscr{V}}_i \big)^{-} = \begin{pmatrix} \mathscr{A}_i & \widetilde{\mathscr{L}}_i^{\mathrm{T}} \mathscr{J}_{1,2} \\ \mathscr{J}_{2,1} \widetilde{\mathscr{L}}_i & \mathscr{J}_{2,2} \end{pmatrix}^{-}$$
$$= \begin{pmatrix} \mathscr{A}_i^{-1} & -\mathscr{A}_i^{-1} \widetilde{\mathscr{L}}_i^{\mathrm{T}} \mathscr{J}_{1,2} \mathscr{J}_{2,2}^{-} \\ -\mathscr{J}_{2,2}^{-} \mathscr{J}_{2,1} \widetilde{\mathscr{L}}_i \mathscr{A}_i^{-1} & \mathscr{J}_{2,2}^{-} + \mathscr{J}_{2,2}^{-} \mathscr{J}_{2,1} \widetilde{\mathscr{L}}_i \mathscr{A}_i^{-1} \widetilde{\mathscr{L}}_i^{\mathrm{T}} \mathscr{J}_{1,2} \mathscr{J}_{2,2}^{-} \end{pmatrix}.$$

This result implies that $\widetilde{\mathscr{V}}_i\big(\widetilde{\mathscr{V}}_i^{\mathrm{T}}\mathscr{J}\widetilde{\mathscr{V}}_i\big)^{-}\widetilde{\mathscr{V}}_i^{\mathrm{T}}$ is equal to

$$\begin{pmatrix} \widetilde{\mathscr{L}}_i\mathscr{A}_i^{-1}\widetilde{\mathscr{L}}_i^{\mathrm{T}} & -\widetilde{\mathscr{L}}_i\mathscr{A}_i^{-1}\widetilde{\mathscr{L}}_i^{\mathrm{T}}\mathscr{J}_{1,2}\mathscr{J}_{2,2}^{-} \\ -\mathscr{J}_{2,2}^{-}\mathscr{J}_{2,1}\widetilde{\mathscr{L}}_i\mathscr{A}_i^{-1}\widetilde{\mathscr{L}}_i^{\mathrm{T}} & \mathscr{J}_{2,2}^{-}+\mathscr{J}_{2,2}^{-}\mathscr{J}_{2,1}\widetilde{\mathscr{L}}_i\mathscr{A}_i^{-1}\widetilde{\mathscr{L}}_i^{\mathrm{T}}\mathscr{J}_{1,2}\mathscr{J}_{2,2}^{-} \end{pmatrix}.$$

Consequently, it holds that

$$\begin{aligned} \big|\big|\Pi_{\mathcal{V}_i}(s)\big|\big|^2_{\mathscr{J}^-} &= s^{\mathrm{T}}\widetilde{\mathscr{V}}_i\big(\widetilde{\mathscr{V}}_i^{\mathrm{T}}\mathscr{J}\widetilde{\mathscr{V}}\big)^{-}\widetilde{\mathscr{V}}_i^{\mathrm{T}}s \\ &= s^{\mathrm{T}}\begin{pmatrix} \widetilde{\mathscr{L}}_i(\widetilde{\mathscr{L}}_i^{\mathrm{T}}\mathscr{J}^*\widetilde{\mathscr{L}}_i)^{-1}\widetilde{\mathscr{L}}_i^{\mathrm{T}}\mathscr{U}^{\mathrm{T}}s \\ -\mathscr{J}_{2,2}^{-}\mathscr{J}_{2,1}\widetilde{\mathscr{L}}_i(\widetilde{\mathscr{L}}_i^{\mathrm{T}}\mathscr{J}^*\widetilde{\mathscr{L}}_i)^{-1}\widetilde{\mathscr{L}}_i^{\mathrm{T}}\mathscr{U}^{\mathrm{T}}s+\mathscr{J}_{2,2}^{-}s_2 \end{pmatrix} \\ &= s^{\mathrm{T}}\mathscr{U}\widetilde{\mathscr{L}}_i(\widetilde{\mathscr{L}}_i^{\mathrm{T}}\mathscr{J}^*\widetilde{\mathscr{L}}_i)_i^{-1}\widetilde{\mathscr{L}}_i^{\mathrm{T}}\mathscr{U}^{\mathrm{T}}s+s_2\mathscr{J}_{2,2}^{-}s_2 \\ &= \big|\big|\Pi_{\mathcal{L}_i}(\mathscr{U}^{\mathrm{T}}s,\mathscr{J}^*)\big|\big|^2_{(\mathscr{J}^*)^-}+s_2\mathscr{J}_{2,2}^{-}s_2, \end{aligned}$$

where $s_2 = \pi^{r+q}_{\{r+1,\dots,r+q\}}(s)$, $\mathscr{U} = \mathscr{U}(\mathscr{J})$ and we use (4.18) and (4.19) as well as $\mathscr{U}^{\mathrm{T}}\mathscr{J}\mathscr{U} = \mathscr{J}^*$ and Proposition B.3.4.c. Using the last equation and Proposition B.2.4.f gives

$$\begin{aligned} T(s) &= \big|\big|\Pi_{\mathcal{V}_1}(s)-\Pi_{\mathcal{V}_0}(s)\big|\big|^2_{\mathscr{J}^-} \\ &= \big|\big|\Pi_{\mathcal{V}_1}(s)\big|\big|^2_{\mathscr{J}^-}-\big|\big|\Pi_{\mathcal{V}_0}(s)\big|\big|^2_{\mathscr{J}^-} \\ &= \big|\big|\Pi_{\mathcal{L}_1}(\mathscr{U}^{\mathrm{T}}s,\mathscr{J}^*)\big|\big|^2_{(\mathscr{J}^*)^-}-\big|\big|\Pi_{\mathcal{L}_0}(\mathscr{U}^{\mathrm{T}}s,\mathscr{J}^*)\big|\big|^2_{(\mathscr{J}^*)^-} \\ &= \big|\big|\Pi_{\mathcal{L}_1}(\mathscr{U}^{\mathrm{T}}s,\mathscr{J}^*)-\Pi_{\mathcal{L}_0}(\mathscr{U}^{\mathrm{T}}s,\mathscr{J}^*)\big|\big|^2_{(\mathscr{J}^*)^-}. \end{aligned}$$

This is the assertion. □

4.2.3 Corollary. It holds that

$$T(S) = L_{\mathcal{L}_0,\mathcal{L}_1,2}\big(U,\mathscr{J}^*\big) \quad P_\xi\text{-almost surely,}$$

where $U = \mathscr{U}(\mathscr{J})^{\mathrm{T}}S$.

Proof. As $P_\xi \ll P_0$ it suffices to show the assertion for P_0. Witting [71, Hilfssatz 1.90] shows that $P_0\big(S \in \mathrm{Im}(\mathscr{J})\big) = 1$. Consequently, Lemma 4.2.2 implies the assertion. □

As the statistic U corresponds with the statistic $U_n = \mathscr{U}(\mathscr{J})^{\mathrm{T}} S_n$ for finite n, the previous result suggests that $L_{\mathfrak{L}_0,\mathfrak{L}_1,2}\big(U_n, \mathscr{J}^*\big)$ is a reasonable test statistic for the testing problem $\widetilde{\mathcal{H}}_2^{\mathfrak{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathfrak{L}_1}$. However, this statistic still depends on asymptotic quantities. In order to get some applicable test we have to replace these quantities by suitable estimators. This is done completely analogously to Section 4.1.

4.2.4 Theorem. Under Assumption 4.1.12 it holds that

a) $L_{\mathfrak{L}_0,\mathfrak{L}_1,2}(X_n, \mathscr{A}) - L_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\widehat{X}_n, \widehat{\mathscr{A}_n}) \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$.

b) $L_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\widehat{X}_n, \widehat{\mathscr{A}_n}) \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} L_{\mathfrak{L}_0,\mathfrak{L}_1,2}\big(X, \mathscr{A}\big)$, where $L_{\mathfrak{L}_0,\mathfrak{L}_1,2}(X, \mathscr{A}) \sim \chi^2_l(\delta)$,

$$l = \operatorname{rank}(\mathscr{A}\mathscr{L}_1) - \operatorname{rank}(\mathscr{A}\mathscr{L}_0) \quad \text{and} \quad \delta = L_{\mathfrak{L}_0,\mathfrak{L}_1,2}\big(\kappa, \mathscr{A}\big).$$

Moreover, if $\kappa \in \operatorname{Im}(\mathscr{A}\mathscr{L}_1)$, then $\delta = 0$ is equivalent to $\kappa \in \operatorname{Im}(\mathscr{A}\mathscr{L}_0)$.

Proof. The first assertion can be seen as follows. Lemma 4.1.10.b and Proposition B.4.6 give that

$$\Pi_{\mathfrak{L}_i}(X_n, \mathscr{A}) - \Pi_{\mathfrak{L}_i}(X_n, \widehat{\mathscr{A}_n}) \longrightarrow_{\mathbb{P}_n} 0, \quad \text{as } n \to \infty, \qquad i = 0, 1.$$

Consequently, it holds that

$$\Pi_{\mathfrak{L}_0,\mathfrak{L}_1}(X_n, \mathscr{A}) - \Pi_{\mathfrak{L}_0,\mathfrak{L}_1}(\widehat{X}_n, \widehat{\mathscr{A}_n}) \longrightarrow_{\mathbb{P}_n} 0, \quad \text{as } n \to \infty, \tag{4.20}$$

and therefore

$$\mathscr{A}^- \, \Pi_{\mathfrak{L}_0,\mathfrak{L}_1}(X_n, \mathscr{A}) - \widehat{\mathscr{A}_n^-} \, \Pi_{\mathfrak{L}_0,\mathfrak{L}_1}(\widehat{X}_n, \widehat{\mathscr{A}_n}) \longrightarrow_{\mathbb{P}_n} 0, \quad \text{as } n \to \infty, \tag{4.21}$$

where one uses Lemma 4.1.10.a and Proposition B.4.6. Because of (4.20) and (4.21), Proposition B.4.6 yields the assertion.

Proof of b). Because of a), Slutsky's Lemma and the Continuous Mapping Theorem, cf. Witting and Müller-Funk [72, Korollar 5.84, Satz 5.43], it follows the first part of the assertion. Completely analogously to the proof of Theorem 4.2.1.a, see Appendix A.4, one establishes the second part of the assertion. □

4.2.5 Corollary. Suppose that Assumption 4.1.12 is satisfied. Moreover, let $c_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\alpha,\mathscr{B})$ denote the $(1-\alpha)$ quantile of a χ^2-distribution with l degrees of freedom, where

$$l = \operatorname{rank}(\mathscr{B}\mathscr{L}_1) - \operatorname{rank}(\mathscr{B}\mathscr{L}_0).$$

Then it holds that

$$c_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\alpha,\widehat{\mathscr{A}_n}) - c_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\alpha,\mathscr{A}) \longrightarrow_{\mathbb{P}_n} 0, \quad \text{as } n\to\infty,$$

for all $\alpha \in (0,1)$.

Proof. Without loss of generality we can assume that all random variables are defined on the same probability space. Using the sub-sub-sequence principle for random variables that converge in probability, cf. Proposition B.4.8,we can also assume that $\widehat{\mathscr{A}_n} \to \mathscr{A}$ almost surely.

Moreover, one sees that $\ker(\mathscr{A}) \subset \ker(\widehat{\mathscr{A}_n})$ $\mathbb{P}_n$-almost surely implies that $\ker(\mathscr{A}\mathscr{L}_i) \subset \ker(\widehat{\mathscr{A}_n}\mathscr{L}_i)$ $\mathbb{P}_n$-almost surely. Thus, it holds that $\operatorname{rank}(\mathscr{A}\mathscr{L}_i) \geq \operatorname{rank}(\widehat{\mathscr{A}_n}\mathscr{L}_i)$ $\mathbb{P}_n$-almost surely for all sufficiently large $n\in\mathbb{N}$.

As the set of all matrices with rank greater or equal to $\operatorname{rank}(\mathscr{A}\mathscr{L}_i)$ is open, one gets that $\operatorname{rank}(\mathscr{A}\mathscr{L}_i) = \operatorname{rank}(\widehat{\mathscr{A}_n}\mathscr{L}_i)$, $i=0,1$, for all sufficiently large $n\in\mathbb{N}$ and $c_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\alpha,\widehat{\mathscr{A}_n}) = c_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\alpha,\mathscr{A})$ for all sufficiently large $n\in\mathbb{N}$. The sub-sub-sequence principle for random variables that converge in probability gives the assertion. $\square$

After this preparation we propose the sequence of tests $\varphi_{n,2}$, $n\in\mathbb{N}$, where

$$\varphi_{n,2} = \begin{cases} 1, \\ 0, \end{cases} L_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\widehat{U}_n,\widehat{V}_n) - c_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\alpha,\widehat{V}_n) \begin{array}{c} > \\ \leq \end{array} 0\,, \quad n\in\mathbb{N},$$

for the testing problem $\widetilde{\mathcal{H}}_2^{\mathfrak{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathfrak{L}_1}$. The next result summarizes the asymptotic properties of this sequence of tests.

4.2.6 Corollary. Under Assumption 4.1.16, it holds that

$$\mathbb{E}_{n,\xi}\big(\varphi_{n,2}\big) \longrightarrow P_\xi\big(L_{\mathfrak{L}_0,\mathfrak{L}_1,2}(U,\mathscr{J}^*) > c_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\alpha,\mathscr{J}^*)\big), \quad \text{as } n\to\infty,$$

where $U = \mathscr{U}(\mathscr{J})^{\mathrm{T}} S$ and $\mathfrak{L}(U \mid P_\xi) = \mathcal{N}(\mathscr{J}^*\beta, \mathscr{J}^*)$. Especially, the sequence of tests $\varphi_{n,2}$, $n \in \mathbb{N}$, keeps asymptotically the level on the hypothesis $\widetilde{\mathcal{H}}_2^{\mathfrak{L}_0}$ and is unbiased. Additionally, we have that

$$P_\xi\big(L_{\mathfrak{L}_0,\mathfrak{L}_1,2}(U, \mathscr{J}^*) > c_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\alpha, \mathscr{J}^*)\big) = \mathbb{E}_\xi(\varphi),$$

where φ, is the most powerful, invariant α-test for the testing problem $\widetilde{\mathcal{H}}_2^{\mathfrak{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathfrak{L}_1}$ given in Theorem 4.2.1.c.

Proof. The Continuous Mapping Theorem, cf. Witting and Müller-Funk [72, Satz 5.43] gives that $U_n \xrightarrow{\mathfrak{D}}_{P_{n,\xi}} U$, where one checks that $U \sim \mathcal{N}(\mathscr{J}^*\beta, \mathscr{J}^*)$, see proof of Corollary 4.1.17. Setting $\mathscr{A} = \mathscr{J}^*$, $\widehat{\mathscr{A}_n} = \widehat{V}_n$, $X_n = U_n$, $\widehat{X}_n = \widehat{U}_n$ and $\mathbb{P}_n = P_{n,0}$, $n \in \mathbb{N}$, one sees that Assumption 4.1.12 holds. Theorem 4.2.4.a, Corollary 4.2.5 and Remark 4.1.2 give that

$$L_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\widehat{U}_n, \widehat{V}_n) - c_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\alpha, \widehat{V}_n) \\ - \big(L_{\mathfrak{L}_0,\mathfrak{L}_1,2}(U_n, \mathscr{J}^*) - c_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\alpha, \mathscr{J}^*)\big) \longrightarrow_{P_{n,\xi}} 0.$$

Setting $\mathscr{A} = \widehat{\mathscr{A}_n} = \mathscr{J}^*$, $X_n = \widehat{X}_n = U_n$, and $\mathbb{P}_n = P_{n,\xi}$, $n \in \mathbb{N}$, one sees that Assumption 4.1.12 holds. Theorem 4.2.4.b and Slutsky's Lemma, cf. Witting and Müller-Funk [72, Korollar 5.84], yield that

$$L_{\mathfrak{L}_0,\mathfrak{L}_1,2}\big(\widehat{U}_n, \widehat{V}_n\big) - c_{\mathfrak{L}_0,\mathfrak{L}_1,2}\big(\alpha, \widehat{V}_n\big) \xrightarrow{\mathfrak{D}}_{P_{n,\xi}} L_{\mathfrak{L}_0,\mathfrak{L}_1,2}\big(U, \mathscr{J}^*\big) - c_{\mathfrak{L}_0,\mathfrak{L}_1,2}\big(\alpha, \mathscr{J}^*\big),$$

Obviously, we have, cf. Theorem 4.2.4.b,

$$P_\xi\big(L_{\mathfrak{L}_0,\mathfrak{L}_1,2}(U, \mathscr{J}^*) - c_{\mathfrak{L}_0,\mathfrak{L}_1,2}(\alpha, \mathscr{J}^*) = 0\big) = 0.$$

The Portmanteau Theorem, cf. Witting and Müller-Funk [72, Satz 5.40], gives the first assertion. The equivalence stated in Theorem 4.2.4.b gives that the test keeps asymptotically the level on the hypothesis and is asymptotically unbiased as well. The last assertion is an immediate consequence of Corollary 4.2.3. □

In the last part of this section we show that the sequence of tests $\varphi_{n,2}$, $n \in \mathbb{N}$, is asymptotically admissible for the testing problem $\widetilde{\mathcal{H}}_2^{\mathfrak{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathfrak{L}_1}$, see Definition 4.1.18. The proceeding is exactly the same as in Section 4.1.

4.2.7 Theorem. Let $\xi_k \in \mathbb{R}^{r+q}$ and $a_k \in \mathbb{R}$, $k \in \mathbb{N}$, be arbitrary. If $\xi_k \in \widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$, $k \in \mathbb{N}$, then the test $\phi'(S)$, where

$$\phi'(s) = \begin{cases} 1, & s^{\mathrm{T}}\xi_k > a_k \text{ for some } k \in \mathbb{N}, \\ 0, & s^{\mathrm{T}}\xi_k \le a_k \text{ for all } k \in \mathbb{N}, \end{cases}$$

is admissible for the testing problem $\widetilde{\mathcal{H}}_2^{\mathcal{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$ and uniquely determined by its distribution.

Proof. The proof of Theorem 71.14 in Strasser [68] is also applicable for this Theorem. The crucial point is the fact that the ξ_k, $k \in \mathbb{N}$, belong to the alternative $\widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$. □

4.2.8 Proposition. Under Assumption 4.1.16, the test $\phi_2'(S)$, where

$$\phi_2'(s) = \begin{cases} 1, & L_{\mathcal{L}_0,\mathcal{L}_1,2}\big(\mathscr{U}(\mathscr{J})^{\mathrm{T}}s, \mathscr{J}^*\big) - c_{\mathcal{L}_0,\mathcal{L}_1,2}(\alpha, \mathscr{J}^*) \begin{smallmatrix} > \\ \le \end{smallmatrix} 0, \\ 0, & \end{cases}$$

and $\alpha \in (0,1)$, is admissible for $\widetilde{\mathcal{H}}_2^{\mathcal{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$ and uniquely determined by its distribution.

Proof. We show that ϕ_2' has a representation as the test considered in Theorem 4.2.7. In the following we use the concept provided in Section B.2, especially Proposition B.2.5.

Remember that

$$L_{\mathcal{L}_0,\mathcal{L}_1,2}\big(\mathscr{U}(\mathscr{J})^{\mathrm{T}}s, \mathscr{J}^*\big) = T(s), \qquad s \in \mathrm{Im}(\mathscr{J}),$$

see Lemma 4.2.2. First, we show that $T : \mathrm{Im}(\mathscr{J}) \longrightarrow \mathbb{R}_+$ is a convex and therefore continuous function, cf. Borwein and Lewis [10, Theorem 4.1.3]. Clearly, it holds that

$$T(s) = \big|\big|\Pi_{\mathcal{V}_1}(s) - \Pi_{\mathcal{V}_0}(s)\big|\big|^2_{\mathscr{J}^-}, \qquad s \in \mathrm{Im}(\mathscr{J}),$$

where $\mathcal{V}_i = \mathrm{Im}(\mathscr{J}\mathscr{V}_i)$. The triangle inequality and the fact that $t \mapsto t^2$, $t \ge 0$, is convex and non-decreasing gives the assertion.

Thus, the set $C^* = \{T \leq c_{\mathcal{L}_0,\mathcal{L}_1,2}(\alpha, \mathscr{I}^*)\}$ is convex and closed. By basic separation theorems we know that there exists $\xi_k \in \operatorname{Im}(\mathscr{I})\backslash\{0\}$ and $a_k \in \mathbb{R}$, $k \in \mathbb{N}$, such that

$$C^* = \bigcap_{k=1}^{\infty} \{s \in \operatorname{Im}(\mathscr{I}) \mid \xi_k^{\mathrm{T}} u \leq a_k\}.$$

It holds that

$$\{s \in \operatorname{Im}(\mathscr{I}) \mid \xi_k^{\mathrm{T}} u \leq a_k\} = \{s \in \operatorname{Im}(\mathscr{I}) \mid <s, \mathscr{I}\xi_k>_{\mathscr{I}^-} \leq a_k\}.$$

First, we show that $\mathscr{I}\xi_k \in \mathcal{V}_1$. Assume that $\widetilde{s} = \lambda \cdot \big(\mathscr{I}\xi_k - \Pi_{\mathcal{V}_1}(\mathscr{I}\xi)\big) \neq 0$. as $<\widetilde{s}, \mathscr{I}\xi_k>_{\mathscr{I}^-} = \lambda ||\widetilde{s}||^2_{\mathscr{I}^-} > a_k$ for sufficiently large λ, where we use Proposition B.2.3.b, it results that $\widetilde{s} \notin C^*$. On the other hand, it holds that

$$0 \leq T(\widetilde{s}) = \big|\big|\Pi_{\mathcal{V}_1}(\widetilde{s}) - \Pi_{\mathcal{V}_0}(\widetilde{s})\big|\big|^2 = \big|\big|\Pi_{\mathcal{V}_0}(\widetilde{s})\big|\big|^2 \leq \big|\big|\Pi_{\mathcal{V}_1}(\widetilde{s})\big|\big|^2 = 0,$$

where we use Proposition B.2.4.b and Proposition B.2.4.h. This means $\widetilde{s} \in C^*$, this is a contradiction.

Now, we show that $\mathscr{I}\xi_k \notin \mathcal{V}_0$. Assume that $\mathscr{I}\xi_k \in \mathcal{V}_0$. For sufficiently large λ it holds that $\lambda \cdot <\mathscr{I}\xi_k, \mathscr{I}\xi_k>_{\mathscr{I}^-} > a_k$, that is to say $\lambda\mathscr{I}\xi \notin C^*$, having said this $T(\lambda\mathscr{I}\xi) = 0$, *i.e.* $\lambda\mathscr{I}\xi \in C^*$, where we use Proposition B.2.4.a and Proposition B.2.4.c. $\mathscr{I}\xi_k \in \mathcal{V}_1\backslash\mathcal{V}_0$ is equivalent to $\xi_k \in \widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$. And as $P_\xi\big(S \in \operatorname{Im}(\mathscr{I})\big) = 1$, cf. Witting [71, Hilfssatz 1.90], we have that

$$\phi_2'(S) = \begin{cases} 1, & S^{\mathrm{T}}\xi_k > a_k \text{ for some } k \in \mathbb{N}, \\ 0, & S^{\mathrm{T}}\xi_k \leq a_k \text{ for all } k \in \mathbb{N}. \end{cases}$$

Theorem 4.2.7 yields the assertion. □

4.2.9 Theorem. Assume that Assumption 4.1.16 holds and that $\alpha \in (0,1)$. The sequence of tests $\varphi_{n,2}$, $n \in \mathbb{N}$, is asymptotically admissible for the testing problem $\widetilde{\mathcal{H}}_2^{\mathcal{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$.

Proof. The proof is identical to the proof of Theorem 4.1.21. Instead of Corollary 4.1.17 and Proposition 4.1.20, one uses Corollary 4.2.6 and Proposition 4.2.8. □

The last result means that there exists no sequence of tests that is uniformly better than the proposed sequence of tests. However, other sequences of admissible tests can be constructed with the help of Theorem 4.2.7.

4.3 Test for Sequences of Hardest Parametric Sub-Models

In this section the results of Section 4.1 and Section 4.2 are applied to sequences of hardest parametric sub-models As a first step we verify that Assumption 4.1.1 is satisfied.

4.3.1 Proposition. Let $\big(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{P}_n\big)$, $n \in \mathbb{N}$, denote a sequence of localized, q-dimensional parametric sub-models of the modified Cox regression model, see Definition 1.3.4. Assume that $\big(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{P}_n\big)$, $n \in \mathbb{N}$, restricted to time τ is asymptotically normal with asymptotic information matrix $\mathscr{I}(\tau)$ and central sequence $S_n(\tau)$, $n \in \mathbb{N}$, then the sequence of statistical experiments $\big(\Omega_n, \mathfrak{F}_{n,\tau}, \{P_{n,\xi}^{(\tau)} \mid \xi \in \mathbb{R}^{r+q}\}\big)$, $n \in \mathbb{N}$, satisfies Assumption 4.1.1, where $S_n = S_n(\tau)$, $n \in \mathbb{N}$, and $\mathscr{I} = \mathscr{I}(\tau)$.

Proof. Paying attention to Definition 2.2.2 and Remark 2.2.3 gives the result. $\square$

Conditions implying the assumptions of Proposition 4.3.1 are stated in Theorem 2.3.10.a. In the following let us assume that $\big(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{P}_n\big)$, $n \in \mathbb{N}$, is a sequence of hardest parametric sub-models (SHPSM) restricted to time τ, see Definition 3.1.5. In the previous sections we saw that reasonable test statistics for multivariate one-sided testing problems and linear testing problems were dependents on the statistic $U_n(\tau) = \mathscr{U}(\mathscr{I})^{\mathrm{T}} S_n(\tau)$, see pp. 99 as well as pp. 121 and Lemma 4.1.7. Using equation (3.1) gives that

$$U_n(\tau) \longrightarrow_{P_{n,\xi}^{(\tau)}} \mathcal{N}\big(\mathscr{I}^*(\tau)\beta, \mathscr{I}^*(\tau)\big), \qquad \text{as } n \to \infty.$$

In the following discussion we show why it is useful to study models with degenerated asymptotic information matrix $\mathscr{I}^*(\tau)$, if one considers multivariate one-sided testing problems.

4.3.2 Discussion. As we consider a sequence of hardest parametric submodels it holds that $\mathscr{I}^*(\tau) = \mathscr{I}^{*,\,\mathrm{can}}(\tau)$, see Theorem 3.2.4. One sees that $\beta \in \ker\big(\mathscr{I}^{*,\,\mathrm{can}}(\tau)\big)$ is equivalent to

$$\beta^{\mathrm{T}} \mathscr{I}^*(\tau)\beta = \int_{I(\tau)} \sum_{u=1}^{p} \sum_{v=1}^{p} c^{(u)}\, c^{(u)} \left(\mu_2^{(u,v)} - \frac{\mu_1^{(u)}\,\mu_1^{(v)}}{\mu_0}\right)\alpha_0 \,\mathrm{d}s = 0, \qquad (4.22)$$

where we set

$$c^{(u)}(s) = \sum_{k=1}^{r_u} \widetilde{\beta}_u^{(k)} \gamma^{(u,k)}(s), \qquad \widetilde{\beta}_u = \big(\beta^{(\sum_{l=1}^{u-1} r_l + k)} \mid k = 1, \dots, r_u\big)^{\mathrm{T}}, \quad (4.23)$$

r_u, $u = 1, \dots, p$, are given in Definition 1.3.2 and the same calculations as in Remark 3.2.14.a are used. Equation (4.22) can imply several things. To simplify matter let us assume that the covariates do not have a linear dependence structure, *i.e.* no component of the covariate vector can be expressed as a linear function of the remaining components. This situation can be achieved by a reasonable experiment design. Therefore, we can suppose that

$$\left(\mu_2^{(u,v)}(s) - \frac{\mu_1^{(u)}(s)\,\mu_1^{(v)}(s)}{\mu_0(s)}\right) \quad u, v = 1, \dots, p,$$

has full rank for $\widetilde{\Lambda}_0^{(\tau)}$-almost all s. This means that $c^{(u)}(s) = 0$ for $\widetilde{\Lambda}_0$-almost all s, and $u = 1, \dots, r$, *i.e.* the weight functions belonging to the u-th component are linearly dependent, if $\widetilde{\beta}_u$ is not the null-vector. Thus, the weight functions for at least one component of the covariate vector are linearly dependent. The reasons for this dependency might be due to the weight functions, that $\tau < \tau_0$ is chosen too small, the right censoring, *i.e.* the linearly independent part of the weight functions are censored. Another reason for the linear dependency can be that the baseline hazard is zero on the sets where the weight functions are linearly independent.

However, in multivariate one-sided testing problems the linear dependency of the weight functions might arise very naturally. Let us consider the case of a univariate, non-negative covariate. For example, we want to test, if larger values of the covariate correlate with shorter survival times. Under the MCRM, see Definition 1.3.2, we can model this situation by setting

$$A^{(i)}_{\beta,\alpha}(\cdot) = \int_{[0,\cdot]} Y^{(i)}(s) \exp\biggl(Z_i \sum_{u=1}^{r} \beta^{(u)} \gamma^{(u)} \biggr) \alpha(s) \, \mathrm{d}s,$$

where $\gamma^{(u)}$, $u = 1, \ldots, r$, are some non-negative functions that determine the direction of the alternatives. Larger values of β given the covariate imply shorter survival times. Therefore, the above mentioned test problem turns out to be $\beta = 0$ versus $\beta \geq 0$, $\beta \neq 0$, *i.e.* a multivariate one-sided testing problem.

The cone $\bigl\{ \sum_{u=1}^{r} \beta^{(u)} \gamma^{(u)} \mid \beta \in \mathbb{R}^r \bigr\}$ gives the possible directions of the alternatives. For illustration, we choose $r = 3$ and

$$\gamma^{(1)}(s) = 1, \quad \gamma^{(2)}(s) = F_0(s), \quad \gamma^{(3)}(s) = F_0^2(s),$$

where $F_0(s)$ is a continuous, strictly increasing cumulative distribution function on $\mathbb{R}_+$. $\gamma^{(1)}$ corresponds with the case of proportional hazard rates, whereas $\gamma^{(2)}$ and $\gamma^{(3)}$ correspond with increasing differences in the hazard rates for large s, Obviously, the weight function

$$\gamma^{(4)}(s) = \bigl(\gamma^{(1)}(s) - \gamma^{(2)}(s) \bigr)^2 = \gamma^{(1)}(s) - 2\gamma^{(2)}(s) + \gamma^{(3)}(s)$$

does not belong to the cone $\bigl\{ \sum_{u=1}^{3} \beta^{(u)} \gamma^{(u)} \mid \beta \geq 0 \bigr\}$. However, the weight function $\gamma^{(4)}$ is non-negative like the others on the whole interval, *i.e.* it generates the same stochastic ordering as the other functions and might be therefore considered as a possible direction of the alternatives. A way out is to extend the model by adding the weight function $\gamma^{(4)}$. As a result, one gets not only a wider range of alternatives, but that the asymptotic information matrix of the MCRM is degenerated, since the functions $\gamma^{(u)}$, $u = 1, \ldots, 4$ are linearly dependent.

In the next step we show that one cannot find three weight functions spanning the same cone as $\gamma^{(u)}$, $u = 1, \dots, n$. Obviously, we have to show that there does not exist $x_u \in \mathbb{R}^3$, $u = 1, 2, 3$, such that

$$\Big\{\sum\nolimits_{u=1}^{4} \beta^{(u)} w_u \mid \beta \geq 0\Big\} = \Big\{\sum\nolimits_{u=1}^{3} \widetilde{\beta}^{(u)} x_u \mid \widetilde{\beta} \geq 0\Big\}, \tag{4.24}$$

where

$$w_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad w_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad w_4 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Assume we could find x_u, $u = 1, 2, 3$, such that (4.24) holds, then there exits a (3×4) matrix $\mathscr{B}$ and a (4×3) matrix $\mathscr{C}$ with non-negative entries, such that

$$\mathscr{W} = \mathscr{X}\mathscr{B} = \mathscr{W}\mathscr{C}\mathscr{B}$$

where we set

$$\mathscr{X} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \quad \text{and} \quad \mathscr{W} = \begin{pmatrix} w_1 & w_2 & w_3 & w_4 \end{pmatrix}.$$

On the one hand, it holds that $\operatorname{rank}(\mathscr{C}\mathscr{B}) \leq 3$. On the other hand we see that the entries of the matrix $\mathscr{C}\mathscr{B}$ are non-negative and that the only non-negative solution of the system of linear equations $\mathscr{W}y = w_u$ is given by $y = \big(\delta_{1,u}, \dots, \delta_{4,u}\big)^{\mathrm{T}}$, $u = 1, \dots, 4$, where $\delta_{i,j}$ denotes the Kronecker symbol. Clearly, this means that $\operatorname{rank}(\mathscr{C}\mathscr{B}) = 4$, a contradiction.

Consequently, there does not exist three weight functions that generate the cone given by $\gamma^{(u)}$, $u = 1, \dots, 4$. All in all, it is worth considering the case of degenerated limit distributions that are due to linear dependencies of the weight functions, if one treats multivariate one-sided testing problems.

The asymptotic information matrix of the MCRM for a model with a one-dimensional covariate and weight functions $\gamma^{(u)}$ is given by $\mathscr{I}^{*,\mathrm{can}}(\tau)$, where

$$\mathscr{I}^{*,\mathrm{can}(u,v)}(\tau) = \int_{I(\tau)} \gamma^{(u)}(s)\,\gamma^{(v)}(s)\left(\mu_2^{(1,1)}(s) - \frac{\mu_1^{(1)}(s)\mu_1^{(1)}(s)}{\mu_0(s)}\right)\alpha_0(s)\,\mathrm{d}s.$$

Under this model it holds that $\ker(\mathscr{I}^*) = \{\kappa \cdot (1, -1, 1, -1)^{\mathrm{T}} \mid \kappa \in \mathbb{R}\}$. Consequently, the conditions $\beta \not\geq 0$ and $-\beta \not\geq 0$ hold for all $\beta \in \ker(\mathscr{I}^*)\setminus\{0\}$, see equation (4.3). This means that the hypothesis $\mathcal{H}_1^{\mathcal{J}}$ and the alternative $\mathcal{K}_1^{\mathcal{J}}$ are disjoint under the limit model, cf. Proposition 4.1.5.

As $\gamma^{(u)}(s)\,\gamma^{(v)}(s) \geq 0$, $s \in \mathbb{R}_+$, $u, v = 1, \ldots, 4$, we get that $\mathscr{I}^{*,\mathrm{can}}\beta \geq 0$ for all $\beta \in \mathbb{R}^4$, $\beta \geq 0$. Consequently, the condition (4.12) holds, *i.e.* the sequence of tests $\varphi_{n,1}$, $n \in \mathbb{N}$, is asymptotically unbiased.

In the next step it is shown that Assumption 4.1.16 is satisfied by SHPSM under certain regularity conditions.

4.3.3 Assumption. Let $(\Omega_n, \mathcal{F}_n, \mathbb{F}_n, \mathfrak{P}_n)$, $n \in \mathbb{N}$, be a sequence of hardest parametric sub-models restricted to time τ with asymptotic information matrix $\mathscr{I}(\tau)$ and central sequence $S_n(\tau)$, $n \in \mathbb{N}$, see Definition 3.1.5. Let us assume that

$$\gamma^{(\dot{u},\ddot{u})} = \gamma_0^{(\dot{u},\ddot{u})} \circ H_0, \quad u = 1, \ldots, r, \qquad \text{where} \quad \gamma_0^{(\dot{u},\ddot{u})} : [0,1] \longrightarrow \mathbb{R},$$

is some measurable function and H_0 is some cumulative distribution function. Moreover, we suppose that the following conditions hold.

i) Assumption 2.2.1 and Assumption 2.3.9.i – viii hold.

ii) Assumption 3.2.1 holds.

iii) Suppose that $\widehat{\gamma}_n^{(\dot{u},\ddot{u})} = \gamma_0^{(\dot{u},\ddot{u})} \circ \widehat{H}_n$, $u = 1, \ldots, r$, where $\widehat{H}_n$ is some estimator H_0, such that $0 \leq \widehat{H}_n \leq 1$ $P_{n,0}$-almost surely. Moreover, let Assumption 3.2.11 hold.

iv) Furthermore, let us assume that $\beta \in \ker(\mathscr{I}^{*,\,\mathrm{can}}(\tau))$ implies

$$\sum_{v=1}^{r_u} \widetilde{\beta}_u^{(v)} \gamma_0^{(u,v)}(s) = 0 \quad \text{for all } s \in [0,1], \qquad u = 1, \ldots, p$$

where $\widetilde{\beta}_u$ is defined in equation (4.23).

v) $\rho_{\{i\}}^r(\mathscr{I}^{*,\mathrm{can}}(\tau)) > 0$, $i = 1, \ldots, r$.

vi) In the case that $\tau = \tau_0^c$ additionally assume that the conditions (3.11), (3.12), (3.13) and (3.17) hold.

4.3.4 Remark. Assumption 4.3.3.iv means that the degeneracy of the asymptotic information matrix of the MCRM $\mathscr{I}^{*,\mathrm{can}}(\tau)$ is only due to some linear dependency of the weight functions. Assumption 4.3.3.v excludes weight functions that are $\widetilde{\Lambda}^{(\tau_0^c)}$-almost surely 0.

Now, we can state the main result of this section. This result enables us to apply the results derived in Section 4.1 and Section 4.2 to testing problems under SHPSM.

4.3.5 Theorem. Set

$$U_n = \mathscr{U}\big(\mathscr{I}(\tau)\big)^{\mathrm{T}} S_n(\tau), \quad \widehat{U}_n = \widehat{U}_n(\tau), \quad \widehat{V}_n = \widehat{V}_n(\tau), \qquad n \in \mathbb{N},$$

and $\mathscr{I} = \mathscr{I}(\tau)$, where $\widehat{U}_n(\tau)$ is defined in Theorem 3.2.9 and $\widehat{V}_n(\tau)$ is defined in Theorem 3.2.13. Assumption 4.3.3 implies Assumption 4.1.16. (Note the representation of $\widehat{U}_n(\tau)$ given in Remark 3.2.10.b.)

Proof. As $\big(\Omega_n, \mathfrak{F}_n, \mathbb{F}_n, \mathfrak{P}_n\big)$, $n \in \mathbb{N}$, is asymptotically normal to time τ, one sees that Assumption 4.1.1 holds by paying attention to Definition 2.2.2 and Remark 2.2.3. Theorem 3.2.4, Theorem 3.2.9, Remark 3.2.10.a and Theorem 3.2.13 imply Assumption 4.1.16.a and Assumption 4.1.16.b. Since the u-th component of $\widehat{V}_n\beta$, $u = 1, \dots, r$ is given by

$$\sum_{i=1}^{n} \int_{I(\tau)} \widehat{\gamma}_n^{(\dot{u},\ddot{u})} \sum_{v=1}^{p} \left(\widehat{\mu}_{n,2}^{(\dot{u},v)} - \frac{\widehat{\mu}_{n,1}^{(u)} \widehat{\mu}_{n,1}^{(v)}}{\widehat{\mu}_{n,0}} \right) \left(\sum_{l=1}^{r_u} \widehat{\gamma}_n^{(v,l)} \widetilde{\beta}_v^{(l)} \right) \cdot \frac{1}{\widehat{\mu}_{n,0}} \, \mathrm{d}N_n^{(i)},$$

where $\widetilde{\beta}_v$ is defined in equation (4.23). Assumption 4.3.3.iii and Assumption 4.3.3.iv yield that Assumption 4.1.16.c is valid. Assumption 4.3.3.v is exactly Assumption 4.1.16.d. □

4.3.6 Corollary. In the situation of Theorem 4.3.5, the following assertions hold true.

a) For the testing problem $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathcal{J}}$, the sequence of tests $\varphi_{n,1}$, $n \in \mathbb{N}$, where

$$\varphi_{n,1} = \begin{cases} 1, & \\ 0, & \end{cases} L_{\mathcal{J},1}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big) - c_{\mathcal{J},1}\big(\alpha, \widehat{V}_n(\tau)\big) \begin{array}{c} > \\ \le \end{array} 0 \,,$$

and $\alpha \in (0, 1/2)$, keeps asymptotically the level on the hypothesis. Moreover, if the condition stated in (4.12) holds, then the sequence of tests is also asymptotically unbiased.

b) For the testing problem $\widetilde{\mathcal{H}}_2^{\mathcal{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$, the sequence of tests $\varphi_{n,2}$, $n \in \mathbb{N}$, where

$$\varphi_{n,2} = \begin{cases} 1, & \\ 0, & \end{cases} L_{\mathcal{L}_0,\mathcal{L}_1,2}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big) - c_{\mathcal{L}_0,\mathcal{L}_1,2}\big(\alpha, \widehat{V}_n(\tau)\big) \begin{array}{c} > \\ \le \end{array} 0$$

and $\alpha \in (0, 1)$, keeps asymptotically the level on the hypothesis and is asymptotically unbiased.

Moreover, both sequences of tests are asymptotically admissible, if one only considers tests that use information up to time τ. In the case that $\tau = \tau_0^c$, we use all available information, because all censored survival times are almost surely smaller than τ_0^c.

Proof. Corollary 4.1.17 and Corollary 4.2.6 yield Corollary 4.3.6.a and Corollary 4.3.6.b. The asymptotic admissibility is implied by Theorem 4.1.21 and Theorem 4.2.9. □

4.3.7 Remark. Note that the sequences of tests $\varphi_{n,1}$, $n \in \mathbb{N}$, and $\varphi_{n,2}$, $n \in \mathbb{N}$, do not depend on the choice of the foot-point α_0 and the nuisance direction $\widetilde{\gamma}$, since we consider a SHPSM. Obviously, one could extend the underlying localized, q-dimensional parametric sub-models with further nuisance directions without any effect on the asymptotic properties of the tests, as the sequence of the extended parametric sub-models is also a SHPSM. For the last conclusion it is assumed that the sequence of the extended parametric sub-models is also asymptotically normal restricted to time τ.

In particular, this means that the sequence of tests $\varphi_{n,1}$, $n \in \mathbb{N}$, is asymptotically distribution free under the hypothesis if $\mathcal{J} = \{1, \ldots, r\}$. Analogously, one sees that the sequence of tests $\varphi_{n,2}$, $n \in \mathbb{N}$, is asymptotically distribution free under the hypothesis if $\mathcal{L}_0 = 0 \in \mathbb{R}^r$.

In the case that $\mathcal{J} \neq \{1, \ldots, r\}$, the sequence of tests $\varphi_{n,1}$, $n \in \mathbb{N}$, is only asymptotically distribution free under the hypothesis if the part of the model based on the weight functions $\gamma^{(\dot{u},\ddot{u})}$, $u \in \mathcal{J}^{\complement}$, is correct. A similar consideration also holds true for the case that $\mathcal{L}_0 \neq 0$.

4.4 The Connection to Projective-Type Tests

In this section it is shown that the testing procedures derived in Section 4.3 are generalizations of well-known testing procedures by proving that our tests are projective-type tests. The latter property provides a descriptive interpretation of the test statistics $L_{\mathcal{J},1}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big)$ and $L_{\mathcal{L}_1,\mathcal{L}_0,2}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big)$. In order to keep notation simple we consider only the case that no concomitant covariates are present, this means we assume that $\mathcal{J} = \big\{1, \ldots, r\big\}$ for multivariate one-sided testing problems and $\mathcal{L}_1 = \mathbb{R}^r$ and $\mathcal{L}_0 = \{0\}$ for linear testing problems.

As it is intended to obtain a different representation of $L_{\mathcal{J},1}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big)$ and $L_{\mathcal{L}_1,\mathcal{L}_0,2}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big)$, it is necessary to introduce some more notation. We define

$$\widehat{\Lambda}_n^{(i)}(B) = \int_B \frac{1}{\widehat{\mu}_{n,0}(s)} \,\mathrm{d}N_n^{(i)}(s), \quad B \in \mathbb{B},$$

which can be interpreted as the empirical hazard measure belonging to the i-th observation, and $\widehat{\Lambda}_n^{\bullet}(B) = \sum_{i=1}^n \widehat{\Lambda}_n^{(i)}(B)$ $B \in \mathbb{B}$, which is the cumulative empirical hazard measure. Obviously, it holds that $\widehat{\Lambda}_n^{(i)} \ll \widehat{\Lambda}_n^{\bullet}$, $i = 1, \ldots, n$. The matrix $\widehat{\sigma}_n(s) = \big(\widehat{\sigma}_n^{(u,v)}(s) \mid u, v = 1, \ldots, p\big)$ is given by

$$\widehat{\sigma}_n^{(u,v)}(s) = \widehat{\mu}_{n,2}^{(u,v)}(s) - \frac{\widehat{\mu}_{n,1}^{(u)}(s)\widehat{\mu}_{n,1}^{(v)}(s)}{\widehat{\mu}_{n,0}(s)}.$$

Note that $\widehat{\sigma}_n(s)$ is positive semi-definite, see Remark 3.2.14.a. One easily sees that

$$\mathcal{V}_{\widehat{\Lambda}_n^\bullet,\widehat{\sigma}_n} = \Big\{ f \mid f : \big(I(\tau), I(\tau)\cap\mathbb{B}\big) \to \big(\mathbb{R}^p, \mathbb{B}^p\big),\ f(t) \in \mathrm{Im}\big(\widehat{\sigma}_n(t)\big)\,\forall\, t \in I(\tau)\Big\}$$

is a real vector space. An empirical pseudo inner product can be defined by

$$\langle f_1, f_2\rangle_{\widehat{\Lambda}_n^\bullet,\widehat{\sigma}_n} = \int_{I(\tau)} f_1(s)^{\mathrm{T}}\big(\widehat{\sigma}_n(s)\big)^- f_2(s)\,\mathrm{d}\widehat{\Lambda}_n^\bullet(s), \qquad f_1, f_2 \in \mathcal{V}_{\widehat{\Lambda}_n^\bullet,\widehat{\sigma}_n}.$$

By the introduction of the equivalence relation $f_1 \cong f_2$, if $||f_1 - f_2||_{\widehat{\Lambda}_n^\bullet,\widehat{\sigma}_n} = 0$, where $||f||^2_{\widehat{\Lambda}_n^\bullet,\widehat{\sigma}_n} = <f, f>_{\widehat{\Lambda}_n^\bullet,\widehat{\sigma}_n}$, one can partition the vector space $\mathcal{V}_{\widehat{\Lambda}_n^\bullet,\widehat{\sigma}_n}$ into equivalence classes. The vector space of the equivalence classes is a real Hilbert space. As a consequence of this procedure, the results provided in Appendix B.2 are applicable to $\big(\mathcal{V}_{\widehat{\Lambda}_n^\bullet,\widehat{\sigma}_n}, <\cdot,\cdot>_{\widehat{\Lambda}_n^\bullet,\widehat{\sigma}_n}\big)$.

4.4.1 Proposition. Define the function $\widehat{h}_n : I(\tau) \longrightarrow \mathbb{R}^p$, where

$$\widehat{h}_n^{(u)}(s) = \sum_{i=1}^n \big(Z_{n,i}^{(u)}(s)\,\widehat{\mu}_{n,0}(s) - \widehat{\mu}_{n,1}^{(u)}(s)\big)\,\frac{\mathrm{d}\widehat{\Lambda}_n^{(i)}}{\mathrm{d}\widehat{\Lambda}_n^\bullet}(s) \qquad s \in I(\tau).$$

It holds that

$$\widehat{\sigma}_n(s)\big(\widehat{\sigma}_n(s)\big)^-\widehat{h}_n(s) = \widehat{h}_n(s) \quad \text{for } \widehat{\Lambda}_n^\bullet\text{-all } s \in I(\tau).$$

In particular, this means that we can assume that $\widehat{h}_n \in \mathcal{V}_{\widehat{\Lambda}_n^\bullet,\widehat{\sigma}_n}$. One only has to choose reasonable versions of $\mathrm{d}\widehat{\Lambda}_n^{(i)}/\mathrm{d}\widehat{\Lambda}_n^\bullet$, $i = 1,\ldots,n$.

Proof. For verifying the assertion the same ideas as in the proof of Theorem 3.2.4.c are used. It holds that

$$\widehat{\sigma}_n^{(u,v)}(s) = \frac{1}{n}\sum_{i=1}^n \left(Z_{n,i}^{(u)}(s) - \frac{\widehat{\mu}_{n,1}^{(u)}(s)}{\widehat{\mu}_{n,0}(s)}\right)\left(Z_{n,i}^{(v)}(s) - \frac{\widehat{\mu}_{n,1}^{(v)}(s)}{\widehat{\mu}_{n,0}(s)}\right) Y_n^{(i)}(s). \quad (4.25)$$

Let us consider the three cases $\mathrm{rank}\big(\widehat{\sigma}_n(s)\big) = 0$, $0 < \mathrm{rank}\big(\widehat{\sigma}_n(s)\big) < p$ and $\mathrm{rank}\big(\widehat{\sigma}_n(s)\big) = p$. Using (4.25) one sees that $\mathrm{rank}\big(\widehat{\sigma}_n(s)\big) = 0$ implies that $Y_n^{(i)}(s) = 0$, $i = 1,\ldots,n$. Therefore we can choose $\big(\mathrm{d}\widehat{\Lambda}_n^{(i)}/\mathrm{d}\widehat{\Lambda}_n^\bullet\big)(s) = 0$, $i =$

$1, \ldots, n$. This means that $\widehat{h}_n(s) = 0$. If $\widehat{\sigma}_n(s)$ has full rank the assertion is also trivial.

Let us assume that $0 < \operatorname{rank}\big(\widehat{\sigma}_{\mathrm{n}}(\mathrm{s})\big) = k < p$ and that $\mathcal{M} = \big\{v_1, \ldots, v_k\big\}$ are the indices of k linearly independent columns of $\widehat{\sigma}_n(s)$. For every $u \in \big\{1, \ldots, p\big\} \backslash \mathcal{M}$ there exists $c_u \in \mathbb{R}^k$, such that

$$\begin{aligned}
\widehat{\sigma}_n^{(u,u)}(s) &= \frac{1}{n} \sum_{i=1}^{n} \left(\left(Z_{n,i}^{(u)}(s) - \frac{\widehat{\mu}_{n,1}^{(u)}(s)}{\widehat{\mu}_{n,0}(s)} \right) \sum_{l=1}^{k} c_u^{(l)} \left(Z_{n,i}^{(v_l)}(s) - \frac{\widehat{\mu}_{n,1}^{(v_l)}(s)}{\widehat{\mu}_{n,0}(s)} \right) \right) Y_n^{(i)}(s) \\
&= \frac{1}{n} \sum_{i=1}^{n} \left(\sum_{l=1}^{k} c_u^{(l)} \left(Z_{n,i}^{(v_l)}(s) - \frac{\widehat{\mu}_{n,1}^{(v_l)}(s)}{\widehat{\mu}_{n,0}(s)} \right) \right)^2 Y_n^{(i)}(s)
\end{aligned}$$

implying

$$\frac{1}{n} \sum_{i=1}^{n} \left(Z_{n,i}^{(u)}(s) - \frac{\widehat{\mu}_{n,1}^{(u)}(s)}{\widehat{\mu}_{n,0}(s)} - \sum_{l=1}^{k} c_u^{(l)} \left(Z_{n,i}^{(v_l)}(s) - \frac{\widehat{\mu}_{n,1}^{(v_l)}(s)}{\widehat{\mu}_{n,0}(s)} \right) \right)^2 Y_n^{(i)}(s) = 0.$$

With the same considerations as before we get that

$$\widehat{h}_n^{(u)}(s) = \sum_{l=1}^{k} c_u^{(l)} \widehat{h}_n^{(v_l)}(s), \quad \text{if } \widehat{\Lambda}_n^{\bullet}\big(\{s\}\big) > 0.$$

As $\widehat{\sigma}_n(s)$ is symmetric, we can conclude that the rank of the extended matrix $\big(\widehat{\sigma}_n(s) \mid \widehat{h}_n(s)\big)$ is also k. Proposition B.1.5 yields the assertion. □

The function $\widehat{h}_n$ can be interpreted as a primitive estimator for the influence of the covariates on the survival function of the survival times in question. For the u-th component we have

$$\begin{aligned}
Z_{n,i}^{(u)}(s)\widehat{\mu}_{n,0}(s) - \widehat{\mu}_{n,1}^{(u)}(s) &= Z_{n,i}^{(u)}(s)\, \frac{1}{n} \sum_{i=1}^{n} Y_n^{(i)}(s) - \frac{1}{n} \sum_{i=1}^{n} Z_{n,i}^{(u)}(s)\, Y_n^{(i)}(s) \\
&\approx \mathbb{E}\big(Z^{(u)}(s)\big)\mathbb{E}\big(Y(s)\big) - \mathbb{E}\big(Z(s)^{(u)}\, Y(s)\big) = -\operatorname{Cov}\big(Z^{(u)}(s), Y(s)\big),
\end{aligned}$$

where $\mathbb{E}\big(Y(s)\big)$ is the probability that a censored observation is larger than s. Without censoring, it would be the probability that an individual survives longer than s. So the function $\widehat{h}_n$ measures the correlation of covariates and survival times.

Under the hypothesis we expect the survival times and the covariates to be uncorrelated, so that the function $\widehat{h}_n$ should vary around 0.

Before we start rewriting the statistic $\widehat{U}_n(\tau)$ and the variance estimator, we artificially rewrite the weight functions as functions that are elements of $\mathcal{V}_{\widehat{\Lambda}^{\bullet}_n,\widehat{\sigma}_n}$.

$$\widehat{w}_{n,u} : I(\tau) \longrightarrow \mathbb{R}^p, \quad \widehat{w}_{n,u}(s) = \frac{\widehat{\gamma}_n^{(\dot{u},\ddot{u})}(s)}{\sqrt{n}} \cdot \widehat{\sigma}_n(s)\mathscr{T}^p_{\{\dot{u}\}}, \quad s \in I(\tau),$$

$u = 1, \ldots, r$, where $\mathscr{T}^p_{\{\dot{u}\}}$ is given in Definition 4.1.3.

4.4.2 Theorem. Let us define the closed, convex cone

$$\Gamma_n^+ = \Big\{ \sum\nolimits_{u=1}^{r} \beta^{(u)} \cdot \widehat{w}_{n,u} \;\big|\; \beta^{(u)} \geq 0,\ u = 1, \ldots, r \Big\} \subset \mathcal{V}_{\widehat{\Lambda}^{\bullet}_n,\widehat{\sigma}_n}$$

and the linear space

$$\Gamma_n = \Big\{ \sum\nolimits_{u=1}^{r} \beta^{(u)} \cdot \widehat{w}_{n,u} \;\big|\; \beta \in \mathbb{R}^r \Big\} \subset \mathcal{V}_{\widehat{\Lambda}^{\bullet}_n,\widehat{\sigma}_n}.$$

If $\mathcal{J} = \big\{1, \ldots, r\big\}$, $\mathcal{L}_1 = \mathbb{R}^r$ and $\mathcal{L}_0 = \{0\}$ then the following assertions hold true, where $\Pi_{\Gamma_n^+}(h_n)$ and $\Pi_{\Gamma_n}(h_n)$ denote the projections of $\widehat{h}_n$ on Γ_n^+ and Γ_n with respect to $<\cdot,\cdot>_{\widehat{\Lambda}^{\bullet}_n,\widehat{\sigma}_n}$.

a) $L_{\mathcal{J},1}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big) = \big|\big|\Pi_{\Gamma_n^+}(h_n)\big|\big|^2_{\widehat{\Lambda}^{\bullet}_n,\widehat{\sigma}_n}$.

b) $L_{\mathcal{L}_1,\mathcal{L}_0,2}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big) = \big|\big|\Pi_{\Gamma_n}(h_n)\big|\big|^2_{\widehat{\Lambda}^{\bullet}_n,\widehat{\sigma}_n}$.

Proof. We readily check that

$$\begin{aligned}
\widehat{V}_n^{(u,v)}(\tau) &= \frac{1}{n}\sum_{i=1}^{n} \int_{I(\tau)} \widehat{\gamma}_n^{(\dot{u},\ddot{u})}\, \widehat{\gamma}_n^{(\dot{v},\ddot{v})} \left(\widehat{\mu}_{n,2}^{(\dot{u},\dot{v})} - \frac{\widehat{\mu}_{n,1}^{(\dot{u})}\widehat{\mu}_{n,1}^{(\dot{v})}}{\widehat{\mu}_{n,0}} \right) \mathrm{d}\widehat{\Lambda}_n^{(i)} \\
&= \frac{1}{n}\int_{I(\tau)} \widehat{\gamma}^{(\dot{u},\ddot{u})} \cdot \mathscr{T}^{p\;\mathrm{T}}_{\{\dot{u}\}} \widehat{\sigma}_n \mathscr{T}^p_{\{\dot{v}\}} \cdot \gamma^{(\dot{v},\ddot{v})}\, \mathrm{d}\widehat{\Lambda}^{\bullet}_n \\
&= \big\langle \widehat{w}_{n,u}, \widehat{w}_{n,v} \big\rangle_{\widehat{\Lambda}^{\bullet}_n,\widehat{\sigma}_n}.
\end{aligned}$$

and that

$$
\begin{aligned}
\widehat{U}_n^{(u)}(\tau) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{I(\tau)} \widehat{\gamma}_n^{(\dot{u},\ddot{u})} \left(Z_{n,i}^{(\dot{u})} - \frac{\widehat{\mu}_{n,1}^{(\dot{u})}}{\widehat{\mu}_{n,0}} \right) \mathrm{d}N_n^{(i)} \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{I(\tau)} \widehat{\gamma}_n^{(\dot{u},\ddot{u})} \left(Z_{n,i}^{(\dot{u})} \widehat{\mu}_{n,0} - \widehat{\mu}_{n,1}^{(\dot{u})} \right) \frac{1}{\widehat{\mu}_{n,0}} \,\mathrm{d}N_n^{(i)} \\
&= \frac{1}{\sqrt{n}} \int_{I(\tau)} \widehat{\gamma}_n^{(\dot{u},\ddot{u})} \sum_{i=1}^{n} \left(Z_{n,i}^{(\dot{u})}\, \widehat{\mu}_{n,0} - \widehat{\mu}_{n,1}^{(\dot{u})} \right) \frac{\mathrm{d}\widehat{\Lambda}_n^{(i)}}{\mathrm{d}\widehat{\Lambda}_n^{\bullet}} \,\mathrm{d}\widehat{\Lambda}_n^{\bullet} \\
&= \frac{1}{\sqrt{n}} \int_{I(\tau)} \widehat{\gamma}_n^{(\dot{u},\ddot{u})}\, \widehat{h}_n^{(\dot{u})} \,\mathrm{d}\widehat{\Lambda}_n^{\bullet} \\
&= \frac{1}{\sqrt{n}} \int_{I(\tau)} \gamma^{(\dot{u},\ddot{u})} \cdot \mathscr{T}_{\{\dot{u}\}}^{p}{}^{\mathrm{T}} \widehat{\sigma}_n\, \widehat{\sigma}_n^{-}\, \widehat{h}_n \,\mathrm{d}\widehat{\Lambda}_n^{\bullet} \\
&= \left\langle \widehat{w}_{n,u}, \widehat{h}_n \right\rangle_{\widehat{\Lambda}_n^{\bullet}, \widehat{\sigma}_n}.
\end{aligned}
$$

For any $\beta \in \ker\big(\widehat{V}_n(\tau)\big)$ we get that

$$
\Big|\Big| \sum\nolimits_{u=1}^{r} \beta^{(u)} \widehat{w}_{n,u} \Big|\Big|^2_{\widehat{\Lambda}_n^{\bullet}, \widehat{\sigma}_n} = 0,
$$

see Proposition B.3.2.b, implying $\beta^{\mathrm{T}} \widehat{U}_n(\tau) = 0$ and $\widehat{U}_n(\tau) \in \mathrm{Im}\big(\widehat{V}_n(\tau)\big)$, where we use Proposition B.3.2.a. Applying Proposition B.2.5.e and Proposition B.2.5.b gives

$$
\begin{aligned}
L_{\mathcal{J},1}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big) &= 2 \sup_{\beta \geq 0} \left(\beta^{\mathrm{T}} \widehat{U}_n(\tau) - \frac{1}{2} \beta^{\mathrm{T}} \widehat{V}_n(\tau) \beta \right) \\
&= 2 \sup_{w \in \Gamma_n^{+}} \left(<\widehat{h}_n, w>_{\widehat{\Lambda}_n^{\bullet}, \widehat{\sigma}_n} - \frac{1}{2} <w, w>_{\widehat{\Lambda}_n^{\bullet}, \widehat{\sigma}_n} \right).
\end{aligned}
$$

Proposition B.2.5.b yields the first assertion. Moreover, Proposition B.2.5.d and Proposition B.2.5.b imply that

$$
\begin{aligned}
L_{\mathcal{L}_1, \mathcal{L}_0, 2}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big) &= \Big|\Big| \Pi_{\mathrm{Im}\left(\widehat{V}_n(\tau)\right)} \big(\widehat{U}_n(\tau)\big) \Big|\Big|^2_{\left(\widehat{V}_n(\tau)\right)^{-}} \\
&= 2 \sup_{\beta \in \mathbb{R}^r} \left(\beta^{\mathrm{T}} \widehat{U}_n(\tau) - \frac{1}{2} \beta^{\mathrm{T}} \widehat{V}_n(\tau) \beta \right) \\
&= 2 \sup_{w \in \Gamma_n} \left(<\widehat{h}_n, w>_{\widehat{\Lambda}_n^{\bullet}, \widehat{\sigma}_n} - \frac{1}{2} <w, w>_{\widehat{\Lambda}_n^{\bullet}, \widehat{\sigma}_n} \right)
\end{aligned}
$$

Again, Proposition B.2.5.b yields the assertion. □

4.4.3 Remark. In the situation of Theorem 4.4.2, one sees that the test statistics $L_{\mathcal{J},1}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big)$ and $L_{\mathcal{L}_1,\mathcal{L}_0,2}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big)$ are projections of a primitive estimator for the correlation between covariates and survival function on the closed, convex cones $\widehat{\Gamma}_n^+$ and $\widehat{\Gamma}_n$, respectively. If this projection is too large, *i.e.* is not too close to 0, the test rejects the hypothesis, since there seems to be some influence of the covariates on the survival times. Moreover, one sees that the cones determine the alternatives the test is sensitive for.

In the case that the weight functions are of the form $\widehat{\gamma}_n^{(\dot{u},\ddot{u})} = \gamma_0^{(\dot{u},\ddot{u})} \circ \widehat{H}_n$, $u = 1, \ldots, r$, cf. Assumption 4.3.3.iii, where $\widehat{H}_n$ is either a left continuous version of the Kaplan-Meyer estimator, see Andersen *et al.* [4, Chapter IV.3], or $\widehat{H}_n = 1 - \frac{1}{n}\sum_{i=1}^n Y_n^{(i)}$, the testing procedures derived in the previous sections are indeed non-parametric procedures. This can be seen as follows. By choosing the functions $\gamma_0^{(\dot{u},\ddot{u})} : [0,1] \longrightarrow \mathbb{R}$, the statistician decides whether it is intended to weight early or late influences of the covariates on the survival times. As the functions $\gamma_0^{(\dot{u},\ddot{u})}$, $u = 1, \ldots, r$, are defined on the interval $[0,1]$, the terms *early* and *late* can be given a meaning. The empirical cumulative distribution function $\widehat{H}_n$ provides the right transformation of the interval $[0,1]$ onto $\mathbb{R}_+$. Some aspects concerning the sign of the weight functions were already worked out in Discussion 4.3.2. In the case that $\widehat{\gamma}_n^{(\dot{u},\ddot{u})} = \gamma_0^{(\dot{u},\ddot{u})} \circ \widehat{H}_n$, $u = 1, \ldots, r$, one can also easily see that our tests are invariant with respect to changes of the time scale in the sense of Remark 1.3.3.c. Results by Janssen [36] suggest, that any test keeping the level on the hypothesis can have reasonable power only for a finite number of orthogonal directions of alternatives. Therefore, restricting ourselves to finite dimensional, closed, convex cones and linear spaces is no restriction in practice, but reveals the advantage of the testing procedure suggested in Section 4.3. The statistician can control the weight functions and the number of weight functions. This means that the statistician can control the alternatives the test is sensitive for.

Behnen and Neuhaus [7] and Mayer [53] have introduced similar tests for the two-sample problem. Behnen and Neuhaus [7, Chapter 3] derive some primitive estimator for the difference of the distributions of the two samples and project this estimator on a cone of score functions. An extension of this proceeding to right censored data can be found in Behnen and Neuhaus [8].

Mayer [53] introduced some empirical inner product and showed that log-rank statistics are projections of a primitive estimator on some one-dimensional cone, if one considers a one-sided testing problem, or on a one-dimensional linear space, if one considers two-sided testing problems. Then Mayer replaces the one-dimensional cone and linear space by higher dimensional cones and spaces that are generated by weight functions and investigates the asymptotic properties of the new test statistics. One easily sees that the projective-type tests of Mayer are special cases of the tests proposed in Section 4.3, cf. also Example 5.3.2.

If we consider the case $p = r = 1$ then our test statistic $\widehat{U}_n(\tau_0)$ belongs to the general class of non-parametric test statistics introduced by Jones and Crowley [39, 40]. Hence, the non-parametric test statistics $L_{\mathcal{J},1}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big)$ and $L_{\mathcal{L}_1,\mathcal{L}_0,2}\big(\widehat{U}_n(\tau), \widehat{V}_n(\tau)\big)$ generalize the statistics introduced by Jones and Crowley, because we allow multivariate covariates and several weight functions instead of one weight function.

5 Examples and Applications

In this chapter we provide applications of the theory developed in the previous chapters. In Section 5.3 several statistical questions are modelled with help of the modified Cox Regression Model (MCRM). And once again, it is shown that our results are extensions of well-known results. But initially, in Section 5.1 and Section 5.2 the existence of parametric sub-model and the Assumptions of Chapter 2 and Chapter 3 are discussed.

5.1 On the Existence of the Modified Cox Regression Model

In this Section, it is aimed to explicitly construct sequences of filtered probability spaces satisfying Assumption 2.2.1. The starting point is given by some stochastic processes, whose paths are supposed to determine the distribution of survival times. The whole construction is carried out in the spirit of Proposition B.5.4. First, let us introduce some notation and premises.

5.1.1 Assumption. i) Let

$$\big(\Omega^*_{n,i}, \mathcal{F}^*_{n,i}, \mathbb{F}^*_{n,i}, Q^*_{n,i}\big), \quad \mathbb{F}^*_{n,i} = \{\mathcal{F}^*_{n,i,t} \mid t \in \mathbb{R}_+\}, \qquad i = 1, \ldots, n,$$

be filtered probability spaces and

$$\big(\Omega'_{n,i}, \mathcal{N}'_{n,i}\big) = \big(\{0,1\} \times (0,\infty), \mathcal{P}\{0,1\} \otimes \mathbb{B}(0,\infty)\big), \qquad i = 1, \ldots, n,$$

measurable spaces, where $\mathcal{P}\{0,1\}$ denotes the power set of $\{0,1\}$ and $\mathbb{B}(0,\infty) = \big\{B \cap (0,\infty) \mid B \in \mathbb{B}\big\}$. Moreover, we set

$$\Omega_n = \bigtimes_{i=1}^{n} \Omega_{n,i}, \qquad \text{where} \quad \Omega_{n,i} = \Omega^*_{n,i} \times \Omega'_{n,i}, \quad i = 1, \dots, n,$$

and

$$\mathcal{F}_n = \bigotimes_{i=1}^{n} \mathcal{F}_{n,i}, \qquad \text{where} \quad \mathcal{F}_{n,i} = \mathcal{F}^*_{n,i} \otimes \mathcal{N}'_{n,i}, \quad i = 1, \dots, n.$$

In the following, the notation

$$\omega_n = (\omega_{n,1}, \dots, \omega_{n,n}) = (\omega^*_{n,1}, \omega'_{n,1} \dots, \omega^*_{n,n}\omega'_{n,n}) \in \Omega_n$$

is used.

ii) Assume that there exits measurable mappings

$$Z^*_{n,i,t} : \Omega^*_{n,i} \longrightarrow \mathbb{R}^p, \quad t \in \mathbb{R}_+, \qquad i = 1, \dots, n,$$

that satisfy the following conditions.

a) $\mathcal{F}^*_{n,i,t} = \bigcap_{s>t} \mathcal{F}^{*,0}_{n,i,s}$, where $\mathcal{F}^{*,0}_{n,i,s} = \sigma\big(Z^*_{n,i,u} \mid u \le s\big)$, $i = 1, \dots, n$.

b) $\bigvee_{t \ge 0} \mathcal{F}^*_{n,i,t} = \mathcal{F}^*_{n,i}$, $i = 1, \dots, n$.

c) The process $\{Z^*_{n,i,t} \mid t \in \mathbb{R}_+\}$ is progressively measurable, *i.e.* for all $t \in \mathbb{R}_+$, the mapping $(\omega^*_{n,i}, s) \mapsto Z^*_{n,i,s}(\omega^*_{n,i})$, $s \le t$, is $\mathcal{F}^*_{n,i,t} \otimes \mathbb{B}[0,t]$–$\mathbb{B}^p$ measurable.

iii) Define the mapping $h_n : \Omega_n \longrightarrow \mathbb{R}_+$, $h_n(\omega_n) = \prod_{i=1}^n h_{n,i}(\omega^*_{n,i}, \omega'_{n,i})$,

$$h_{n,i}(\omega^*_{n,i}, \omega'_{n,i}) = \exp\Big(- \int_{[0,u_{n,i}]} h_{n,i,1}\big(Z^*_{n,i,s}(\omega^*_{n,i}), s\big) \, \mathrm{d}\lambda(s)\Big) \times h_{n,i,2}\big(Z^*_{n,i,u_{n,i}}(\omega^*_{n,i}), \omega'_{n,i}\big),$$

where $\omega'_{n,i} = (\delta_{n,i}, u_{n,i})$, λ denotes the Lebesgue measure on $\mathbb{R}_+$ and $h_{n,i,j} : \mathbb{R}^{p+j} \longrightarrow \mathbb{R}_+$, $j = 1, 2$, $i = 1, \dots, n$, are $\mathbb{B}^{p+j}$–$\mathbb{B}_+$ measurable.

iv) Assume that α_0 and $\widetilde{\alpha}_{n,i}$, $i = 1, \dots, n$, are hazard rates of some probability measures on $\mathbb{B}(0, \infty)$. Set

$$h_{n,i,1}(z, s) = \big(R_n(z, s)\, \alpha_0(s) + \widetilde{\alpha}_{n,i}(s)\big) \cdot \mathbb{1}(s < \tau_{n,i})$$

and

$$h_{n,i,2}(z, \delta, s) = \big(R_n(z, s)\, \alpha_0(s) \cdot \mathbb{1}_{\{1\}}(\delta) + \widetilde{\alpha}_{n,i}(s) \cdot \mathbb{1}_{\{0\}}(\delta)\big) \cdot \mathbb{1}(s < \tau_{n,i}),$$

$(z, s) \in \mathbb{R}^{p+1}$, $(z, \delta, s) \in \mathbb{R}^{p+2}$, $i = 1, \dots, n$, where

$$\tau_{n,i} = \sup\Big\{ t \;\Big|\; \int_{[0,t]} \alpha_0 + \widetilde{\alpha}_{n,i} \,\mathrm{d}\lambda < \infty \Big\}$$

and $R_n : \mathbb{R}^{p+1} \longrightarrow \mathbb{R}_+$ is some $\mathbb{B}^{p+1}$–$\mathbb{B}_+$ measurable mapping. Moreover, let us assume that $\alpha_0(t) = 0$ for all $t \geq \tau_0$, if $\tau_0 < \infty$, where

$$\tau_0 = \sup\Big\{ t \;\Big|\; \int_{I(t)} \alpha_0(s) \,\mathrm{d}s < \infty \Big\}.$$

v) $\mu_n = \bigotimes_{i=1}^n \mu_{n,i}$, $\mu_{n,i} = Q^*_{n,i} \otimes \nu_c \otimes \lambda$, where ν_c denotes the counting measure on $\mathcal{P}\{0, 1\}$. .

vi) We define the mappings $N'_{n,i,t}, \widetilde{N}_{n,i,t} : \Omega'_{n,i} \longrightarrow \mathbb{R}$, where

$$N'_{n,i,t}(\delta_{n,i}, s_{n,i}) = \mathbb{1}(s_{n,i} \leq t)\delta_{n,i}$$

and

$$\widetilde{N}'_{n,i,t}(\delta_{n,i}, s_{n,i}) = \mathbb{1}(s_{n,i} \leq t)(1 - \delta_{n,i}),$$

and $Y'_{n,i,t} : \Omega'_{n,i} \longrightarrow \mathbb{R}_+$,

$$Y'_{n,i,t}(\delta_{n,i}, s_{n,i}) = \mathbb{1}(s_{n,i} \geq t),$$

$\omega'_{n,i} = (\delta_{n,i}, s_{n,i})$, $t \in \mathbb{R}_+$, $i = 1, \dots, n$.

vii) We define the mappings $A'_{n,i,t}, \widetilde{A}'_{n,i,t} : \Omega_{n,i} \longrightarrow \bar{\mathbb{R}}_+$, where

$$A'_{n,i,t}(\omega^*_{n,i}, \omega'_{n,i}) = \int_{[0,t]} Y'_{n,i,s}(\omega'_{n,i})\, R_n\big(Z^*_{n,i,s}(\omega^*_{n,i}), s\big)\, \alpha_0(s) \,\mathrm{d}\lambda(s)$$

and
$$\widetilde{A}'_{n,i,t}(\omega^*_{n,i},\omega'_{n,i}) = \int_{[0,t]} Y'_{n,i,s}(\omega'_{n,i})\,\widetilde{\alpha}_{n,i}(s)\,\mathrm{d}\lambda(s),$$
$(\omega^*_{n,i},\omega'_{n,i}) \in \Omega_{n,i}$, $\omega'_{n,i} = (\delta_{n,i}, u_{n,i})$, $t \in \mathbb{R}_+$, $i = 1,\dots,n$.

viii) Moreover, we define the mappings $Z_{n,i}(t) : \Omega_n \longrightarrow \mathbb{R}^p$ and
$$N_n^{(i)}(t), \widetilde{N}_n^{(i)}(t), Y_n^{(i)}(t), A_n^{(i)}(t), \widetilde{A}_n^{(i)}(t) : \Omega_n \longrightarrow \mathbb{R}_+,$$
by setting
$$\begin{aligned} N_n^{(i)}(t) &= N'_{n,i,t}\circ\varpi'_{n,i}, & \widetilde{N}_n^{(i)}(t) &= \widetilde{N}'_{n,i,t}\circ\varpi'_{n,i},\\ Z_{n,i}(t) &= Z^*_{n,i,t}\circ\varpi^*_{n,i}, & Y_n^{(i)}(t) &= Y'_{n,i,t}\circ\varpi'_{n,i},\\ A_n^{(i)}(t) &= A'_{n,i,t}\circ(\varpi^*_{n,i},\varpi'_{n,i}), & \widetilde{A}_n^{(i)}(t) &= \widetilde{A}'_{n,i,t}\circ(\varpi^*_{n,i},\varpi'_{n,i}),\end{aligned}$$
$t \in \mathbb{R}_+$, $i = 1,\dots,n$, where $\varpi^*_{n,i} : \Omega_n \longrightarrow \Omega^*_{n,i}$ and $\varpi'_{n,i} : \Omega_n \longrightarrow \Omega'_{n,i}$ denote coordinate projections, *i.e.* $\varpi^*_{n,i}(\omega_n) = \omega^*_{n,i}$ and $\varpi'_{n,i}(\omega_n) = \omega'_{n,i}$.

ix) Let $\mathcal{S}_n$ be some σ-algebra on the space Ω_n. In the following the important case will be that $\mathcal{S}_n$ is generated by the subsets of negligible sets of some probability measure. Finally, we define the σ-algebras
$$\mathcal{N}'_{n,i,t} = \sigma\big(N'_{n,i,s}, \widetilde{N}'_{n,i,s}, \mid s \le t\big), \qquad t \in \mathbb{R}_+,$$
and set $\mathbb{G}_n = \{\mathcal{G}_{n,t} \mid t \in \mathbb{R}_+\}$, where $\mathcal{G}_{n,t} = \mathcal{S}_n \vee \bigotimes_{i=1}^n \big(\mathcal{F}^*_{n,i} \otimes \mathcal{N}'_{n,i,t}\big)$. Moreover, we define the filtration $\mathbb{H}_n = \{\mathcal{H}_{n,t} \mid t \in \mathbb{R}_+\}$, where $\mathcal{H}_{n,t} = \bigcap_{s>t} \mathcal{H}^0_{n,s}$, $\mathcal{H}^0_{n,s} = \mathcal{S}_n \vee \bigotimes_{i=1}^n \big(\mathcal{F}^*_{n,i,s} \otimes \mathcal{N}'_{n,i,s}\big)$.

5.1.2 Proposition (Properties of the Filtration). Under Assumption 5.1.1, it holds that

a) $\mathcal{G}_{n,t} = \mathcal{S}_n \vee \mathcal{G}_n \vee \sigma\big(N_n^{(i)}(s), \widetilde{N}_n^{(i)}(s) \mid s \le t,\, i = 1,\dots,n\big)$, $t \in \mathbb{R}_+$, where
$$\mathcal{G}_n = \sigma\left(\bigtimes_{i=1}^n (F_i \times \Omega'_{n,i}) \mid F_i \in \mathcal{F}^*_{n,i} \right).$$

b) $\mathbb{G}_n$ and $\mathbb{H}_n$ are indeed increasing and right continuous.

c) $\mathcal{H}_{n,t} \subset \mathcal{G}_{n,t}$.

d) $\bigvee_{t\geq 0} \mathcal{G}_{n,t} = \bigvee_{t\geq 0} \mathcal{H}_{n,t} = \mathcal{S}_n \vee \mathcal{F}_n$.

Proof. It holds that

$$\begin{aligned}\mathcal{G}_{n,t} &= \mathcal{S}_n \vee \sigma\Bigg(\bigtimes_{i=1}^{n} (F_i^* \times N_i') \mid F_i^* \in \mathcal{F}_{n,i}^*,\, N_i \in \mathcal{N}_{n,i,t}'\Bigg)\\ &= \mathcal{S}_n \vee \sigma\Bigg(\bigtimes_{i=1}^{n} (F_i^* \times \Omega_{n,i}') \mid F_i^* \in \mathcal{F}_{n,i}^*\Bigg) \vee \sigma\Bigg(\bigtimes_{i=1}^{n} (\Omega_{n,i}^* \times N_i') \mid N_i \in \mathcal{N}_{n,i,t}'\Bigg).\end{aligned}$$

As $\sigma\big(\bigtimes_{i=1}^{n}(\Omega_{n,i}^* \times N_i') \mid N_i \in \mathcal{N}_{n,i,t}'\big) = \sigma\big(N_{n,i,s} \mid s \leq t,\, i = 1, \ldots, n\big)$ the proof of a) is complete. Proof of b). We note that $\{\mathcal{H}_{n,t}^0 \mid t \in \mathbb{R}_+\}$ is increasing by construction, therefore $\mathbb{H}_n$ is increasing. The right continuity is also given by the construction of $\mathbb{H}_n$. a) and Proposition B.5.1 imply the result for $\mathbb{G}_n$. Assume $H \in \mathcal{H}_{n,t}$. Because of $\mathcal{H}_{n,s} \subset \mathcal{H}_{n,s}^0$ and b), we get $H \in \mathcal{G}_{n,s}$ for all $s > t$. The right continuity of $\mathbb{G}_n$ gives c). Proof of d). Because of the previous inclusion, it suffices to show the assertion $\bigvee_{t\geq 0} \mathcal{H}_{n,t} = \mathcal{F}_n \vee \mathcal{S}_n$. One readily checks that $\bigvee_{t\geq 0} \mathcal{N}_{n,i,t}' = \mathcal{N}_{n,i}'$. Using this result and Assumption 5.1.1.ii make the assertion an easy consequence of Proposition B.5.4.a. □

The function h_n is a candidate for a μ_n-density of some probability measure. As a first step we show that h_n is a measurable mapping.

5.1.3 Proposition. Under Assumption 5.1.1 without articles iv) and v) the mappings $h_{n,i}$, $i = 1, \ldots, n$, are $\mathcal{F}_{n,i}^* \otimes \mathcal{N}_{n,i}'$–$\mathbb{B}_+$ measurable. Consequently, the mapping h_n is $\mathcal{F}_n$–$\mathbb{B}_+$ measurable.

Proof. Consider the measurable space $\big(\Omega_{n,i}^* \times \{0,1\}, \mathcal{F}_{n,i}^* \otimes \mathcal{P}\{0,1\}\big)$. Clearly, the processes $\big\{(Z_{n,i,t}^*, t) \mid t \in \mathbb{R}_+\big\}$ and $\big\{(Z_{n,i,t}^*, \delta, t) \mid t \in \mathbb{R}_+\big\}$ are progressively measurable with respect to the filtration $\{\mathcal{F}_{n,i,t}^* \otimes \mathcal{P}\{0,1\} \mid t \in \mathbb{R}_+\}$, see Proposition B.5.4.b. Proposition B.5.2.c yields that the mappings $g_1 : \Omega_{n,i} \times \{0,1\} \times [0,t] \longrightarrow \mathbb{R}_+$ and $g_2 : \Omega_{n,i}^* \times \{0,1\} \times [0,t] \longrightarrow \mathbb{R}_+$, where

$$g_1(\omega, \delta, u) = \exp\bigg(-\int_{[0,u]} h_{n,i,1}\big(Z_{n,i,s}^*(\omega), s\big)\,\mathrm{d}\lambda(s)\bigg)$$

and

$$g_2(\omega, \delta, u) = h_{n,i,2}\big(Z^*_{n,i,u}(\omega), \delta, u\big),$$

are $\mathcal{F}^*_{n,i,t} \otimes \mathcal{P}\{0,1\} \otimes \mathbb{B}[0,t]$–$\mathbb{B}_+$ measurable. Consequently, $g_1 \cdot g_2 : \Omega^*_{n,i} \times \{0,1\} \times [0,t] \longrightarrow \mathbb{R}_+$ is $\mathcal{F}^*_{n,i,t} \otimes \mathcal{P}\{0,1\} \otimes \mathbb{B}[0,t]$–$\mathbb{B}_+$ measurable. Now, we consider the mapping $g = g_1 \cdot g_2$ on the space $\Omega^*_{n,i} \times \{0,1\} \times \mathbb{R}_+$. It holds that

$$\big\{(\omega, \delta, s) \mid g(\omega, \delta, s) \in B\big\} = \bigcup_{t=1}^{\infty} \big\{(\omega, \delta, s) \mid g(\omega, \delta, s) \in B\big\} \cap \{(\omega, \delta, s) \mid s \leq t\},$$

$B \in \mathbb{B}_+$, and

$$\big\{(\omega, \delta, s) \mid g(\omega, \delta, s) \in B\big\} \cap \{(\omega, \delta, s) \mid s \leq t\} \in \mathcal{F}^*_{n,i,t} \otimes \mathcal{P}\{0,1\} \otimes \mathbb{B}_+[0,t].$$

As $\mathcal{F}^*_{n,i,t} \otimes \mathcal{P}\{0,1\} \otimes \mathbb{B}_+[0,t] \subset \mathcal{F}^*_{n,i} \otimes \mathcal{P}\{0,1\} \otimes \mathbb{B}_+$, it holds that

$$\big\{(\omega, \delta, s) \mid g(\omega, \delta, s) \in B\big\} \cap \big\{(\omega, \delta, s) \mid 0 < s < \infty\big\} \in \mathcal{F}^*_{n,i} \otimes \mathcal{P}\{0,1\} \otimes \mathbb{B}_+.$$

$\mathcal{F}^*_{n,i} \otimes \mathcal{P}\{0,1\} \otimes \mathbb{B}_+ \cap \big\{(\omega, \delta, s) \mid 0 < s < \infty\big\} = \mathcal{F}^*_{n,i} \otimes \mathcal{N}'_{n,i}$ and Bauer [6, Bemerkung 2, p. 153] give the assertion. □

5.1.4 Remark. Let $T, C : \big(\Omega, \mathcal{A}, \mathbb{P}\big) \longrightarrow \big(\mathbb{R}_+, \mathbb{B}_+\big)$ be stochastically independent random variables. Furthermore, assume that α_T and α_C are hazard rates of the measures $\mathbb{P}^T$ and $\mathbb{P}^C$. If $X = T \wedge C$ and $\Delta = \mathbb{1}(T \leq C)$ then it holds that

$$\begin{aligned}\mathbb{P}\big(X \leq x, \, \Delta = \delta\big) = \mathbb{1}_{\{0\}}(\delta) \cdot &\int_{[0,x]} \exp\Big(-\int_{[0,s]} \alpha_T + \alpha_C \, \mathrm{d}\lambda\Big) \cdot \alpha_C(s) \, \mathrm{d}\lambda(s) \\ + \mathbb{1}_{\{1\}}(\delta) \cdot &\int_{[0,x]} \exp\Big(-\int_{[0,s]} \alpha_T + \alpha_C \, \mathrm{d}\lambda\Big) \cdot \alpha_T(s) \, \mathrm{d}\lambda(s)\end{aligned}$$

The result can be used to construct μ_n-densities of probability measures on $\mathcal{F}_n$. The previous remark gives a probability measure on the measurable space $\big(\Omega'_{n,i}, \mathcal{N}'_{n,i}\big)$ and the $(\nu_C \otimes \lambda)$-density of that probability measure. Now, one just constructs for almost every paths of the covariate processes $Z^*_{n,i}$ a probability measure of the above type. The next but one result justifies this proceeding.

5.1.5 Lemma. Set

$$\tau^*_{n,i}(\omega^*_{n,i}) = \sup\Big\{ t \;\Big|\; \int_{[0,t]} R_n\big(Z^*_{n,i,s}(\omega^*_{n,i}), s\big)\, \alpha_0 + \widetilde{\alpha}_{n,i} \,\mathrm{d}\lambda < \infty \Big\}. \tag{5.1}$$

It holds that $\tau^*_{n,i}$ is $\mathcal{F}^*_{n,i}$–$\bar{\mathbb{B}}$ measurable.

Proof. One readily shows that

$$X_{n,i}(t) = \int_{[0,t]} R_n\big(Z^*_{n,i,s}(\cdot), s\big)\, \alpha_0 + \widetilde{\alpha}_{n,i} \,\mathrm{d}\lambda,$$

is $\mathcal{F}^*_{n,i}$–$\bar{\mathbb{B}}_+$ measurable, see Proposition B.5.2. As

$$\big\{\tau^*_{n,i} > c\big\} = \big\{X_{n,i}(c) < \infty\big\}, \qquad c \in \mathbb{R},$$

where we use the continuity of the paths, it results that $\tau^*_{n,i}$ is $\mathcal{F}^*_{n,i}$–$\bar{\mathbb{B}}_+$ measurable. □

5.1.6 Proposition (Existence of Probability Measures). Suppose that Assumption 5.1.1 holds and that the processes $\{Z^*_{n,i,t} \mid t \in \mathbb{R}_+\}$, $i = 1, \dots, n$, are predictable and locally bounded. In particular this means that the processes are progressively measurable, cf. Dellacherie and Meyer [16, IV.67]. If

$$\int_{I(\tau_{n,i})} h_{n,i,1}\big(Z^*_{n,i,s}(\omega^*_{n,i}), s\big)\,\mathrm{d}\lambda(s) = \infty \tag{5.2}$$

for $Q^*_{n,i}$-almost all $\omega^*_{n,i}$ then h_n is a μ_n-density of a probability measure on $\mathcal{F}_n$.

Proof. According to Proposition 5.1.3, the mappings $h_{n,i}$ and h_n are measurable. The local boundedness guarantees that for $Q^*_{n,i}$-almost all $\omega^*_{n,i}$ it holds that $\tau^*_{n,i}(\omega^*_{n,i}) > 0$, for the measurability see Lemma 5.1.5. Consequently, the condition (5.2) and Proposition B.5.2.b guarantee that $h_{n,i,1}\big(Z^*_{n,i,\cdot}(\omega^*_{n,i}), \cdot\big)$ is a hazard rate of some probability measure on $\mathbb{B}(0,\infty)$ for $Q^*_{n,i}$-almost all $\omega^*_{n,i}$. Using Remark 5.1.4 gives that $h_{n,i}\big(Z^*_{n,i}(\omega^*_{n,i}), \cdot, \cdot\big)$ is a $\nu_C \otimes \lambda$-density of some probability measure for $Q^*_{n,i}$-almost all $\omega^*_{n,i}$. Therefore, $h_{n,i}$ is a $(Q^*_{n,i} \otimes \nu_C \otimes \lambda)$-density of some probability measure on $\mathcal{F}^*_{n,i} \otimes \mathcal{N}^*_{n,i}$. Bauer [6, Satz 23.11] gives the assertion. □

The next result gives the dual predictable projections of the counting processes.

5.1.7 Proposition (Counting Processes and Dual Predictable Projection). In the situation of Proposition 5.1.6 and considering the probability space $(\Omega_n, \mathcal{F}_n, P_n)$, where $P_n(F) = \int_F h_n \, \mathrm{d}\mu_n$, $F \in \mathcal{F}_n$, the processes

$$\big\{A_n^{(i)}(t) \mid t \in \mathbb{R}_+\big\} \quad \text{and} \quad \big\{\widetilde{A}_n^{(i)}(t) \mid t \in \mathbb{R}_+\big\}$$

are predictable with respect to $\mathbb{G}_n$ and $\mathbb{H}_n$, where we assume that $\mathcal{S}_n = \{\Omega_n, \emptyset\}$, see Assumption 5.1.1.ix. Moreover, the processes

$$\big\{N_n^{(i)}(t) - A_n^{(i)}(t) \mid t \in \mathbb{R}_+\big\} \quad \text{and} \quad \big\{\widetilde{N}_n^{(i)}(t) - \widetilde{A}_n^{(i)}(t) \mid t \in \mathbb{R}_+\big\}$$

are $\mathbb{G}_n$ and $\mathbb{H}_n$ martingales, where we assume that $\mathcal{S}_n = \{\Omega_n, \emptyset\}$, again.

Proof. One readily checks that the process

$$\big\{Y'_{n,i,s}(\cdot)\, R_n\big(Z^*_{n,i,s}(\cdot), s\big)\, \alpha_0(s) \mid s \in \mathbb{R}_+\big\}$$

is progressively measurable with respect to $\{\mathcal{F}^*_{n,i,s} \otimes \mathcal{N}'_{n,i,s} \mid s \in \mathbb{R}_+\}$. Proposition B.5.2.c and Proposition B.5.4.b yield that $A_n^{(i)} = \big\{A_n^{(i)}(t) \mid t \in \mathbb{R}_+\big\}$ is progressively measurable with respect to $\mathbb{H}_n$ and $\mathbb{G}_n$, where we use Proposition 5.1.2.c. Consequently, the processes $A_{n,k}^{(i)} = \big\{A_n^{(i)}(t) \wedge k \mid t \in \mathbb{R}_+\big\}$ are progressively measurable with respect to $\mathbb{H}_n$ and $\mathbb{G}_n$ As the process $A_{n,k}^{(i)}$ is real-valued with continuous paths, it follows that the process $A_{n,k}^{(i)}$ is predictable with respect to $\mathbb{H}_n$ and $\mathbb{G}_n$. Finally, we receive that $A_n^{(i)}(t, \omega_n) = \sup_{k \in \mathbb{N}} A_{n,k}^{(i)}(t, \omega_n)$ for all $\omega_n \in \Omega_n$ and $t \in \mathbb{R}_+$. Thus, we can conclude that the process $A_n^{(i)}$ is predictable with respect to $\mathbb{H}_n$ and $\mathbb{G}_n$. Additionally, Lemma 5.1.5 implies that $F_{n,i} = \big\{\tau^*_{n,i}(\omega^*_{n,i}) > u_{n,i}\big\} \in \mathcal{F}^*_{n,i} \otimes \mathcal{N}'_{n,i}$, where $\tau^*_{n,i}$ is defined in equation (5.1). By construction it holds that $\int_{F_{n,i}} h_{n,i} \, \mathrm{d}\mu_{n,i} = 1$, therefore the process $A_n^{(i)}$ is almost surely finite. Proving that the process $\big\{\widetilde{A}_n^{(i)}(t) \mid t \in \mathbb{R}_+\big\}$ is predictable is done by the same means.

Because of the product structure it suffices to show the second assertion for $i = n$. For $H \in \mathcal{H}_{n,t}$, we define $H_{\widetilde{\omega}} = \big\{\omega_{n,n} \mid (\widetilde{\omega}, \omega_{n,n}) \in H\big\}$, $\widetilde{\omega} \in \widetilde{\Omega} = \bigtimes_{i=1}^{n-1} \Omega_{n,i}$, and $H_{\widetilde{\omega},\omega^*_{n,n}} = \{\omega'_{n,n} \mid (\omega^*_{n,n}, \omega'_{n,n}) \in H_{\widetilde{\omega}}\}$. As $H \in \mathcal{H}_{n,t}$ implies $H \in \mathcal{H}^0_{n,s}$, $s > t$, it holds that $H_{\widetilde{\omega}} \in \mathcal{F}^*_{n,n,s} \otimes \mathcal{N}'_{n,n,s}$, $s > t$, and $H_{\widetilde{\omega},\omega^*_{n,n}} \in \mathcal{N}'_{n,n,s}$, $s > t$. As the filtration $\{\mathcal{N}'_{n,n,s} \mid s \in \mathbb{R}_+\}$ is right continuous, see Proposition B.5.1, it follows that $H_{\widetilde{\omega},\omega^*_{n,n}} \in \mathcal{N}'_{n,n,t}$.

Note that for $Q^*_{n,n}$-almost all $\omega^*_{n,n}$ and all $s, t \geq 0$, it holds that

$$\mathbb{E}\big[N'_{n,n,t+s} - A'_{n,n,t+s}(\omega^*_{n,n}, \cdot) \mid \mathcal{N}'_{n,n,t}\big] = N'_{n,n,t} - A'_{n,n,t}(\omega^*_{n,n}, \cdot), \tag{5.3}$$

almost surely with respect to the probability measure given by the $\nu_C \otimes \lambda$-density $h_{n,n}(\omega^*_{n,n}, \cdot)$, cf. Fleming and Harrington [19, Theorem 1.3.1]. Using Fubini's Theorem, cf. Bauer [6, Korollar 23.7], and equation (5.3) yields

$$\int_H \big(N_n^{(n)}(t+s) - A_n^{(n)}(t+s)\big)\, h_n \, \mathrm{d}\mu_n = \int_H \big(N_n^{(n)}(t) - A_n^{(n)}(t)\big)\, h_n \, \mathrm{d}\mu_n,$$

$H \in \mathcal{H}_{n,t}$, $s, t \geq 0$, *i.e.* the assertion. By exactly the same arguments, one proves the assertion for $\mathbb{G}_n$ and $\big\{\widetilde{N}_n^{(i)}(t) - \widetilde{A}_n^{(i)}(t) \mid t \in \mathbb{R}_+\big\}$. □

Now, we can state conditions that imply the existence of the MCRM and (localized) q-dimensional parametric sub-models.

5.1.8 Proposition. a) Let us suppose that the processes $\{Z^*_{n,i,t} \mid t \in \mathbb{R}_+\}$, $i = 1, \dots, n$, are predictable, locally bounded and that for every $\omega^*_{n,i} \in \Omega^*_{n,i}$ there exists a $C(\omega^*_{n,i}) \in \mathbb{R}_+$, such that $\sup\big\{\big|Z^*_{n,i,t}(\omega^*_{n,i})\big| \mid t \in \mathbb{R}_+\big\} \leq C(\omega^*_{n,i})$. Moreover, assume that the function $R_n : \mathbb{R}^{p+1} \longrightarrow \mathbb{R}$ is given by

$$R_n\big(z, s\big) = R_{n,\xi}\big(z, s\big) = \exp\bigg(\frac{1}{\sqrt{n}}\,\beta^{\mathrm{T}} z \circledcirc \gamma(s) + \frac{1}{\sqrt{n}}\,\eta^{\mathrm{T}}\widetilde{\gamma}(s)\bigg), \tag{5.4}$$

$\xi = (\beta^{\mathrm{T}}, \eta^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{r+q}$, see Definition 1.3.4, and that the functions $\gamma^{(\dot{u},\ddot{u})}$, $u = 1, \dots, r$ and $\widetilde{\gamma}^{(u)}$, $u = 1, \dots, q$, are bounded. Then the condition (5.2) holds and $\tau^*_{n,i} = \tau_{n,i}$ $Q^*_{n,i}$-almost surely.

b) Assume that the processes $\{Z^*_{n,i,t} \mid t \in \mathbb{R}_+\}$, $i = 1, \dots, n$, are predictable and locally bounded. Moreover, suppose that $\tau_0 = \infty$ and $\tau_{n,i} = \infty$, $i = 1, \dots, n$, as well as that $\gamma^{(\dot{u},\ddot{u})}$, $u = 1, \dots, r$, and $\widetilde{\gamma}^{(u)}$, $u = 1, \dots, r$, are bounded on every interval $[0, t]$, $t \in \mathbb{R}_+$. Analog to Proposition 5.1.8.a let $R_n : \mathbb{R}^{p+1} \longrightarrow \mathbb{R}$ be given by (5.4). Then the condition (5.2) holds.

Proof. Straightforward. □

Of course the conditions stated in Proposition 5.1.8 are not the most general conditions one can find to guarantee the existence of parametric sub-models. The cricial point is ensuring that $h_{n,i,1}$ is a hazard rate of a probability measure on $\mathcal{N}'_{n,i}$. On closer inspection this requirement is not really difficult, because of the right censoring. However, we also intend to get that $P_{n,\xi} \ll P_{n,0}$. This point turns out to be responsible for most of the conditions in Proposition 5.1.8.

5.1.9 Remark. Let $P_{n,0}$ be some probability measure on $\mathcal{F}_n$ and $\mathcal{Z}_{n,0}$ the σ-algebra generated by all subsets of $P_{n,0}$ negligible sets. It is well known that $P_{n,0}$ can be uniquely extended to a probability measure $P^c_{n,0}$ on $\mathcal{F}^{c,0}_n = \mathcal{Z}_{n,0} \vee \mathcal{F}_n$, such that $P^c_{n,0}(F) = P_{n,0}(F)$, $F \in \mathcal{F}_n$, see Dellacherie and Meyer [16, Theorem II.31, Remark II.32].

The next result is essential for proving that Assumption 2.2.1 holds.

5.1.10 Lemma. Consider the probability space $\big(\Omega_n, \mathcal{F}_n, \{P_{n,0}, P_{n,1}\}\big)$ and assume that $P_{n,1} \ll P_{n,0}$.

a) The Probability measure $P_{n,1}$ can be uniquely extended to a measure $P^{c,0}_{n,1}$ on $\mathcal{F}^{c,0}_n$, such that $P^{c,0}_{n,1}\big(F\big) = P_{n,1}(F)$, $F \in \mathcal{F}_n$. Moreover, it holds that $P^{c,0}_{n,1} \ll P^c_{n,0}$.

b) Assume that under $P_{n,1}$ the process $M_n = \{M_n(t) \mid t \in \mathbb{R}_+\}$ is a $\mathbb{H}_n$-martingale with $\mathcal{S}_n = \{\Omega_n, \emptyset\}$. Then under $P^{c,0}_{n,1}$ the process M_n is a $\mathbb{H}_n$-martingale with $\mathcal{S}_n = \mathcal{Z}_n$.

Proof. Assume that f is a $P_{n,0}$ density of $P_{n,1}$. Set

$$P^{c,0}_{n,1}(F) = \int_F f \,\mathrm{d}P^c_{n,0}, \qquad F \in \mathcal{F}^{c,0}_n.$$

Remark 5.1.9 implies the first assertion. Proof of b). Any set $H \in \mathcal{H}_{n,t}$, can be represented as $H = (G \cup N_c)\backslash(G \cap N_c)$, where $G \in \bigotimes_{i=1}^n (F^*_{n,i} \otimes \mathcal{N}'_{n,i,t})$

and $N_c \subset N \in \mathcal{F}_n$ with $P_{n,0}(N) = 0$. Consequently, $G \backslash N \subset H \subset G \cup N$, cf. Dellacherie and Meyer [16, Remark II.32]. Thus, it holds that

$$\int_H M_n(t+s)\,\mathrm{d}P^{c,0}_{n,1} = \int_G M_n(t+s)\,\mathrm{d}P_{n,1} = \int_G M_n(t)\,\mathrm{d}P_{n,1} = \int_H M_n(t)\,\mathrm{d}P^{c,0}_{n,1},$$

for all $s \geq 0$. □

5.1.11 Discussion. Under Assumption 5.1.1 and in the situation of Proposition 5.1.8, we proved the existence of a probability space $\big(\Omega_n, \mathcal{F}_n, \mathfrak{P}_n\big)$, $\mathfrak{P}_n = \{P_{n,\xi} \mid \xi \in \mathbb{R}^{r+q}\}$, such that

$$P_{n,\xi}(F) = \int_F h_{n,\xi}(\omega_n)\,\mathrm{d}\mu_n(\omega_n), \quad F \in \mathcal{F}_n, \qquad h_{n,\xi} = \prod_{i=1}^n h_{n,i,\xi},$$

where

$$h_{n,i,\xi} = \exp\bigg(-\int_{[0,\cdot]} h_{n,i,\xi,1}\big(Z_{n,i}(s,\omega_n), \varpi'_{n,i}\big)(\omega_n)\,\mathrm{d}\lambda(s)\bigg) \\ \times h_{n,i,\xi,2}\big(Z_{n,i}(s,\omega_n), \varpi'_{n,i}(\omega_n)\big),$$

with

$$h_{n,i,\xi,1}(z,\delta,s) = \big(R_{n,\xi}(z,s)\,\alpha_0(s) + \widetilde{\alpha}_{n,i}(s)\big) \cdot \mathbb{1}(s < \tau_{n,i})$$

and

$$h_{n,i,\xi,2}(z,\delta,s) = \big(R_{n,\xi}(z,s)\,\alpha_0(s) \cdot \mathbb{1}_{\{1\}}(\delta) + \widetilde{\alpha}_{n,i}(s) \cdot \mathbb{1}_{\{0\}}(\delta)\big) \cdot \mathbb{1}(s < \tau_{n,i}),$$

cf. Assumption 5.1.1.iii, Assumption 5.1.1.iv and equation (5.4). As $h_{n,0} = 0$ implies that $h_{n,\xi} = 0$ for all $\xi \in \mathbb{R}^{r+q}$, it results that $P_{n,\xi} \ll P_{n,0}$ for all $\xi \in \mathbb{R}^{r+q}$. Using Lemma 5.1.10.a and setting $\mathcal{S}_n = \mathcal{Z}_{n,0}$, we can assume that $\mathcal{F}_n$ and $\mathcal{G}_{n,t}$, $t \in \mathbb{R}_+$, are $P_{n,0}$-complete. This means that Assumption 2.2.1.iii and Assumption 2.2.1.iv hold. Proposition 5.1.2 yields that $\mathbb{G}_n$ is indeed a filtration and that Assumption 2.2.1.i and Assumption 2.2.1.vii hold. Since

$$P_{n,0}\big(N^{(i)}_n(s) - N^{(i)}_n(s-) = 1\big) = 0 \quad \text{and} \quad P_{n,0}\big(\widetilde{N}^{(i)}_n(s) - N^{(i)}_n(s-) = 1\big) = 0,$$

$i = 1, \dots, n$, $s \geq 0$, we can modify the process

$$\big(N_n^{(1)}, \dots, N_n^{(n)}, \widetilde{N}_n^{(1)}, \dots, \widetilde{N}_n^{(n)}\big)^{\mathrm{T}}$$

on a $P_{n,0}$ negligible set, such it is a multivariate counting process. Note that this procedure has no impact on the filtration, as it is $P_{n,0}$-complete. Thus, Assumption 2.2.1.ii, Assumption 2.2.1.v and Assumption 2.2.1.vi are valid by construction. Proposition 5.1.7, Lemma 5.1.10, Proposition B.5.3 and Proposition B.5.4 yield Assumption 2.2.1.viii and Assumption 2.2.1.ix

5.1.12 Remark. Proposition 5.1.6 and Proposition 5.1.7 emphasize that Assumption 2.2.1.vii is mainly due to the fact that we have to guarantee that all local martingales have representation property with respect to the counting process $(N_n^{\mathrm{T}}, \widetilde{N}_n^{\mathrm{T}})$. Basically, Assumption 2.2.1.vii secures that the distribution of the covariates does not change with the parameter $\xi \in \mathbb{R}^{r+q}$.

5.2 Checking Further Conditions

In this section we always suppose that the following premises, for which we gave sufficient condition in Discussion 5.1.11, hold

5.2.1 Assumption. Suppose that Assumption 5.1.1, Assumption 2.2.1, Assumption 2.3.9.i, Assumption 2.3.9.ii and Assumption 2.3.9.viii are satisfied. Moreover, suppose that

$$P_{n,0}(F) = \int_F h_n \, \mathrm{d}\mu_n, \qquad F \in \mathfrak{F}_n,$$

where we assume that $R_n = 1$, see also Assumption 5.1.1.iii and Assumption 5.1.1.iv.

In the following paragraphs we consider the remaining articles of Assumption 2.3.9. Note that conditions similar to Assumption 2.3.9 are also used by other author, cf. Andersen *et al.* [4, Condition VII.2.1] or Andersen and Gill [5]. Now we state a result derived in empirical process theory that helps us to verify Assumption 2.3.9.iv – Assumption 2.3.9.vi

5.2.2 Proposition (Abstract Law of Large Numbers). Under Assumption 5.2.1, let us suppose that every path of the covariate processes

$$\{Z_{n,i}(t) \mid t \in \mathbb{R}_+\}, \qquad i = 1, \dots, n, \ n \in \mathbb{N},$$

is left continuous. Set $\widetilde{Z}_{n,i} = \sup_{t \in I(\tau)} \big|\big|Z_{n,i}(t)\big|\big|_\infty$, where $||\cdot||_\infty$ denotes the sup-norm on $\mathbb{R}^p$, see Definition B.4.3, and define the covering number

$$\mathrm{CN}(\varepsilon, \rho) = \min\Big\{|\mathcal{T}| \;\Big|\; \mathcal{T} \subset \mathbb{Q}_\tau, \ \inf_{s \in \mathcal{T}} \rho(t,s) \le \varepsilon \text{ for all } t \in \mathbb{Q}_\tau\Big\},$$

where $\mathbb{Q}_\tau = I(\tau) \cap \mathbb{Q}$ and ρ is a pseudo metric on $I(\tau)$, as well as the empirical pseudo metrics

$$\rho_{n,1}^{(u)}(s,t) = \frac{1}{n} \sum_{i=1}^{n} \Big| Z_{n,i}^{(u)}(s) Y_n^{(i)}(s) - Z_{n,i}^{(u)}(t) Y_n^{(i)}(t) \Big|, \quad s,t \in \mathbb{Q}_\tau,$$

and

$$\rho_{n,2}^{(u,v)}(s,t) = \frac{1}{n} \sum_{i=1}^{n} \Big| Z_{n,i}^{(u)}(s)\, Z_{n,i}^{(v)}(s)\, Y_n^{(i)}(s) - Z_{n,i}^{(u)}(t)\, Z_{n,i}^{(v)}(t)\, Y_n^{(i)}(t) \Big|,$$

$s,t \in \mathbb{Q}_\tau$. Assume that

$$\lim_{C \to \infty} \sup_{n \in \mathbb{N}} \sup_{1 \le i \le n} \int \mathbb{1}\big(\widetilde{Z}_{n,i}^2 > C\big)\, \widetilde{Z}_{n,i}^2 \,\mathrm{d}P_{n,0} = 0$$

and

$$\frac{1}{n} \log \mathrm{CN}(\varepsilon, \rho_{n,1}^{(u)}) \longrightarrow_{P_{n,0}} 0, \quad \text{and} \quad \frac{1}{n} \log \mathrm{CN}(\varepsilon, \rho_{n,2}^{(u,v)}) \longrightarrow_{P_{n,0}} 0, \tag{5.5}$$

as $n \to \infty$, for all $\varepsilon > 0$, $u,v = 1, \dots, n$. Then it holds that

a) $\mathbb{E}_{n,0}\Big(\sup_{t \in I(\tau)} \big|\widehat{\mu}_{n,0}(t) - \mathbb{E}_{n,0}\widehat{\mu}_{n,0}(t)\big|\Big) \to 0$, as $n \to \infty$.

b) $\mathbb{E}_{n,0}\Big(\sup_{t \in I(\tau)} \big|\widehat{\mu}_{n,1}^{(u)}(t) - \mathbb{E}_{n,0}\widehat{\mu}_{n,1}^{(u)}(t)\big|\Big) \to 0$, as $n \to \infty$, $u = 1, \dots, p$.

c) $\mathbb{E}_{n,0}\Big(\sup_{t \in I(\tau)} \big|\widehat{\mu}_{n,2}^{(u,v)}(t) - \mathbb{E}_{n,0}\widehat{\mu}_{n,2}^{(u,v)}(t)\big|\Big) \to 0$, as $n \to \infty$, $u,v = 1, \dots, p$.

Proof. As the paths of the processes $\{Y_n^{(i)}(t) \mid t \in \mathbb{R}_+\}$, $i = 1, \dots, n$, $n \in \mathbb{N}$, are left continuous, it holds that

$$\sup_{t \in I(\tau)} \Big| \widehat{\mu}_{n,0}(t) - \mathbb{E}_{n,0}\big(\widehat{\mu}_{n,0}(t)\big) \Big| = \sup_{t \in \mathbb{Q}_\tau} \Big| \widehat{\mu}_{n,0}(t) - \mathbb{E}_{n,0}\big(\widehat{\mu}_{n,0}(t)\big) \Big|,$$

see Proposition B.5.5. Obviously, it also holds that

$$\lim_{C\to\infty} \sup_{n\in\mathbb{N}} \sup_{1\le i\le n} \int \sup_{t\in\mathbb{Q}_\tau} \big|Y_n^{(i)}(t)\big| \mathbb{1}\Big(\sup_{t\in\mathbb{Q}_\tau} \big|Y_n^{(i)}(t)\big| > C\Big) \,\mathrm{d}P_{n,0} = 0$$

and $\mathrm{CN}(\varepsilon, \rho_{n,0}) \le (n+1)$, where

$$\rho_{n,0}(s,t) = \frac{1}{n}\sum_{i=1}^{n} \big|Y_n^{(i)}(s) - Y_n^{(i)}(t)\big|, \quad s,t\in\mathbb{Q}_\tau.$$

The latter assertion can be seen as follows. We have that

$$\Big\{\sup\big\{s\in\mathbb{R}_+ \mid Y_n^{(i)}(s) = 1\big\},\, i=1,\dots,n\Big\} = \{s_{n,1},\dots,s_{n,k_n}\},$$

where $s_{n,i-1} < s_{n,i}$ and $k_n \le n$. Choose rational numbers $t_{n.i}$, such that $s_{n,i-1} < t_{n,i} < s_{n,i}$, $i = 1,\dots,k_n+1$, where $s_{n,0} = 0$ and $s_{n,k_n+1} = \infty$. Setting $\mathcal{T} = \{t_{n,i} \mid i = 1,\dots,k_n+1\}$ we get that

$$\inf_{t\in\mathcal{T}} \rho_{n,0}(t,s) = 0, \quad \text{for all } s\in\mathbb{R}_+.$$

Thus, $n^{-1}\log \mathrm{CN}(\varepsilon, \rho_{n,0}) \longrightarrow_{P_{n,0}} 0$, as $n\to\infty$. Now the assertions are implied by Dümbgen [17, Satz 8.3]. The other assertions are proved completely analogously. Note that $\widetilde{Z}_{n,i}$, $i=1,\dots,n$, $n\in\mathbb{N}$, are measurable, because of Proposition B.5.5.a. $\square$

5.2.3 Corollary. In the situation of Proposition 5.2.2, assume that the following conditions hold.

i) $\sup_{t\in I(\tau)}\big|\frac{1}{n}\sum_{i=1}^n \mathbb{E}_{n,0}\big(Y_n^{(i)}(t)\big) - \mu_0(t)\big| \to 0$, as $n\to\infty$,

ii) $\sup_{t\in I(\tau)}\big|\frac{1}{n}\sum_{i=1}^n \mathbb{E}_{n,0}\big(Y_n^{(i)}(t)\big)\,\mathbb{E}_{n,0}\big(Z_{n,i}^{(u)}(t)\big) - \mu_1^{(u)}(t)\big| \to 0$, as $n\to\infty$, $u=1,\dots,p$,

iii) $\sup_{t\in I(\tau)}\big|\frac{1}{n}\sum_{i=1}^n \mathbb{E}_{n,0}\big(Y_n^{(i)}(t)\big)\,\mathbb{E}_{n,0}\big(Z_{n,i}^{(u)}(t)\,Z_{n,i}^{(v)}(t)\big) - \mu_2^{(u,v)}(t)\big| \to 0$, as $n\to\infty$, $u,v=1,\dots,p$,

where μ_0, $\mu_1^{(u)}$ and $\mu_2^{(u,v)}$, $u,v=1,\dots,p$, are bounded and left continuous. Then Assumption 2.3.9.iv, Assumption 2.3.9.v and Assumption 2.3.9.vi hold for all $t\le\tau$.

Proof. Proposition B.5.5.a gives that $\sup_{t\in I(\tau)}\big|\widehat{\mu}_{n,0}(t)-\mu_0(t)\big|$ is measurable. The Markov-inequality, cf. Gänssler and Stute [20, Lemma 1.18.1], implies that $\sup_{t\in I(\tau)}\big|\widehat{\mu}_{n,0}(t)-\mathbb{E}_{n,0}\,\widehat{\mu}_{n,0}(t)\big| \longrightarrow_{P_{n,0}} 0$, as $n\to\infty$. Therefore, the first assertion is an immediate consequence of the triangle inequality. Noting that

$$\mathbb{E}_{n,0}\big(Y_n^{(i)}(t)\big)\,\mathbb{E}_{n,0}\big(Z_{n,i}^{(u)}(t)\big)=\mathbb{E}_{n,0}\big(Y_n^{(i)}(t)\,Z_{n,i}^{(u)}(t)\big)$$

and

$$\mathbb{E}_{n,0}\big(Y_n^{(i)}(t)\big)\,\mathbb{E}_{n,0}\big(Z_{n,i}^{(u)}(t)\,Z_{n,i}^{(v)}(t)\big)=\mathbb{E}_{n,0}\big(Y_n^{(i)}(t)\,Z_{n,i}^{(u)}(t)\,Z_{n,i}^{(v)}(t)\big),$$

one proves the second and third assertion analogously to the first one. □

5.2.4 Example (Proposition 5.2.2). a) Let us assume that the paths of the processes $\big\{Z_{n,i}(t)\mid t\in\mathbb{R}_+\big\}$, $i=1,\dots,n$, $n\in\mathbb{N}$, are left continuous with right hand limits (caglad) and that

$$\big|Z_{n,i}^{(u)}(s)-Z_{n,i}^{(u)}(t)\big|\le L(\tau)|t-s| \tag{5.6}$$

and

$$\big|Z_{n,i}^{(u)}(s)\,Z_{n,i}^{(v)}(s)-Z_{n,i}^{(u)}(t)\,Z_{n,i}^{(v)}(t)\big|\le L(\tau)|t-s|, \tag{5.7}$$

whenever the process $Z_{n,i}^{(u)}$ and $Z_{n,i}^{(u)}$ do not jump in the interval $[s,t]$, where $L(\tau)\in\mathbb{R}_+$ is independent of u, v, i and n. This means the paths of the processes are piecewise Lipschitz continuous functions. Set $J_n=\max\{J_n^{(u,v)}\mid u,v=1,\dots,p\}$, where $J_n^{(u,v)}$ is the number of jumps of the process $\big\{\widehat{\mu}_{n,2}^{(u,v)}(t)\mid t\in\mathbb{R}_+\big\}$ in the interval $[0,\tau]$. Assume that for a fixed $\omega_n\in\Omega_n$ the paths of the processes $\big\{Z_{n,i}^{(u)}(t\wedge\tau)\,Y_n^{(i)}(t\wedge\tau)\mid t\in\mathbb{R}_+\big\}$, $u=1,\dots,p$, $i=1,\dots,n$, do not jump in the interval $[s_1,s_2]\subset I(\tau)$. Then it holds that $\rho_{n,1}^{(u)}(s_1,s_2)\le L(\tau)\,|s_2-s_1|$ and $\rho_{n,2}^{(u,v)}(s_1,s_2)\le L(\tau)\,|s_2-s_1|$, $u,v=1,\dots,p$. Taking the jumps into account one easily checks the following estimates for the covering numbers

$$\mathrm{CN}(\varepsilon,\rho_{n,1}^{(u)})\le\frac{2\tau L(\tau)}{\varepsilon}+J_n+1\quad\text{and}\quad\mathrm{CN}(\varepsilon,\rho_{n,2}^{(u,v)})\le\frac{2\tau L(\tau)}{\varepsilon}+J_n+1.$$

Without loss of generality we can assume that $2\tau L(\tau)\varepsilon^{-1} \geq 2$ and that $J_n + 1 \geq 2$. Hence, it holds the estimate

$$\log\left(\frac{2\tau L(\tau)}{\varepsilon} + J_n + 1\right) \leq \log(2\tau L(\tau)) - \log(\varepsilon) + \log(J_n + 1).$$

If $n^{-1}\log(J_n + 1) \longrightarrow_{P_{n,0}} 0$ then the condition (5.5) holds.

b) Examples of caglad processes that satisfy the conditions (5.6) and (5.7) are processes with piecewise constant paths.

c) If the paths of the covariate processes do not contain any jumps and conditions (5.6) and (5.7) hold, then $J_n \leq n + 1$. Thus, the condition (5.5) holds.

5.2.5 Example (k**-sample problems).** In the situation of Proposition 5.2.2 assume that $n > k$ and that $n_i = n_i(n)$ $i = 0, \dots, k$, are sequences of natural numbers, such that $n_0 = 0$ and $n_k = n$ and that under $P_{n,0}$ the random variables

$$Z_{n,i}(t) \sim \bar{Z}_l(t) \quad \text{and} \quad Y_n^{(i)}(t) \sim \bar{Y}^{(l)}(t), \qquad n_{l-1} < i \leq n_l$$

$t \in \mathbb{R}_+$, $l = 1, \dots, k$. Suppose that $\sup_{1 \leq l \leq k} \sup_{t \in \mathbb{R}_+} \mathbb{E}_{n,0} \big|\big|\bar{Z}_l(t)\big|\big|_\infty^2 < \infty$. If $\big(n_l - n_{l-1}\big)/n \to \nu_l$, as $n \to \infty$, for $l = 1, \dots, k$, then the conditions of Corollary 5.2.3 hold with $\mu_0(t) = \sum_{l=1}^k \nu_l \mathbb{E}\big(Y^{(l)}(t)\big)$, as well as

$$\mu_1^{(u)}(t) = \sum_{l=1}^{k} \nu_l \mathbb{E}\big(Y^{(l)}(t)\big)\, \mathbb{E}\big(Z_l^{(u)}(t)\big)$$

and

$$\mu_2^{(u,v)}(t) = \sum_{l=1}^{k} \nu_l \mathbb{E}\big(Y^{(l)}(t)\big)\mathbb{E}\big(Z_l^{(u)}(t)\, Z_l^{(v)}(t)\big),$$

$u = 1, \dots, p$. The left continuity of this functions can be proved analogously to Proposition B.5.5.

5.2.6 Proposition. Suppose that Assumption 2.3.9.iv – Assumption 2.3.9.vi and

$$\sup_{n \in \mathbb{N}} \sup_{1 \leq i \leq n} \sup_{t \in I(\tau)} \mathbb{E}_{n,0} \big|\big|Z_{n,i}(t)\big|\big|_\infty^2 \leq K < \infty$$

hold, where $||\cdot||_\infty$ denotes the sup-norm on $\mathbb{R}^p$, see Definition B.4.3. Then the functions μ_0, $\mu_1^{(u)}$ and $\mu_2^{(u,v)}$, $u, v = 1, \dots, p$, are bounded on the interval $[0, \tau]$, *i.e.* Assumption 2.3.9.iii holds for all $t \le \tau$.

Proof. For μ_0 the assertion is straightforward. Let $M > 0$ be arbitrary. For all $t \in [0, \tau]$ it holds that

$$P_{n,0}\Big(\big|\mu_1^{(u)}(t)\big| \ge M\Big) \le P_{n,0}\left(\big|\widehat{\mu}_{n,1}^{(u)}(t) - \mu_1^{(u)}(t)\big| \ge \frac{M}{2}\right) + P_{n,0}\left(\big|\widehat{\mu}_{n,1}^{(u)}(t)\big| \ge \frac{M}{2}\right).$$

The Markov-inequality and the Jensen inequality, cf. Gänssler and Stute [20, Lemma 1.18.1, Satz 5.4.7], yield that

$$P_{n,0}\left(\big|\widehat{\mu}_{n,1}^{(u)}(t)\big| \ge \frac{M}{2}\right) \le \frac{2}{M} \cdot \mathbb{E}_{n,0}\big|\widehat{\mu}_{n,1}^{(u)}(t)\big| \le \frac{2\sqrt{K}}{M}.$$

Using Assumption 2.3.9.v, one receives

$$\limsup_{n\to\infty} P_{n,0}\Big(\big|\mu_1^{(u)}(t)\big| \ge M\Big) \le \frac{2\sqrt{K}}{M} < 1.$$

for all sufficiently large $M > 0$. Note that M is independent of t. As

$$P_{n,0}\Big(\big|\mu_1^{(u)}(t)\big| \ge M\Big) = \begin{cases} 1, & \text{if } \big|\mu_1^{(u)}(t)\big| \ge M, \\ 0, & \text{if } \big|\mu_1^{(u)}(t)\big| < M, \end{cases}$$

it results that $\big|\mu_1^{(u)}(t)\big| \le M$ for all $t \in I(\tau)$. The third assertion is proved completely analogously. □

5.2.7 Example (Assumption 3.2.11). Suppose that

$$\gamma^{(\dot{u},\ddot{u})} = \gamma_0^{(\dot{u},\ddot{u})} \circ H \quad \text{and} \quad \widehat{\gamma}_n^{(\dot{u},\ddot{u})} = \gamma_0^{(\dot{u},\ddot{u})} \circ \widehat{H}_n, \qquad n \in \mathbb{N},\ u = 1, \dots, r,$$

where $\gamma_0^{(\dot{u},\ddot{u})} : [0,1] \longrightarrow \mathbb{R}$ are some continuous functions. Let us discuss the following two cases.

a) Assume that $H = 1 - \mu_0$ and set $\widehat{H}_n = 1 - \widehat{\mu}_{n,0}$. Assumption 2.3.9.iv implies Assumption 3.2.11. This assertion can seen as follows. The processes $\{\widehat{\gamma}_n^{(\dot{u},\ddot{u})}(t) \mid t \in \mathbb{R}_+\}$, $u = 1, \dots, r$, $n \in \mathbb{N}$, are bounded and left continuous, *i.e.* they are especially locally bounded and predictable. In particular, the functions $\gamma_0^{(\dot{u},\ddot{u})}$, $u = 1, \dots, r$, are uniformly continuous. Therefore, we can find for every $\varepsilon > 0$ some $\delta > 0$, such that

$$P_{n,0}\bigg(\sup_{t \in I(\tau)} \big|\widehat{\gamma}_n^{(\dot{u},\ddot{u})} - \gamma^{(\dot{u},\ddot{u})}\big| \geq \varepsilon\bigg) \leq P_{n,0}\bigg(\sup_{t \in I(\tau)} \big|\mu_0(t) - \widehat{\mu}_{n,0}(t)\big| \geq \delta\bigg) \to 0,$$

as $n \to \infty$, for all $\tau < \tau_0$.

b) Suppose that Assumption 3.2.1 holds and that

$$1 - H(t) = \exp\bigg(-\int_{I(t)} \alpha_0(s)\,\mathrm{d}s\bigg), \qquad t \in \mathbb{R}_+.$$

Let $\widehat{H}_n$ denote a left continuous version of the Kaplan-Meyer estimator for H, see Andersen *et al.* [4, Section IV.3]. One readily checks that Assumption 2.3.9.iv implies the conditions of Andersen *et al.* [4, Theorem IV.3.1]. Therefore, we have that $\sup_{t \in I(\tau)}\big|\widehat{H}_n(t) - H(t)\big| \longrightarrow_{P_{n,0}} 0$, as $n \to \infty$, for all $\tau \leq \tau_0^c$. Analogously to Example 5.2.7.a, one shows that Assumption 3.2.11 holds.

Asymptotic normality restricted to time τ_0 always depended on some additional conditions. The next result gives sufficient, handier assumptions for these premises. We also show the existence of the canonical SHPSM restricted to time τ_0.

5.2.8 Proposition. Under Assumption 5.2.1, suppose that Assumption 2.3.9 and Assumption 3.2.1 hold as well as the conditions

$$\frac{1}{\sqrt{n}} \max_{1 \leq i \leq n} \sup_{s \in I(\tau_0^c)} \big|\big|Z_{n,i}(s)\big|\big|_\infty \longrightarrow_{P_{n,0}} 0, \quad \text{as } n \to \infty, \tag{5.8}$$

and

$$\sup_{n \in \mathbb{N}} \sup_{1 \leq i \leq n} \sup_{s \in I(\tau_0^c)} \mathbb{E}_{n,0}\Big(\big|\big|Z_{n,i}(s)\big|\big|_\infty^2\Big) = K < \infty, \tag{5.9}$$

where $||\cdot||_\infty$ denotes the sup-norm on $\mathbb{R}^p$, see Definition B.4.3.

a) The conditions (2.24) and (2.25) hold. In particular, the functions $\mu_1^{(u)}/\mu_0$, $u = 1, \ldots, p$, are bounded. This means that the conditions (2.24) and (2.25) also hold in the context of Theorem 3.2.3.

b) The conditions (3.11) and (3.12) are satisfied.

c) Assume that $|\widehat{\gamma}_n^{(\dot{u},\ddot{u})}| \le C < \infty$, $u = 1, \ldots, r$, $n \in \mathbb{N}$, then the condition (3.13) and (3.17) hold.

Proof. Using the notation of Theorem 2.3.10, we prove that (2.24) holds. Note that we can replace τ_0 by τ_0^c in equation (2.24) and (2.25), because of Assumption 3.2.1. Choose $\delta, \varepsilon > 0$ and define the sets

$$A_n = \left\{ \frac{1}{\sqrt{n}} \max_{1 \le i \le n} \sup_{s \in I(\tau_0)} \big|\big| Z_{n,i}(s) \big|\big|_\infty \le \delta \right\}, \qquad n \in \mathbb{N},$$

as well as

$$B_n = \left\{ \left| \sum_{i=1}^n \int_{(t,\tau_0)} V_n^{(i)}(s,\xi,\xi')\, \lambda_n^{(i)}(s)\,\mathrm{d}s \right| \ge \varepsilon \right\}, \qquad n \in \mathbb{N}.$$

Clearly, it holds that

$$\limsup_{n\to\infty} P_{n,0}(B_n) \le \limsup_{n\to\infty} P_{n,0}(A_n \cap B_n) + \limsup_{n\to\infty} P_{n,0}(A_n^{\complement})$$

and $\lim_{n\to\infty} P_{n,0}(A_n^{\complement}) = 0$, because of (5.8). A Taylor expansion gives that

$$\exp\left(\frac{1}{\sqrt{n}} \widetilde{\xi}^{\mathrm{T}} \Psi_{n,i}(s)\, Y_n^{(i)}(s) \right) - 1 = \frac{1}{\sqrt{n}} \widetilde{\xi}^{\mathrm{T}} \Psi_{n,i}(s) \exp\left(\frac{\widetilde{\vartheta}(s)}{\sqrt{n}}\, \xi^{\mathrm{T}} \Psi_{,n,i}(s) \right) Y_n^{(i)}(s),$$

where $\widetilde{\vartheta}(s) \in (0,1)$. Therefore, we get that

$$\begin{aligned} \big| V_n^{(i)}(s,\xi,\xi') \big| \le{}& \frac{1}{2n} \big| \xi^{\mathrm{T}} \Psi_{n,i}(s)\, \xi'^{\mathrm{T}} \Psi_{n,i}(s) \big| \\ & \times \exp\left(\frac{1}{2\sqrt{n}} \vartheta(s)\, \xi^{\mathrm{T}} \Psi_{n,i}(s) + \frac{1}{2\sqrt{n}} \vartheta'(s)\, \xi'^{\mathrm{T}} \Psi_{n,i}(s) \right) Y_n^{(i)}(s), \end{aligned}$$

where $\vartheta, \vartheta' : \mathbb{R}_+ \longrightarrow (0,1)$. Using the boundedness of the weight functions, on the set A_n it holds the estimate $\big| V_n^{(i)}(s,\xi,\xi') \big| \le C$, where C is some constant that can be chosen independently of i, n and s. Using Fubini's Theorem, cf. Bauer [6, Korollar 23.7], gives

$$\mathbb{E}_{n,0} \int_{(t,\tau_0^c)} C\, Y_n^{(i)}(s)\, \alpha_0(s)\, \mathrm{d}s = C \int_{(t,\tau_0^c)} \mathbb{E}_{n,0}\big(Y_n^{(i)}(s)\big)\, \alpha_0(s)\, \mathrm{d}s.$$

Using Remark 5.1.4 gives that $\mathbb{E}_{n,0}\big(Y_n^{(i)}(s)\big) \leq 1 - F_0(s)$, where

$$1 - F_0(s) = \exp\Big(- \int_{[0,t]} \alpha_0(u)\, \mathrm{d}u \Big).$$

Therefore, we get that

$$\mathbb{E}_{n,0} \int_{(t,\tau_0)} C\, Y_n^{(i)}(s)\, \alpha_0(s)\, \mathrm{d}s \leq C\big(F(\tau_0^c) - F_0(t)\big). \tag{5.10}$$

The Markov-inequality, cf. Gänssler and Stute [20, Lemma 1.18.1], and (5.10) give that

$$\begin{aligned} \limsup_{n\to\infty} P_{n,0}(A_n \cap B_n) &\leq \limsup_{n\to\infty} P_{n,0}\Big(\frac{1}{n} \sum\nolimits_{i=1}^{n} \int_{(t,\tau_0^c)} C \;\; \lambda_n^{(i)}(s)\, \mathrm{d}s \geq \varepsilon \Big) \\ &\leq \frac{C}{\varepsilon}\big(F_0(\tau_0^c) - F_0(t)\big) \to 0, \end{aligned}$$

as $t \to \tau_0^c$. With a similar consideration one proves that (2.25) holds. In the next step we show that the functions $\mu_1^{(u)}/\mu_0$, $u = 1, \ldots, p$, are bounded. For this proof we also use the same idea as in Proposition 5.2.6. If $\tau_0^c < \infty$ we can assume without loss of generality that $\mu_1^{(u)}(\tau_0^c)/\mu_0(\tau_0^c) = 0$. Let $M > 0$ be arbitrary. For all $t \in [0, \tau_0^c)$ it holds that

$$\begin{aligned} \limsup_{n\to\infty} P_{n,0}\left(\left| \frac{\mu_1^{(u)}(t)}{\mu_0(t)} \right| \geq M \right) &\leq \limsup_{n\to\infty} P_{n,0}\left(\left| \frac{\widehat{\mu}_{n,1}^{(u)}(t)}{\widehat{\mu}_{n,0}(t)} \right| \geq \frac{M}{2} \right) \\ &\quad + \limsup_{n\to\infty} P_{n,0}\left(\left| \frac{\mu_1^{(u)}(t)}{\mu_0(t)} - \frac{\widehat{\mu}_{n,1}^{(u)}(t)}{\widehat{\mu}_{n,0}(t)} \right| \geq \frac{M}{2} \right) \end{aligned}$$

Obviously, it holds that

$$\left| \frac{\widehat{\mu}_{n,1}^{(u)}(t)}{\widehat{\mu}_{n,0}(t)} \right| \leq \frac{\widehat{\mu}_{n,2}^{(u,u)}(t) \cdot \widehat{\mu}_{n,0}(t)}{\widehat{\mu}_{n,0}(t)} \leq \frac{1}{n} \sum_{i=1}^{n} \big|\big| Z_{n,i}(t) \big|\big|_\infty^2,$$

where we use the Cauchy-Schwarz inequality, see *e.g.* Gänssler and Stute [20, Satz 1.13.3]. The Markov inequality and Lemma 3.2.7, give that

$$\limsup_{n\to\infty} P_{n,0}\left(\left|\frac{\mu_1^{(u)}(t)}{\mu_0(t)}\right| \geq M\right) \leq \frac{2K}{M} < 1$$

for sufficiently large M. As K and M are independent of t and

$$P_{n,0}\left(\left|\frac{\mu_1^{(u)}(t)}{\mu_0(t)}\right| \geq M\right) = \begin{cases} 1, & \text{if } \left|\mu_1^{(u)}(t)/\mu_0(t)\right| \geq M, \\ 0, & \text{if } \left|\mu_1^{(u)}(t)/\mu_0(t)\right| < M, \end{cases}$$

it follows the assertion.

First, we show that (3.13) holds. Again, using the estimate $(\widehat{\mu}_{n,1}^{(\dot{u})})^2/\widehat{\mu}_{n,0} \leq \widehat{\mu}_{n,2}^{(\dot{u},\dot{u})}$, applying Markov's inequality and Fubini's Theorem yield

$$\begin{aligned}
P_{n,0}&\left(\int_{(t,\tau_0^c)} \big(\widehat{\gamma}_n^{(\dot{u},\ddot{u})}(s)\big)^2 \left(\widehat{\mu}_{n,2}^{(\dot{u},\dot{u})}(s) - \frac{(\widehat{\mu}_{n,1}^{(\dot{u})}(s))^2}{\widehat{\mu}_{n,0}(s)}\right)\alpha_0(s)\,\mathrm{d}s \geq \varepsilon\right) \\
&\leq \frac{2}{\varepsilon n}\sum_{i=1}^n \int_{(t,\tau_0^c)} \mathbb{E}_{n,0}\big(Z_{n,i}^{(\dot{u})}(s)\big)^2\, \mathbb{E}_{n,0}\Big(\big(\widehat{\gamma}_n^{(\dot{u},\ddot{u})}(s)\big)^2\, Y_n^{(i)}(s)\Big)\alpha_0(s)\,\mathrm{d}s \\
&\leq \frac{2KC^2}{\varepsilon}\int_{(t,\tau_0^c)} \mathbb{E}_{n,0}\big(Y_n^{(i)}(s)\big)\;\alpha_0(s)\,\mathrm{d}s \leq \frac{2KC^2}{\varepsilon}\cdot\big(F_0(\tau_0^c) - F_0(t)\big),
\end{aligned}$$

where $C > 0$ is some suitable constant. For the last estimates we also use boundedness of the weight functions and the stochastic independence of the covariates and the at-risk processes. $t \to \tau_0^c$ gives that (3.13) holds. With the same method one proves that the conditions (3.11) and (3.12) are satisfied. Using the estimates

$$\big|\widehat{\mu}_{n,2}^{(\dot{u},\dot{v})}(s)\big| \leq \frac{1}{n}\sum_{i=1}^n \big|\big|Z_{n,i}(s)\big|\big|_\infty^2 Y_n^{(i)}(s)$$

and

$$\begin{aligned}
\big|\widehat{\mu}_{n,1}^{(\dot{u})}(s)\,\widehat{\mu}_{n,1}^{(\dot{v})}(s)\big| &\leq \left(\frac{1}{n}\sum_{i=1}^n \big|\big|Z_{n,i}(s)\big|\big|_\infty Y_n^{(i)}(s)\right)^2 \\
&\leq \left(\frac{1}{n}\sum_{i=1}^n \big|\big|Z_{n,i}(s)\big|\big|_\infty^2 Y_n^{(i)}(s)\right)\cdot \widehat{\mu}_{n,0}(s),
\end{aligned}$$

where we applied the Cauchy-Schwarz inequality, one sees that the condition (3.17) is proved completely analogously to (3.13), cf. also Remark 3.2.14.b.

□

5.2.9 Remark. a) Abbreviating $\widetilde{Z}_{n,i} = \sup_{s\in I(\tau_0^c)} \big|\big|Z_{n,i}(s)\big|\big|_\infty$ the condition (5.8) is implied by

$$\lim_{C\to\infty} \sup_{n\in\mathbb{N}} \sup_{1\le i\le n} \int \mathbb{1}\big(\widetilde{Z}^2_{n,i} > C\big)\widetilde{Z}^2_{n,i}\,\mathrm{d}P_{n,0} = 0. \tag{5.11}$$

For $\delta > 0$ choose $C > 0$, such that $\sup_{1\le i\le n,\, n\in\mathbb{N}} \mathbb{E}_{n,0}\big(\mathbb{1}(\widetilde{Z}^2_{n,i} > C)\widetilde{Z}^2_{n,i}\big) \le \delta$, and define the sets $A_n = \big\{\max_{1\le i\le n} \widetilde{Z}^2_{n,i} > C\big\}$, $n \in \mathbb{N}$. It holds that

$$\begin{aligned} P_{n,0}\bigg(\max_{1\le i\le n} \widetilde{Z}_{n,i} \ge \sqrt{n}\varepsilon\bigg) &\le P_{n,0}\bigg(\Big\{\max_{1\le i\le n} \widetilde{Z}^2_{n,i} \ge n\varepsilon^2\Big\} \cap A_n^{\complement}\bigg) \\ &\quad + P_{n,0}\bigg(\max_{1\le i\le n} \widetilde{Z}^2_{n,i}\mathbb{1}(\widetilde{Z}^2_{n,i} > C) \ge n\varepsilon^2\bigg) = p_{n,1} + p_{n,2}. \end{aligned}$$

Obviously, it holds that $p_{n,1} \to 0$, as $n \to \infty$, and

$$p_{n,2} \le \frac{1}{n\varepsilon^2} \sum_{i=1}^{n} \mathbb{E}_{n,0}\big(\mathbb{1}(\widetilde{Z}^2_{n,i} > C)\widetilde{Z}^2_{n,i}\big) \le \frac{\delta}{\varepsilon^2},$$

since δ was arbitrary, $\delta \downarrow 0$ yields the assertion. Additionally, the condition (5.11) yields that the condition (5.9) holds.

b) The condition (5.9) does not generally imply the condition (5.8). This can be seen as follows. Using the notation of a) assume that $\widetilde{Z}_{n,i}$, $i = 1, \ldots, n$, $n \in \mathbb{N}$, are stochastically independent and

$$P_{n,0}\big(\widetilde{Z}_{n,i} = z\big) = \begin{cases} 1 - 1/i, & \text{if } z = 0, \\ 1/i, & \text{if } z = \sqrt{i}. \end{cases}$$

Clearly, it holds that $\mathbb{E}_{n,0}(\widetilde{Z}_{n,i}) = 1/\sqrt{i}$ and $\mathbb{E}_{n,0}(\widetilde{Z}^2_{n,i}) = 1$. However, some tedious computation, where one uses the properties of the gamma function, gives that

$$\lim_{n\to\infty} P_{n,0}\bigg(\frac{1}{\sqrt{n}} \max_{1\le i\le n} \widetilde{Z}_{n,i} \ge \varepsilon\bigg) = 1 - \min\{1, \varepsilon^2\}, \qquad \varepsilon > 0.$$

c) Assume that the random variables $\widetilde{Z}_{n,i}$, $i = 1, \ldots, n$, $n \in \mathbb{N}$, see Remark 5.2.9.a, have the same distribution and that $\mathbb{E}_{n,0}\big(\widetilde{Z}^2_{n,i}\big) < \infty$, then the condition (5.11) holds.

d) Using the notation of a) assume that for some $\delta > 0$, it holds that

$$\sup_{n\in\mathbb{N}} \sup_{1\le i\le n} \mathbb{E}_{n,0}\big(\widetilde{Z}^{2+\delta}_{n,i}\big) = K < \infty. \tag{5.12}$$

Using the Markov inequality yields that

$$\begin{aligned} &P_{n,0}\Big(\frac{1}{\sqrt{n}} \max_{1\le i\le n} \sup_{s\in I(\tau_0)} \big|\big|Z_{n,i}(s)\big|\big|_\infty \ge \varepsilon\Big) \\ &\qquad\qquad \le \varepsilon^{-(2+\delta)} \cdot n^{-(1+\delta/2)} \sum_{i=1}^{n} \mathbb{E}_{n,0}\Big(\widetilde{Z}^{2+\delta}_{n,i}\Big) \le \frac{K}{\varepsilon^{2+\delta}} n^{-\delta/2}, \end{aligned}$$

i.e. the condition (5.8) holds. Obviously, the condition (5.9) is also implied by the condition (5.12).

e) If all paths of all covariate processes are bounded by the same constant then clearly the condition (5.11) is satisfied.

f) In particular, the condition (5.11) or (5.12) imply Assumption 2.3.9.iii, see Proposition 5.2.6, and Assumption 2.3.9.vii.

5.2.10 Proposition (Local Boundedness of Covariate Processes). Assume that $\big|\big|Z_{n,i}(0)\big|\big|_\infty \le C_{n,i}$, where $C_{n,i} \in \mathbb{R}_+$, $i = 1, \ldots, n$, and that the condition (5.11) holds. Then the processes $\big\{Z_{n,i}(t \wedge \tau_0^c) \mid t \in \mathbb{R}_+\big\}$, $i = 1, \ldots, n$, are locally bounded with respect to the filtration $\mathbb{G}_n$

Proof. Set

$$\tau_{n,i,k} = \left\{ \begin{array}{ll} 0, & \text{if } \widetilde{Z}_{n,i} \wedge C_{n,i} > k, \\ \infty, & \text{otherwise.} \end{array} \right\}$$

Clearly, it holds that $\{\tau_{n,i,k} \le t\} \in \mathcal{G}_{n,0} \subset \mathcal{G}_{n,t}$ and $\big|\big|Z_{n,i}(t \wedge \tau_{n,i,k} \wedge \tau_0^c)\big|\big|_\infty \le \max(C_{n,i}, k)$ for all $t \in \mathbb{R}_+$. The condition (5.11) gives that $\widetilde{Z}_{n,i}$ is integrable. Thus $\widetilde{Z}_{n,i} < \infty$ $P_{n,0}$-almost surely. Thus, we get that $\tau_{n,i,k} \to \infty$ $P_{n,0}$-almost surely. □

Let us end this section with an important and prominent type of covariate.

5.2.11 Example (Time-Independent Covariates). If the covariates $Z'_{n,i}$, $i = 1, \dots, n$, $n \in \mathbb{N}$, are time independent, we can identify them with the stochastic processes

$$\left\{ Z_{n,i}(t) = Z'_{n,i} \cdot \mathbb{1}(t > 0) \mid t \in \mathbb{R}_+ \right\}, \quad i = 1, \dots, n, \ n \in \mathbb{N}.$$

Furthermore, assume that the condition (5.11) holds. According to Proposition 5.2.10 these processes are locally bounded. As these processes are left continuous, using Example 5.2.4.a we see that the conditions of Proposition 5.2.2 hold. Remark 5.2.9.c yields that the premises of Proposition 5.2.8 are satisfied. Sufficient conditions for the assumptions of Corollary 5.2.3 can be found in Example 5.2.5.

5.3 Applications

In the previous sections we basically discussed Assumption 4.3.3, except Assumption 4.3.3.ii – Assumption 4.3.3.v, and showed how to construct models that satisfy these premises. Assumption 4.3.3.ii – Assumption 4.3.3.v guarantee that our statistical model is reasonable, see Remark 3.2.2.b and Remark 4.3.4. Therefore, we suppose that Assumption 4.3.3 holds and that our observation are given by the tuples $\left(X_{n,i}, \Delta_{n,i}, Z_{n,i}\right)$, $i = 1, \dots, n$, where $X_{n,i}$ and $\Delta_{n,i}$ denote a censored survival time and the corresponding censoring indicator. $Z_{n,i}$ is the covariate process associated with $(X_{n,i}, \Delta_{n,i})$, $i = 1, \dots, n$, $n \in \mathbb{N}$. In particular, it holds that $N_n^{(i)}(t) = \mathbb{1}(X_{n,i} \le t) \cdot \Delta_{n,i}$ and $Y_n^{(i)}(t) = \mathbb{1}(X_{n,i} \ge t)$, $t \in \mathbb{R}_+$, $i = 1, \dots, n$, $n \in \mathbb{N}$. Further information on the modelling can be found in Chapter 1. In the next example it is shown, once again, how a nonparametric testing problem can be transformed into parametric testing problem with help of the Modified Cox Regression Model (MCRM) and sequences of hardest parametric sub-models (SHPSM). As this procedure is always the same, the other examples are only discussed on the level of SHPSM.

5.3.1 Example (One-Sided Tests). Let $Z = \{Z(t) \mid t \in \mathbb{R}_+\}$ be some multivariate, non-negative covariate process possibly having some impact on a survival time T. Let

$$\lambda(t \mid Z = z) = \lim_{h \to 0} \mathbb{P}\big(t \leq T < t + h \mid T \geq t, Z = z\big).$$

denote the conditional hazard rate of T given $Z = z$, where $z = \{z(t) \,|\, t \in \mathbb{R}_+\}$ is a fixed path of the covariate process. In the following it is assumed that z is not identically zero for every component, *i.e.* $z^{(u)}(t_u) > 0$ for some $t_u \in \mathbb{R}_+$, $u = 1, \ldots, p$. Moreover, we suppose that $\lambda(t \mid Z = z) = \lambda\big(t \mid Z(t) = z(t)\big)$, *i.e.* the conditional hazard rate of T at t depends only at the value of the covariate process at time t.

Now, we intended to test the hypothesis that the covariates have no influence on the survival times versus the alternative that the larger the values of the covariates the shorter the survival times. More precisely, we want to test

$$\mathcal{H}_1 : \; \lambda\big(t \mid Z(t) = z(t)\big) = \lambda\big(t \mid Z(t) = 0\big) \quad \text{for all } t \in \mathbb{R}_+$$

versus

$$\mathcal{K}_1 : \; \lambda\big(t \mid Z(t) = z(t)\big) > \lambda\big(t \mid Z(t) = \widetilde{z}(t)\big) \text{ for all } t \text{ with } z(t) \gneqq \widetilde{z}(t),$$

where $z(t) \gneqq \widetilde{z}(t)$ means that $z^{(u)}(t) \geq \widetilde{z}^{(u)}(t)$ for all u and at least one inequality is strict.

This testing problem can be modelled with the MCRM. Let us assume that under the probability measure $P_{\beta,\alpha}$ the hazard rate of the survival time is given by

$$\begin{aligned}\lambda_{\beta,\alpha}(t \mid z) &= \exp\big(\beta^{\mathrm{T}} z \circledcirc \gamma_\alpha(t)\big)\, \alpha(t) \\ &= \exp\left(\sum_{u=1}^{p} z^{(u)}(t) \sum_{v=1}^{r_u} \overline{\beta}^{(u,v)} \gamma_\alpha^{(u,v)}(t)\right) \alpha(t),\end{aligned}$$

$\gamma_\alpha^{(\dot{u},\ddot{u})} = \gamma_0^{(\dot{u},\ddot{u})} \circ H_\alpha$, $u = 1, \ldots, r$, are positive weight functions and H_α is some cumulative distribution that might depend on the baseline hazard α and the distribution of the censoring times and $\overline{\beta}^{(u,v)} = \beta^{(\sum_{l=1}^{u-1} r_l + v)}$, $v = 1, \ldots, r_u$, $u = 1, \ldots, p$, see Definition 1.3.2 and Remark 1.3.3.b. Clearly, this setting means that the weight functions $\gamma_0^{(\dot{u},\ddot{u})}$, $u = 1, \ldots, r$, are positive. In other words, we assume that the predictable dual projection of $N^{(i)}$ under $P_{\beta,\alpha}$ is given by

$$A^{(i)}_{\beta,\alpha}(\cdot) = \int_{I(\cdot)} Y^{(i)}(s)\, \lambda_{\beta,\alpha}\big(s \mid Z_i(s)\big)\, \mathrm{d}s.$$

Under the MCRM our testing problem $\mathcal{H}_1$ versus $\mathcal{K}_1$ is equivalent to

$$\widetilde{\mathcal{H}}_1 : \ \beta = 0 \quad \text{versus} \quad \widetilde{\mathcal{K}}_1 : \ \beta \geq 0, \ \beta \neq 0,$$

where we use the notation of Definition 1.3.2. Localizing and embedding our observations $\big(X_{n,i}, \Delta_{n,i}, Z_{n,i}\big)$, $i = 1, \ldots, n$, in a SHPSM the testing problem transforms into $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathcal{J}}$, where $\mathcal{J} = \{1, \ldots, r\}$, see Section 4.1. Consequently,

$$\varphi_{n,1} = \begin{cases} 1, \\ 0, \end{cases} L_{\mathcal{J},1}\big(\widehat{U}_n(\infty), \widehat{V}_n(\infty)\big) - c_{\mathcal{J},1}\big(\alpha, \widehat{V}_n(\infty)\big) \ \begin{matrix} > \\ \leq \end{matrix} \ 0\,,$$

is an admissible test for our testing problem, see Corollary 4.3.6.a. The statistic $\widehat{U}_n(\infty)$ and the variance estimator $\widehat{V}_n(\infty)$ are defined in Theorem 3.2.9, Remark 3.2.10.a and Theorem 3.2.13.

If one wants to test $\mathcal{H}_1$ versus

$$\mathcal{K}_2 : \ \lambda\big(t \mid Z(t) = z(t)\big) < \lambda\big(t \mid Z(t) = \widetilde{z}(t)\big) \text{ for all } t \text{ with } z(t) \gneqq \widetilde{z}(t),$$

one merely has to replace the positive weight functions by negative ones or use the covariate processes $-Z_{n,i} = \big\{-Z_{n,i}(t) \mid t \in \mathbb{R}_+\big\}$, $i = 1, \ldots, n$ instead of $Z_{n,i}$, $i = 1, \ldots, n$. The second proposal is more suitable, if one wants to check, if the condition (4.12) holds.

5.3.2 Example (Two-Sample Problem, Tests of Mayer [53]). Assume that we observe a one-dimensional covariate, *i.e.* $p = 1$ that can only attain the values 0 and 1, where $Z_{n,i} = 1$ (0) means that the i-th observation belongs to the first or the second sample. Assume that we want to test the hypothesis that the distribution of the survival times in both samples is equal versus the alternative that the distribution in the second sample is stochastically larger than the distribution in the first sample This means larger values of the covariates correspond with shorter survival times. Therefore, this testing problem is a special case of Example 5.3.1, see also Example 1.2.1 and Example 1.3.1.

Moreover, assume that $Z_{n,i} = \mathbb{1}(1 \leq i \leq \nu_n)$ where $1 \leq \nu_n < n$. Thus, the sample-size of the first (second) sample is given by ν_n $(n - \nu_n)$. Obviously, in this case the covariates are non-random. If one computes the test statistic $L_{\mathcal{J},1}(\widehat{U}_n, \widehat{V}_n)$, see Section 4.1 and Section 4.3, we receive the one-sided projective-type of Mayer [53]. Setting

$$\widehat{\mu}_{n,0,1} = \frac{1}{n} \sum_{j=1}^{\nu_n} Y_n^{(j)}, \qquad \widehat{\mu}_{n,0,2} = \frac{1}{n} \sum_{j=\nu_n+1}^{n} Y_n^{(j)},$$
$$N_{n,1} = \sum_{i=1}^{\nu_n} N_n^{(i)}, \qquad N_{n,2} = \sum_{i=\nu_n+1}^{n} N_n^{(i)},$$

it hold that

$$\begin{aligned}\widehat{U}_n^{(u)}(\infty) &= \frac{1}{\sqrt{n}} \left(\sum_{i=1}^{\nu_n} \int_{I(\infty)} \widehat{\gamma}_n^{(1,u)} \frac{\widehat{\mu}_{n,0,2}}{\widehat{\mu}_{n,0}} \, \mathrm{d}N_n^{(i)} - \sum_{i=\nu_n+1}^{n} \int_{I(\infty)} \widehat{\gamma}_n^{(1,u)} \frac{\widehat{\mu}_{n,0,1}}{\widehat{\mu}_{n,0}} \, \mathrm{d}N_n^{(i)} \right) \\ &= \frac{1}{\sqrt{n}} \int_{I(\infty)} \widehat{\gamma}_n^{(1,u)} \frac{\widehat{\mu}_{n,0,1}\, \widehat{\mu}_{n,0,2}}{\widehat{\mu}_{n,0}} \, \mathrm{d}\left(\frac{N_{n,1}}{\widehat{\mu}_{n,0,1}} - \frac{N_{n,2}}{\widehat{\mu}_{n,0,2}} \right),\end{aligned}$$

$u = 1, \ldots, r$. Consequently, $\widehat{U}_n(\tau)$ is in this special case a r-dimensional vector of log-rang statistics that are frequently applied for the two-sample problem in survival analysis. Furthermore, the estimator of the covariance matrix boils down to

$$V_n^{(u,v)}(\infty) = \frac{1}{n} \int_{I(\infty)} \widehat{\gamma}_n^{(1,u)}\, \widehat{\gamma}_n^{(1,v)} \frac{\widehat{\mu}_{n,0,1}\, \widehat{\mu}_{n,0,2}}{\widehat{\mu}_{n,0}} \frac{1}{\widehat{\mu}_{n,0}} \, \mathrm{d}\big(N_{n,1} + N_{n,2}\big),$$

$u, v = 1, \ldots, r$, a multivariate version of the variance estimator V_2 in Gill [22, Equation (3.3.12)]. Extensive simulation results for the two-sample problem can be found in Behnen and Neuhaus [8] and Mayer [53, Section 3.6].

Clearly, Assumption 2.3.9.vii is satisfied and Assumption 2.3.9.iv – Assumption 2.3.9.vi hold, if

$$\sup_{s\in[0,t]}\left|\widehat{\mu}_{n,0,1}(s)-\mu_{0,1}(s)\right| \quad \text{and} \quad \sup_{s\in[0,t]}\left|\widehat{\mu}_{n,0,2}(s)-\mu_{0,2}(s)\right|,$$

for all $t<\tau_0$ and $\lim_{n\to\infty}\frac{\nu_n}{n}=\nu\in(0,1)$, where the last condition provides that $\mathscr{J}^*(\tau_0)\neq 0$. Note that these are the classical conditions needed for treatment of the two-sample problem.

5.3.3 Example (Trend Test for Discrete Stages). In a study, the subjects are classified by the stage of their disease, when they start participating in the study.

Assume that we have $p+1$ different stages of the disease. Let λ_j denote the hazard rate determining the distribution of the survival times at stage j. We want to test the hypothesis $\mathcal{H}:\ \lambda_1=\lambda_2=\ldots=\lambda_{p+1}$ versus the alternative $\mathcal{K}:\ \lambda_1\leq\lambda_2\leq\ldots\leq\lambda_{p+1}$ with at least one strict inequality. In other words, we want to test that there is no difference in the hazard rate among the stages versus the higher the stage, the higher the death rate.

Now, we show how this statistical question can be modelled with the MCRM. Choosing non-negative weight functions, we have merely to define suitable covariates. Suppose that $\overline{Z}_{n,i}=j$, if and only if the i-th subject is classified stage j. A possible choice of the covariates is given by

$$Z_{n,i}=\left(Z_{n,i}^{(1)},\ldots,Z_{n,i}^{(p)}\right), \qquad \text{where } Z_{n,i}^{(u)}=\mathbb{1}\left(u<\overline{Z}_{n,i}\right).$$

As we assume that our observation are embedded in a SHPSM, the dual predictable projection of $N_n^{(i)}$ and $P_{n,\xi}$ is given by

$$A_{n,\xi}^{(i)}(\cdot)=\int_{I(\cdot)}\lambda_\beta(s,Z_{n,i})\,Y_n^{(i)}(s)\,\exp\left(\frac{1}{\sqrt{n}}\eta^{\mathrm{T}}\widetilde{\gamma}(s)\right)\alpha_0(s)\,\mathrm{d}s$$

where

$$\begin{aligned}\lambda_\beta(s, Z_{n,i}) &= \exp\left(\frac{1}{\sqrt{n}} \sum_{u=1}^{r} \beta^{(u)}\, Z_{n,i}^{(\dot u)}\, \gamma^{(\dot u, \ddot u)}(s)\right) \\ &= \exp\left(\frac{1}{\sqrt{n}} \sum_{u=1}^{p} Z_{n,i}^{(u)} \sum_{v=1}^{r_u} \overline{\beta}^{(u,v)} \gamma^{(u,v)}(s)\right) \\ &= \exp\left(\frac{1}{\sqrt{n}} \sum_{u=1}^{\overline{Z}_{n,i}-1} \sum_{v=1}^{r_u} \overline{\beta}^{(u,v)} \gamma^{(u,v)}(s)\right),\end{aligned}$$

$\overline{\beta}^{(u,v)} = \beta^{(\sum_{l=1}^{u-1} r_l + v)}$, $v = 1, \dots, r_u$, $u = 1, \dots, p$, see Remark 1.3.3.b.

Obviously, our testing problem $\mathcal{H}$ versus $\mathcal{K}$ transform into $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathcal{J}}$, where $\mathcal{J} = \{1, \dots, r\}$ – the same testing problem as in Example 5.3.1, where we stated the corresponding test procedure.

5.3.4 Example (Trend Test for Continuous Stages). Now we want to extend Example 5.3.3 and allow as covariate not only discrete stages, but also continuous stages, *i.e.* $\overline{Z}_{n,i}$ can take values in whole $\mathbb{R}$ and not only $1, 2, ., p$. Of course, one can always discretize $\overline{Z}_{n,i}$ but to do this one has to have an idea of which values of $\overline{Z}_{n,i}$ represent a certain stage of a disease.

The following example might serve as an illustration. Assume that we can observe the time between the beginning of a disease and the treatment, this might be the case if you consider organ transplantation. A possible statistical question would be to test, the longer the time between the beginning of the disease and the treatment, the shorter the survival time after the treatment. In this case, $X_{n,i}$ denotes the time between treatment and death or censoring of the i-th subject. And let $\overline{Z}_{n,i}$ be the time before the treatment of the i-th subject. Clearly, in this example we could replace time before the transplantation by any real valued quantity that can be measured, when a subject enters a study, *e.g.* the number of white blood cells.

Choosing non-negative weight functions, we have to define suitable covariates.

A possible choice of the covariates would be $Z_{n,i} = \overline{Z}_{n,i}$ implying that

$$\lambda_\beta(s, Z_{n,i}) = \exp\left(\frac{1}{\sqrt{n}}\overline{Z}_{n,i}\sum_{u=1}^{r}\beta^{(u)}\gamma^{(1,u)}(s)\right),$$

see Example 5.3.3. And again, our testing problem that the stage of a disease does not effect the survival time versus the higher the stage, the higher the death rate transforms into $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathcal{J}}$, where $\mathcal{J} = \{1, \ldots, r\}$. However, a linear influence might be too restrictive. Therefore, let $g_u : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$, $u = 1, \ldots, p$, some strictly increasing, known functions that are linearly independent. A different choice of the covariate could be

$$Z_{n,i} = \left(g_1(\overline{Z}_{n,i}), \ldots, g_p(\overline{Z}_{n,i})\right)^{\mathrm{T}}.$$

Consequently, we get that

$$\lambda_\beta(s, Z_{n,i}) = \exp\left(\frac{1}{\sqrt{n}}\sum_{u=1}^{p} g_u(\overline{Z}_{n,i})\sum_{v=1}^{r_u}\overline{\beta}^{(u,v)}\gamma^{(u,v)}(s)\right)$$

Choosing $\gamma^{(u,v)} = \gamma^{(v)}$, and setting $r_1 = \ldots = r_p = \widetilde{r}$, this expression simplifies to

$$\lambda_\beta(s, Z_{n,i}) = \exp\left(\frac{1}{\sqrt{n}}\sum_{v=1}^{\widetilde{r}}\gamma^{(v)}(s)\sum_{u=1}^{p}\overline{\beta}^{(u,v)}g_u(\overline{Z}_{n,i})\right).$$

In this model, covariates and weight functions change their roles to some extend, as we use the g_u to approximate some unknown link function g. Again, our testing problem transforms into $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathcal{J}}$.

So far, we have mainly considered applications with time-independent covariates, now, let us consider an example with time-dependent covariates.

5.3.5 Example (Application of Example 5.3.1). Assume we want to find out, if disease complications in the recovery phase lead to shorter survival times. Therefore, let us define the covariates as follows

$$Z_{n,i}^{(u)}(t) = \begin{cases} 1, & t \geq \text{time at which disease complication } u \text{ occurs}, \\ 0. & \text{otherwise}, \end{cases}$$

$u = 1, \ldots, p$. In this model we presume that the point in time at which the disease complication occurs is important. With this choice of covariates our statistical question boils down to test the hypothesis that the covariates have no influence on the survival times versus the alternative that the larger the values of the covariates, the shorter the survival times. This problem was already treated in Example 5.3.1.

5.3.6 Example (Two-sided Tests). Assume we want to find out whether the observed covariates have any effect on survival times at all. As we consider a SHPSM, the predictable dual projection of $N_n^{(i)}$ under $P_{n,\xi}$ is given by

$$A_{n,\xi}^{(i)}(\cdot) = \int_{I(\cdot)} \exp\left(\frac{1}{\sqrt{n}} \cdot \beta^{\mathrm{T}} Z_{n,i} \circledcirc \gamma(s) + \frac{1}{\sqrt{n}} \cdot \eta^{\mathrm{T}} \widetilde{\gamma}(s)\right) Y_n^{(i)}(s)\, \alpha_0(s)\, \mathrm{d}s,$$

where $\gamma^{(u,v)}$, $v = 1, \ldots, r_u$, $u = 1, \ldots, p$, are any weight functions. Consequently, our testing problem transforms into $\widetilde{\mathcal{H}}_2^{\mathcal{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$, where $\mathcal{L}_1 = \mathbb{R}^r$ and $\mathcal{L}_0 = 0 \in \mathbb{R}^r$, see Section 4.2. Consequently,

$$\varphi_{n,2} = \begin{cases} 1, & \\ 0, & \end{cases} L_{\mathcal{L}_0,\mathcal{L}_1,2}\big(\widehat{U}_n(\infty), \widehat{V}_n(\infty)\big) - c_{\mathcal{L}_0,\mathcal{L}_1,2}\big(\alpha, \widehat{V}_n(\infty)\big) \begin{array}{c} > \\ \leq \end{array} 0$$

is an admissible test for our testing problem, see Corollary 4.3.6.b. The statistic $\widehat{U}_n(\infty)$ and the variance estimator $\widehat{V}_n(\infty)$ are defined in Theorem 3.2.9, Remark 3.2.10.a and Theorem 3.2.13.

5.3.7 Example (*k*-Sample Problem). Let λ_j denote the hazard rate of the j-th sample and suppose we want to test the hypothesis $\mathcal{H}: \ \lambda_1 = \ldots = \lambda_k$ versus the alternative $\mathcal{K}: \ \lambda_i \neq \lambda_j$ for at least one pair (i,j), $i,j \in \{1, \ldots, k\}$. Choosing arbitrary weight functions and defining the covariates as

$$Z_{n,i}^{(u)} = \begin{cases} 1, & i\text{-th observation belongs to sample } u, \\ 0, & \text{otherwise,} \end{cases} \quad, \quad u = 1, \ldots, k,$$

the testing problem $\mathcal{H}$ versus $\mathcal{K}$ is a special case of Example 5.3.7.

5.3.8 Example (Competing Risks). If one changes the interpretation of the covariate in Example 5.3.7 to

$$Z_{n,i}^{(u)} = \begin{cases} 1, & i\text{-th subject dies of cause } u, \\ 0, & \text{otherwise,} \end{cases} \quad , \quad u = 1, \ldots, k,$$

we receive a model for competing risks.

5.3.9 Example (Two-Sample Problem with Concomitant Covariates). In Example 5.3.2, we considered the two sample problem and presented a testing procedure for the hypothesis no differences between the two samples versus the alternative the distribution of the second sample is stochastically larger, where $Z_{n,i} = 1$ (0) means that the i-th observation belongs to the first (second) sample. In the two sample problem the group membership often depends on the kind of treatment a subject receives, more precisely the subjects in the first (second) sample receive the standard (new) treatment.

However, the individuals within one sample might differ in various characteristics that might have an additional impact on the survival times, but we only want to know, if the group membership leads to difference in the survival times. Such a characteristic could be gender for example.

Of course this situation can be modelled with the MCRM. Assume that our covariate process $\{Z_{n,i}(t) \mid t \in \mathbb{R}_+\}$, $i = 1, \ldots, n$, are interpreted as follows $Z_{n,i}(t)^{(1)} = 1$ (0), if the i-th observation belongs to the first (second) sample. $\{Z_{n,i}^{(u)} \mid t \in \mathbb{R}_+\}$, $u = 2, \ldots, p$ denote the concomitant covariates that might have an impact on the survival times. As we consider a SHPSM the predictable dual projection of $N_n^{(i)}$ under $P_{n,\xi}$ is given by

$$A_{n,\xi}^{(i)}(\cdot) = \int_{I(\cdot)} \lambda_\beta(s, Z_{n,i}(s))\, Y_n^{(i)}(s) \, \exp\left(\frac{1}{\sqrt{n}} \eta^{\mathrm{T}} \widetilde{\gamma}(s)\right) \alpha_0(s) \, \mathrm{d}s,$$

where

$$\lambda_\beta(s, Z_{n,i}(s)) = \exp\left(Z_{n,i}^{(1)}(s) \sum_{u=1}^{r_1} \beta^{(u)} \gamma^{(1,u)}(s) + \sum_{u=r_1+1}^{r} Z_{n,i}^{(\dot{u})}(s) \beta^{(u)} \gamma^{(\dot{u},\ddot{u})}(s) \right).$$

Furthermore, let us assume that the weight functions $\gamma^{(1,1)}, \dots, \gamma^{(1,r_1)}$ are non-negative. The remaining weight functions $\gamma^{(\dot{u},\ddot{u})}$, $u = r_1+1, \dots, r$ can be chosen arbitrarily. In this setting our testing problem in question transforms into the testing problem $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathcal{J}}$, where the set $\mathcal{J} = \{1, \dots, r_1\}$, see Section 4.1. The corresponding test is given by

$$\varphi_{n,1} = \begin{cases} 1, \\ 0, \end{cases} L_{\mathcal{J},1}\big(\widehat{U}_n(\infty), \widehat{V}_n(\infty)\big) - c_{\mathcal{J},1}\big(\alpha, \widehat{V}_n(\infty)\big) \begin{array}{c} > \\ \leq \end{array} 0\ ,$$

see Corollary 4.3.6.a. The statistic $\widehat{U}_n(\infty)$ and the variance estimator $\widehat{V}_n(\infty)$ are defined in Theorem 3.2.9, Remark 3.2.10.a and Theorem 3.2.13.

The testing problem no influence of the group membership versus there is an influence of the group membership on the survival times the under this model transforms into $\mathcal{H}_2^{\mathcal{L}_0}$ versus $\mathcal{K}_2^{\mathcal{L}_1}$, where $\mathcal{L}_1 = \mathrm{Im}(\mathscr{T}^r_{\{1,\dots,r\}})$ and $\mathcal{L}_0 = \mathrm{Im}(\mathscr{T}^r_{\{r_1+1,\dots,r\}})$. See Definition 4.1.3 for the matrices $\mathscr{T}^r_{\{1,\dots,r\}}$ and $\mathscr{T}^r_{\{r_1+1,\dots,r\}}$. The corresponding testing procedure is given by

$$\varphi_{n,2} = \begin{cases} 1, \\ 0, \end{cases} L_{\mathcal{L}_0,\mathcal{L}_1,2}\big(\widehat{U}_n(\infty), \widehat{V}_n(\infty)\big) - c_{\mathcal{L}_0,\mathcal{L}_1,2}\big(\alpha, \widehat{V}_n(\infty)\big) \begin{array}{c} > \\ \leq \end{array} 0\ ,$$

cf. Corollary 4.3.6.b. Analogue to the two-sided testing problem, see Example 5.3.6, we do not have to restrict ourselves to non-negative weight functions $\gamma^{(1,1)}, \dots, \gamma^{(1,r_1)}$ in order to get the "right" stochastic ordering.

5.3.10 Remark. In Example 5.3.9 we extended Example 5.3.2 by concomitant covariates. With an analogue proceeding one can also extended the other Examples discussed in this section so far by concomitant covariates. However, one should keep in mind that an extension by concomitant covariates also increases the number of model parameters. In order to obtain reasonable results in such situations, the number of observations available for the analysis has to be adequately large.

5.3.11 Example (Model Check). Analogue to the previous examples we consider again a SHPSM and assume that the predictable dual projection of

$N_n^{(i)}$ under $P_{n,\xi}$ is given by

$$A_{n,\xi}^{(i)}(\cdot) = \int_{I(\cdot)} \lambda_\beta\big(s, Z_{n,i}(s)\big)\, Y_n^{(i)}(s)\, \exp\Big(\frac{1}{\sqrt{n}}\eta^{\mathrm{T}}\widetilde{\gamma}(s)\Big)\, \alpha_0(s)\, \mathrm{d}s,$$

where

$$\lambda_\beta\big(s, Z_{n,i}(s)\big) = \\ \exp\left(\sum_{u=1}^{p} Z_{n,i}^{(u)}(s)\left(\sum_{v=1}^{\widetilde{r}_u} \overline{\beta}^{(u,v)}\gamma^{(u,v)}(s) + \sum_{v=\widetilde{r}_u+1}^{r_u} \overline{\beta}^{(u,v)}\gamma^{(u,v)}(s)\right)\right),$$

$\overline{\beta}^{(u,v)} = \beta^{(\sum_{l=1}^{u-1} r_l + v)}$, $v = 1,\dots,r_u$, $u = 1,\dots,p$, see Definition 1.3.2 and Remark 1.3.3.b. Moreover, assume that $\widetilde{r}_u \le r_u$, $u = 1,\dots,r$, where at least one of the inequalities is strict. Considering the testing problem $\widetilde{\mathcal{H}}_2^{\mathcal{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$, where $\mathcal{L}_1 = \mathbb{R}^r$ and $\mathcal{L}_0 = \mathrm{Im}\big(\mathscr{T}_{\mathcal{J}}^r\big)$ with

$$\mathcal{J} = \bigcup_{u=1}^{p}\left\{\sum_{v=1}^{u-1} r_v + 1, \dots\dots, \sum_{v=1}^{u-1} r_v + \widetilde{r}_u\right\},$$

is a possible simple way to check, if the family of probability $\{P_{n,\xi} \mid \xi \in \mathcal{L}_0\}$ measures is sufficient to model the given observations, or if we have to introduce more weight functions. The alternative models are specified by the weight functions $\gamma^{(u,v)}$, $v = \widetilde{r}_u + 1,\dots,r_u$, $u = 1,\dots,p$. The corresponding testing procedure is given by Corollary 4.3.6.b.

5.3.12 Example. Assume that $1 \le \widetilde{p} < p$ and that we want to test the hypothesis that only the components $u = 1,\dots,\widetilde{p}$ of the covariate process have some effect on the survival times versus the alternative that also other components of the covariate process have some influence on the survival time. Clearly, this statistical question can be modelled with the MCRM. Considering a SHPSM the non-parametric testing problem transforms to $\widetilde{\mathcal{H}}_2^{\mathcal{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$, where $\mathcal{L}_1 = \mathbb{R}^r$ and $\mathcal{L}_0 = \mathrm{Im}\big(\mathscr{T}_{\mathcal{J}}^r\big)$ with $\mathcal{J} = \big\{1,\dots \sum_{u=1}^{\widetilde{p}} r_u\big\}$, see also Example 5.3.11.

6 Generalized Permutation Tests

In this chapter a different method for determining the critical values for the tests developed in Chapter 4 is presented. The advantage of these critical values is that they do not only converge to the correct asymptotic critical values, but in certain cases they are the exact critical values for finite $n \in \mathbb{N}$, so that our tests hold the level even for finite $n \in \mathbb{N}$. In Section 6.1 we motivate and introduce so called permutation tests. Before we can state our main result – the asymptotic equivalence of permutation tests and the tests developed in Chapter 4 – in Section 6.2, we have to present some rather technical results. The whole proceeding is based on ideas that were developed by Neuhaus [59]. A similar approach is also used by Janssen and Mayer [37] who investigate conditional studentized permutation tests. The assumptions needed in the proof of our main result are discussed and verified for important examples in Section 6.3 in detail.

Unfortunately, this method only applies to external covariates, see Section 1.1, as our method requires that we can observe the covariates determining the survival time of a subject even after the death of that subject. Obviously this condition is satisfied, if we consider time-independent covariates, which are a major example for external covariates.

6.1 Introduction

In the following it is supposed that Assumption 4.3.3 holds with $\tau = \tau_0^c$. Conditions implying these premises were discussed in Chapter 5. So without loss of generality, we can assume that we are in the situation of Discussion 5.1.11, which means in particular that Assumption 5.1.1 is valid.

Remember that the censored survival time $X_{n,i}$ and the censoring status $\Delta_{n,i}$ of the i-th observation are given by

$$X_{n,i} = \sup\bigl\{t \mid Y_n^{(i)}(t) = 1\bigr\}, \quad \Delta_{n,i} = N_n^{(i)}(X_{n,i}), \qquad i = 1, \ldots, n,\ n \in \mathbb{N}.$$

The ranks of the censored survival times $X_{n,i}$, $i = 1, \ldots, n$ are denoted by $R_n = \bigl(R_{n,1}, \ldots, R_{n,n}\bigr)$, where $R_{n,i} = \sum_{j=1}^n \mathbb{1}(X_{n,j} \le X_{n,i})$, $i = 1, \ldots, n$. The inverse ranks $D_n = \bigl(D_{n,1}, \ldots, D_{n,n}\bigr)$ are defined by the identities $D_{n,R_{n,i}} = R_{n,D_{n,i}} = i$, $i = 1, \ldots, n$. The statistic

$$X_{n,\uparrow} = \bigl(X_{n:1}, \ldots, X_{n:n}\bigr) = \bigl(X_{n,D_{n,1}}, \ldots, X_{n,D_{n,n}}\bigr)$$

is the order statistic of the observations. The statistic

$$\Delta_{n,\uparrow} = \bigl(\Delta_{n:1}, \ldots, \Delta_{n:n}\bigr) = \bigl(\Delta_{n,D_{n,1}}, \ldots, \Delta_{n,D_{n,n}}\bigr)$$

is called the concommitant order statistic of the censoring indicators. Finally, the reduced covariate processes are given by $Z_{n,\uparrow} = \bigl(Z_{n,\uparrow,1}, \ldots, Z_{n,\uparrow,n}\bigr)$, where

$$Z_{n,\uparrow,i} = \begin{pmatrix} Z_{n,\uparrow,i}^{(1,1)} & \ldots & Z_{n,\uparrow,i}^{(1,n)} \\ \vdots & \ddots & \vdots \\ Z_{n,\uparrow,i}^{(p,1)} & \ldots & Z_{n,\uparrow,i}^{(p,n)} \end{pmatrix} = \begin{pmatrix} Z_{n,i}^{(1)}(X_{n:1}) & \ldots & Z_{n,i}^{(1)}(X_{n:n}) \\ \vdots & \ddots & \vdots \\ Z_{n,i}^{(p)}(X_{n:1}) & \ldots & Z_{n,i}^{(p)}(X_{n:n}) \end{pmatrix},$$

indicate why we only consider external covariates in this chapter.

Additionally, let us suppose that $\widehat{\gamma}_n^{(\dot{u},\ddot{u})}(\cdot) = \widehat{\gamma}_n^{(\dot{u},\ddot{u})}(\cdot \mid X_{n,\uparrow}, \Delta_{n,\uparrow})$, *i.e.* the estimators for the weight functions only depend on the order statistics. For example, estimators based on the Kaplan-Meyer estimator or on $\widehat{\mu}_{n,0}$ satisfy this condition, see Example 5.2.7 for details. Finally, set $\widehat{\gamma}_{n:i}^{(\dot{u},\ddot{u})} = \widehat{\gamma}_n^{(\dot{u},\ddot{u})}(X_{n:i})$, $u = 1, \ldots, r$, $i = 1, \ldots, n$.

One readily checks that

$$\begin{aligned} \widehat{U}_n^{(u)}(\infty) &= \frac{1}{\sqrt{n}} \sum_{l=1}^n \Delta_{n:l} \widehat{\gamma}_{n:l}^{(\dot{u},\ddot{u})} \left(Z_{n,\uparrow,D_{n,l}}^{(\dot{u},l)} - \frac{\sum_{j=l}^n Z_{n,\uparrow,D_{n,j}}^{(\dot{u},l)}}{n+1-l} \right) \\ &= \widehat{U}_{n,\star}^{(u)}(D_n, W_{n,\uparrow}), \end{aligned}$$

$u = 1, \ldots, r$, and that

$$\widehat{V}_n^{(u,v)}(\infty) = \frac{1}{n} \sum_{l=1}^{n} \Delta_{n:l}\, \widehat{\gamma}_{n:l}^{(\dot{u},\ddot{u})} \widehat{\gamma}_{n:l}^{(\dot{v},\ddot{v})}$$

$$\times \left(\frac{\sum_{j=l}^{n} Z_{n,\uparrow,D_{n,j}}^{(\dot{u},l)} Z_{n,\uparrow,D_{n,j}}^{(\dot{v},l)}}{n+1-l} - \frac{\sum_{j=l}^{n} Z_{n,\uparrow,D_{n,j}}^{(\dot{u},l)} \sum_{k=l}^{n} Z_{n,\uparrow,D_{n,k}}^{(\dot{v},l)}}{(n+1-l)^2} \right)$$

$$= \widehat{V}_{n,\star}^{(u,v)}(D_n, W_{n,\uparrow}),$$

$u, v = 1, \ldots, r$, where $W_{n,\uparrow} = (X_{n,\uparrow}, \Delta_{n,\uparrow}, Z_{n,\uparrow})$. Moreover, we note that the covariate processes $\{Z_{n,i}(t) \mid t \in \mathbb{R}_+\}$, $i = 1, \ldots, n$, and the multivariate counting process $(N_n, \widetilde{N}_n)$ are stochastically independent under $P_{n,0}$. Thus, $\{Z_{n,i}(t) \mid t \in \mathbb{R}_+\}$, $i = 1, \ldots, n$, and $(X_{n,i}, \Delta_{n,i})$, $i = 1, \ldots, n$, are stochastically independent under $P_{n,0}$. If additionally all censoring times have the same distribution, *i.e.*

$$\widetilde{\alpha}_{n,1} = \ldots = \widetilde{\alpha}_{n,n}, \tag{6.1}$$

see Assumption 5.1.1.vii, Remark 5.1.4, and Assumption 2.2.1.ix, it holds that $(X_{n,i}, \Delta_{n,i})$, $i = 1, \ldots, n$, are stochastically independent and identically distributed (i.i.d.). In particular, the ranks of our observations R_n and the inverse ranks D_n are uniformly distributed on $\mathrm{Per}(1, \ldots, n)$ under $P_{n,0}$, where $\mathrm{Per}(1, \ldots, n)$ denotes the set of all permutation of the numbers $1, \ldots, n$. More precisely, it holds that

$$P_{n,0}\big(R_n = r\big) = P_{n,0}\big(D_n = r\big) = \frac{1}{n!}, \qquad r \in \mathrm{Per}(1, \ldots, n).$$

Furthermore, one can easily show that D_n and $\big(X_{n,\uparrow}, \Delta_{n,\uparrow}, Z_{n,\uparrow}\big)$ are stochastically independent under $P_{n,0}$, if condition (6.1) is satisfied. In this situation assume that the order statistics and the reduced covariate processes are given and fixed, *i.e.* $X_{n,\uparrow} = x_{n,\uparrow}$, $\Delta_{n,\uparrow} = \delta_{n,\uparrow}$ and $Z_{n,\uparrow} = z_{n,\uparrow}$. Then the distribution of the statistics, cf. Section 4.3,

$$T_{\mathcal{J},1}^{\star,1,\alpha}(D_n, w_{n,\uparrow}) =$$

$$L_{\mathcal{J},1}\big(\widehat{U}_{n,\star}(D_n, w_{n,\uparrow}), \widehat{V}_{n,\star}(D_n, w_{n,\uparrow})\big) - c_{\mathcal{J},1}\big(\alpha, \widehat{V}_{n,\star}(D_n, w_{n,\uparrow})\big)$$

and

$$T^{\star,2,\alpha}_{\mathfrak{L}_0,\mathfrak{L}_1}(D_n, w_{n,\uparrow}) = \\ L_{\mathfrak{L}_0,\mathfrak{L}_1,2}\big(\widehat{U}_{n,\star}(D_n, w_{n,\uparrow}), \widehat{V}_{n,\star}(D_n, w_{n,\uparrow})\big) - c_{\mathfrak{L}_0,\mathfrak{L}_1,2}\big(\alpha, \widehat{V}_{n,\star}(D_n, w_{n,\uparrow})\big)$$

where we set $w_{n,\uparrow} = \big(x_{n,\uparrow}, \delta_{n,\uparrow}, z_{n,\uparrow}\big)$, is principally known and can be easily approximated by simulations. This observation leads to the introduction of conditional permutation tests.

Let $F^{\star,1,\alpha}_{n,\mathfrak{J},w_{n,\uparrow}}$ and $F^{\star,2,\alpha}_{n,\mathfrak{L}_0,\mathfrak{L}_1,w_{n,\uparrow}}$ denote the cumulative distribution functions of the statistic $T^{\star,1,\alpha}_{\mathfrak{J},1}(D_n, w_{n,\uparrow})$ and $T^{\star,2,\alpha}_{\mathfrak{L}_0,\mathfrak{L}_1}(D_n, w_{n,\uparrow})$, in the case that D_n is uniformly distributed on $\mathrm{Per}(1,\ldots,n)$. Obviously, for every $\alpha \in (0,1)$ there exists real numbers

$$r^{\star,1}_{n,\mathfrak{J}}(\alpha, w_{n,\uparrow}),\ r^{\star,2}_{n,\mathfrak{L}_0,\mathfrak{L}_1}(\alpha, w_{n,\uparrow}) \quad \text{and} \quad k^{\star,1}_{n,\mathfrak{J}}(\alpha, w_{n,\uparrow}),\ k^{\star,2}_{n,\mathfrak{L}_0,\mathfrak{L}_1}(\alpha, w_{n,\uparrow}),$$

such that

$$\int_{[0,\infty)} \phi^{\star,1}_{n,\mathfrak{J}}(s, w_{n,\uparrow})\, \mathrm{d}F^{\star,1,\alpha}_{n,\mathfrak{J},w_{n,\uparrow}}(s) = \alpha$$

and

$$\int_{[0,\infty)} \phi^{\star,2}_{n\mathfrak{L}_0,\mathfrak{L}_1}(s, w_{n,\uparrow})\, \mathrm{d}F^{\star,2,\alpha}_{n,\mathfrak{L}_0,\mathfrak{L}_1,w_{n,\uparrow}}(s) = \alpha,$$

where

$$\phi^{\star,1}_{n,\mathfrak{J}}(s, w_{n,\uparrow}) = \left\{ \begin{array}{cc} 1, & > \\ r^{\star,1}_{n,\mathfrak{J}}(\alpha, w_{n,\uparrow}),\ s - k^{\star,1}_{n,\mathfrak{J}}(\alpha, w_{n,\uparrow}) & = 0 \\ 0, & < \end{array} \right.$$

and

$$\phi^{\star,2}_{n,\mathfrak{L}_0,\mathfrak{L}_1}(s, w_{n,\uparrow}) = \left\{ \begin{array}{cc} 1, & > \\ r^{\star,2}_{n,\mathfrak{L}_0,\mathfrak{L}_1}(\alpha, w_{n,\uparrow}),\ s - k^{\star,2}_{n,\mathfrak{L}_0,\mathfrak{L}_1}(\alpha, w_{n,\uparrow}) & = 0\ . \\ 0, & < \end{array} \right.$$

The sequences of tests

$$\varphi^{\star}_{n,1} = \phi^{\star,1}_{n,\mathfrak{J}}\big(T^{\star,1}_{\mathfrak{J},1,\alpha}(D_n, W_{n,\uparrow}), W_{n,\uparrow}\big), \quad n \in \mathbb{N},$$

and

$$\varphi^{\star}_{n,2} = \phi^{\star,2}_{n,\mathcal{L}_0,\mathcal{L}_1}\big(T^{\star,2,\alpha}_{\mathcal{L}_0,\mathcal{L}_1,2}(D_n, W_{n,\uparrow}), W_{n,\uparrow}\big), \quad n \in \mathbb{N},$$

are called (conditional) permutation tests of level α for the testing problems $\widetilde{\mathcal{H}}_1^{\mathcal{J}}$ versus $\widetilde{\mathcal{K}}_1^{\mathcal{J}}$ and $\widetilde{\mathcal{H}}_2^{\mathcal{L}_0}$ versus $\widetilde{\mathcal{K}}_2^{\mathcal{L}_1}$, respectively, cf. Hájek and Šidák [26, p. 42]. Summarizing the previous discussion gives the following result.

6.1.1 Proposition. In the situation of Discussion 5.1.11 assume that the condition (6.1) is satisfied and that $\widehat{\gamma}_n^{(\dot{u},\ddot{u})}$, $u = 1, \dots, r$, only depend on the order statistics $X_{n,\uparrow}$ and $\Delta_{n,\uparrow}$, then it holds that

$$\int \varphi^{\star}_{n,1} \,\mathrm{d}P_{n,0} = \int \varphi^{\star}_{n,2} \,\mathrm{d}P_{n,0} = \alpha, \qquad n \in \mathbb{N}.$$

Proof. Because of condition (6.1), it holds that $(X_{n,i}, \Delta_{n,i})$, $i = 1, \dots, n$, are i.i.d. implying that D_n is uniformly distributed on $\mathrm{Per}(1, \dots, n)$ under $P_{n,0}$. As D_n and $\big(X_{n,\uparrow}, \Delta_{n,\uparrow}, Z_{n,\uparrow}\big)$ are stochastically independent under $P_{n,0}$, conditioning gives that

$$\begin{aligned}
\mathbb{E}_{n,0}(\varphi^{\star}_{n,1}) &= \int \mathbb{E}_{n,0}\big[\varphi^{\star}_{n,1} \mid (X_{n,\uparrow}, \Delta_{n,\uparrow}, Z_{n,\uparrow}) = w_{n,\uparrow}\big] \,\mathrm{d}P_{n,0}^{(X_{n,\uparrow},\Delta_{n,\uparrow},Z_{n,\uparrow})}(w_{n,\uparrow}) \\
&= \int \mathbb{E}_{n,0}\Big(\phi^{\star,1}_{n,\mathcal{J}}\big(T^{\star,1,\alpha}_{\mathcal{J},1}(D_n, w_{n,\uparrow}), w_{n,\uparrow}\big)\Big) \,\mathrm{d}P_{n,0}^{(X_{n,\uparrow},\Delta_{n,\uparrow},Z_{n,\uparrow})}(w_{n,\uparrow}) \\
&= \alpha,
\end{aligned}$$

where we use that $\mathbb{E}_{n,0}\Big(\phi^{\star,1}_{n,\mathcal{J}}\big(T^{\star,1,\alpha}_{\mathcal{J},1}(D_n, w_{n,\uparrow}), w_{n,\uparrow}\big)\Big) = \alpha$ by construction. The second assertion is shown completely analogously. □

6.2 Asymptotic Equivalence

In this section we show that the sequences of tests $\varphi^{\star}_{n,j}$, $n \in \mathbb{N}$, and $\varphi_{n,j}$, $n \in \mathbb{N}$, are asymptotically equivalent, $j = 1, 2$, which implies that the sequences of tests $\varphi^{\star}_{n,j}$, $n \in \mathbb{N}$, $j = 1, 2$, keep asymptotically the level on the hypothesis, are asymptotically unbiased and asymptotically admissible, cf. Corollary 4.3.6. However, before we can proof such a result, we need some more assumptions and notation.

6.2.1 Assumption. Suppose that Assumption 4.3.3 holds and set

$$\Omega = \bigtimes_{n=1}^{\infty} \Omega_n, \quad \mathcal{F} = \bigotimes_{n=1}^{\infty} \mathcal{F}_n, \quad \Big\{ P_\xi = \bigotimes_{n=1}^{\infty} P_{n,\xi} \;\Big|\; \xi \in \mathbb{R}^{r+q} \Big\}.$$

Moreover, let $\big(\Omega_n', \mathcal{F}_n', P_n'\big)$, $n \in \mathbb{N}$, be a sequence of probability spaces, where $\mathcal{F}_n'$ is P_n'-complete. Assume that $D_n' = \big(D_{n,1}', \ldots, D_{n,n}'\big) : \Omega_n' \longrightarrow \operatorname{Per}(1, \ldots, n)$ are uniformly distributed random permutations, *i.e.*

$$P_n'\big(D_n' = d\big) = \frac{1}{n!}, \qquad d \in \operatorname{Per}(1, \ldots, n).$$

Computing the asymptotic distribution of

$$T_{\mathcal{J},1}^{\star,1,\alpha}\big(D_{k_n}', W_{k_n,\uparrow}(\omega_{k_n})\big) \qquad \text{and} \qquad T_{\mathcal{L}_0,\mathcal{L}_1}^{\star,2,\alpha}\big(D_{k_n}', W_{k_n,\uparrow}(\omega_{k_n})\big),$$

where $W_{k_n,\uparrow}(\omega_{k_n}) = \big(X_{n_k,\uparrow}(\omega_{k_n}), \Delta_{n_k,\uparrow}(\omega_{k_n}), Z_{n_k,\uparrow}(\omega_{k_n})\big)$, for sub-sequences of natural numbers k_n, $n \in \mathbb{N}$, and fixed $\omega = (\omega_1, \omega_2, \ldots,) \in \Omega$ is a key step in the proof of the asymptotic equivalence of $\varphi_{n,j}^\star$, $n \in \mathbb{N}$, and $\varphi_{n,j}$, $n \in \mathbb{N}$, $j = 1, 2$. As these statistics depend on the statistic $\widehat{U}_{k_n,\star}\big(D_{k_n}', W_{k_n,\uparrow}(\omega_{k_n})\big)$ and the variance estimator $\widehat{V}_{n,\star}\big(D_{k_n}', W_{k_n,\uparrow}(\omega_{k_n})\big)$, we mainly have to prove a central limit theorem for $\widehat{U}_{k_n,\star}\big(D_{k_n}', W_{k_n,\uparrow}(\omega_{k_n})\big)$, $n \in \mathbb{N}$, for fixed $\omega \in \Omega$. For deriving such a result it is intended to apply Rebolledo's Martingale Limit Theorem, see Theorem 2.1.2. Unfortunately, we have to introduce some more notation.

6.2.2 Definition. Suppose that Assumption 6.2.1 holds. For fixed $\omega = (\omega_1, \omega_2, \ldots) \in \Omega$, we define

$$\zeta_{n,0} = \zeta_{n,0}(\omega) = 0 \in \mathbb{R}^{p \times n} \;\text{ and }\; \zeta_{n,i} = \zeta_{n,i}(\omega) = Z_{n,\uparrow,D_{n,i}'}(\omega_n),\; i = 1, \ldots, n,$$

and

$$\bar{\delta}_{n,0} = \bar{\delta}_{n,0}(\omega) = 0 \in \mathbb{R} \;\text{ and }\; \bar{\delta}_{n,i} = \bar{\delta}_{n,i}(\delta) = \Delta_{n:i}(\omega_n),\; i = 1, \ldots, n,$$

as well as

$$\bar{\gamma}_{n,0}^{(\dot{u},\ddot{u})} = \bar{\gamma}_{n,0}^{(\dot{u},\ddot{u})}(\omega) = 0 \in \mathbb{R} \;\text{ and }\; \bar{\gamma}_{n,i}^{(\dot{u},\ddot{u})} = \bar{\gamma}_{n,i}^{(\dot{u},\ddot{u})}(\omega) = \widehat{\gamma}_{n:i}^{(\dot{u},\ddot{u})}(\omega_n),\; i = 1, \ldots, n,$$

$u = 1, \ldots, r$, $n \in \mathbb{N}$. Moreover, we set

$$\widehat{\bar{\mu}}_{n,1}^{(u)}(s) = \widehat{\bar{\mu}}_{n,1}^{(u)}(s, \omega) = \frac{1}{n} \sum_{i=\lfloor ns \rfloor}^{n} \zeta_{n,i}^{(u,\lfloor ns \rfloor)},$$

$$\widehat{\bar{\mu}}_{n,2}^{(u,v)}(s) = \widehat{\bar{\mu}}_{n,2}^{(u,v)}(s, \omega) = \frac{1}{n} \sum_{i=\lfloor ns \rfloor}^{n} \zeta_{n,i}^{(u,\lfloor ns \rfloor)} \, \zeta_{n,i}^{(v,\lfloor ns \rfloor)},$$

$s \in [0,1]$, $u, v = 1, \ldots, p$, $n \in \mathbb{N}$, and

$$h_n(s) = h_n(s, \omega) = \bar{\delta}_{n,\lfloor ns \rfloor}, \qquad s \in [0,1],\ n \in \mathbb{N}.$$

Finally, we define the stochastic processes $\big\{ \widehat{\bar{U}}_n(t) \mid t \in [0,1] \big\}$, where $\widehat{\bar{U}}_n(t) = \Big(\widehat{\bar{U}}_n^{(1)}(t), \ldots, \widehat{\bar{U}}_n^{(r)}(t) \Big)^{\mathrm{T}}$ and

$$\widehat{\bar{U}}_n^{(u)}(t) = \widehat{\bar{U}}_n^{(u)}(t, \omega) = \frac{1}{\sqrt{n}} \sum_{l=1}^{\lfloor nt \rfloor} \bar{\delta}_{n,l} \, \bar{\gamma}_{n,l}^{(\dot{u},\ddot{u})} \bigg(\zeta_{n,l}^{(\dot{u},l)} - \frac{\sum_{j=l}^{n} \zeta_{n,j}^{(\dot{u},l)}}{n+1-l} \bigg),$$

as well as $\big\{ \widehat{\bar{V}}_n(t) \mid t \in [0,1] \big\}$, where $\widehat{\bar{V}}_n(t) = \Big(\widehat{\bar{V}}_n^{(u,v)}(t) \;\Big|\; u, v = 1, \ldots, r \Big)$ and

$$\begin{aligned}
\widehat{\bar{V}}_n^{(u,v)}(t) &= \widehat{\bar{V}}_n^{(u,v)}(t, \omega) \\
&= \int_{[0,t]} \bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{u},\ddot{u})} \bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{v},\ddot{v})} \Bigg(\frac{\widehat{\bar{\mu}}_{n,2}^{(\dot{u},\dot{v})}(s)}{1 - \frac{\lfloor ns \rfloor}{n} + \frac{1}{n}} - \frac{\widehat{\bar{\mu}}_{n,1}^{(\dot{u})}(s) \, \widehat{\bar{\mu}}_{n,1}^{(\dot{v})}(s)}{(1 - \frac{\lfloor ns \rfloor}{n} + \frac{1}{n})^2} \Bigg) h_n(s) \, \mathrm{d}s,
\end{aligned}$$

$n \in \mathbb{N}$.

Obviously, it holds that

$$\widehat{\bar{U}}_n(1) = \widehat{U}_{k_n,\star}\big(D'_{k_n}, W_{k_n,\uparrow}(\omega_{k_n})\big) \text{ and } \widehat{V}_{n,\star}\big(D'_{k_n}, W_{k_n,\uparrow}(\omega_{k_n})\big) = \widehat{\bar{V}}_n(1). \quad (6.2)$$

But before we can proceed in computing the asymptotic distribution of $\widehat{\bar{U}}_n(1)$, we have to compute some quantities which are collected in the following results.

6.2.3 Lemma. Under Assumption 6.2.1 let $\omega = (\omega_1, \omega_2, \ldots) \in \Omega$ be fixed. Using the notation provided in Definition 6.2.2, set

$$\bar{\Omega}_n = \bar{\Omega}_n(\omega) = \big\{ Z_{n,\uparrow,i}(\omega_n) \mid i = 1, \ldots, n \big\} = \big\{ \bar{z}_{n,1}, \ldots, \bar{z}_{n,m_n} \big\},$$

where $1 \le m_n \le n$, and $\nu_{n,i} = \sum_{j=1}^{n} \mathbb{1}\big(\bar{z}_{n,i} = Z_{n,\uparrow,j}(\omega_n) \big)$, $\qquad i = 1, \ldots, m_n$.

a) For all $\bar{x} \in \bar{\Omega}_n$ it holds that

$$P'_n\{\zeta_{n,i} = \bar{x}\} = \frac{1}{n} \sum\nolimits_{j=1}^{m_n} \nu_{n,j} \, \mathbb{1}\big(\bar{z}_{n,j} = \bar{x}\big), \qquad i = 1, \ldots, n.$$

b) For all $\bar{x}_1, \ldots, \bar{x}_n \in \bar{\Omega}_n$ it holds that

$$P'_n\{(\zeta_{n,1}, \ldots, \zeta_{n,n}) = (\bar{x}_1, \ldots, \bar{x}_n)\}$$
$$= \frac{1}{n!} \prod_{i=1}^{m_n} \Big(\nu_{n,i}! \cdot \mathbb{1}\Big(\sum\nolimits_{j=1}^{n} \mathbb{1}\big(\bar{z}_{n,i} = \bar{x}_j\big) = \nu_{n,i} \Big)\Big).$$

c) For all $\bar{x}_1, \ldots, \bar{x}_i \in \bar{\Omega}_n$ it holds that

$$P'_n\{\zeta_{n,i} = \bar{x}_i \mid (\zeta_{n,1}, \ldots, \zeta_{n,i-1}) = (\bar{x}_1, \ldots, \bar{x}_{i-1})\} =$$
$$\frac{1}{n+1-i} \sum\nolimits_{j=1}^{m_n} \mathbb{1}\big(\bar{z}_{n,j} = \bar{x}_i\big) \Big(\nu_{n,j} - \sum\nolimits_{k=1}^{i-1} \mathbb{1}\big(\bar{z}_{n,j} = \bar{x}_k\big) \Big),$$

if the left hand side is defined.

Proof. Straightforward and elementary calculations give the results. □

6.2.4 Proposition. Under Assumption 6.2.1 let $\omega = (\omega_1, \omega_2, \ldots) \in \Omega$ be fixed. Using the notation provided in Definition 6.2.2, let $\mathcal{Z}'_n$ denote the σ-algebra that is generated by all subsets of P'_n-negligible sets and set $\mathbb{F}'_n = \big\{\mathcal{F}'_{n,t} \mid t \in [0,1]\big\}$, where

$$\mathcal{F}'_{n,t} = \mathcal{Z}'_n \vee \sigma\big(\zeta_{n,0}, \zeta_{n,1}, \ldots, \zeta_{n,\lfloor nt \rfloor}\big), \qquad t \in [0,1].$$

Then, the process $c^{\mathrm{T}} \widehat{\bar{U}}_n = \big\{c^{\mathrm{T}} \widehat{\bar{U}}_n(t) \mid t \in [0,1]\big\}$ is a $\mathbb{F}'_n$-martingale, $c \in \mathbb{R}^r$.

Proof. Basically, the process $\big\{\widehat{\bar{U}}_n(t) \mid t \in [0,1]\big\}$ is a process in discrete time. Therefore, it suffices to show that

$$\mathbb{E}'_n\big[c^{\mathrm{T}} \widehat{\bar{U}}_n(i/n) \mid \mathcal{F}'_{n,(i-1)/n}\big] = c^{\mathrm{T}} \widehat{\bar{U}}_n\big((i-1)/n\big) \quad P'_n\text{-almost surely}$$

for $i = 1, \ldots, n$. As $\mathcal{F}'_{n,0} = \mathcal{Z}_n$, it holds that either $P'_n(F) = 0$ or $P'_n(F) = 1$ for $F \in \mathcal{F}'_{n,0}$. Hence, we have P'_n-almost surely the following chain of equalities

$$\begin{aligned}
\mathbb{E}'_n\big[c^{\mathrm{T}}\widehat{\widehat{U}}_n(1/n) \mid \mathcal{F}'_{n,0}\big] &= \mathbb{E}'_n\big(c^{\mathrm{T}}\widehat{\widehat{U}}_n(1/n)\big) \\
&= \frac{\bar{\delta}_{n,1}}{\sqrt{n}} \sum_{u=1}^{r} c^{(u)} \bar{\gamma}_{n,1}^{(\dot{u},\ddot{u})}\ \mathbb{E}'_n\left(\zeta_{n,1}^{(\dot{u},1)} - \frac{1}{n}\sum_{j=1}^{n} \zeta_{n,j}^{(\dot{u},1)}\right) \\
&= \frac{\bar{\delta}_{n,1}}{\sqrt{n}} \sum_{u=1}^{r} \bar{\gamma}_{n,1}^{(\dot{u},\ddot{u})}\ \left(\mathbb{E}'_n\big(\zeta_{n,1}^{(\dot{u},1)}\big) - \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}'_n\big(\zeta_{n,j}^{(\dot{u},1)}\big)\right) \\
&= \frac{\bar{\delta}_{n,1}}{\sqrt{n}} \sum_{u=1}^{r} c^{(u)} \bar{\gamma}_{n,1}^{(\dot{u},\ddot{u})}\ \left(\mathbb{E}'_n\big(\zeta_{n,1}^{(\dot{u},1)}\big) - \mathbb{E}'_n\big(\zeta_{n,1}^{(\dot{u},1)}\big)\right) \\
&= 0 = c^{\mathrm{T}}\widehat{\widehat{U}}_n(0),
\end{aligned}$$

where we use Lemma 6.2.3.a. As $\bar{\zeta}_n = \sum_{l=1}^{n} \zeta_{n,l} = \sum_{l=1}^{n} Z_{n,\uparrow,l}(\omega_n)$ is non-random, we get that $\sum_{j=l}^{n} \zeta_{n,j}^{(u,l)}$ is $\mathcal{F}'_{n,(l-1)/n}$–$\mathbb{B}^{p\times n}$-measurable for all $l = 1, \ldots, n$ and $u = 1, \ldots, p$. Consequently, it holds that

$$\begin{aligned}
\mathbb{E}'_n\big[c^{\mathrm{T}}\widehat{\widehat{U}}_n(l/n) \mid \mathcal{F}'_{n,(l-1)/n}\big] &= c^{\mathrm{T}}\widehat{\widehat{U}}_n\Big(\frac{l-1}{n}\Big) \\
&\quad + \frac{\bar{\delta}_{n,l}}{\sqrt{n}} \sum_{u=1}^{r} c^{(u)} \bar{\gamma}_{n,l}^{(\dot{u},\ddot{u})} \left(\mathbb{E}'_n\big[\zeta_{n,l}^{(\dot{u},l)} \mid \mathcal{F}'_{n,(l-1)/n}\big] - \frac{\sum_{j=l}^{n} \zeta_{n,j}^{(\dot{u},l)}}{n-l+1}\right)
\end{aligned}$$

P'_n-almost surely, $l = 2, \ldots, n$. Using Lemma 6.2.3.c we receive that

$$\begin{aligned}
&\mathbb{E}'_n\big[\zeta_{n,l}^{(\dot{u},l)} \mid (\zeta_{n,1}, \ldots, \zeta_{n,l-1}) = (\bar{x}_{n,1}, \ldots, \bar{x}_{n,l-1})\big] \\
&\qquad = \sum_{j=1}^{m_n} \bar{z}_{n,j}^{(\dot{u},l)} \frac{\big(\nu_{n,j} - \sum_{k=1}^{l-1} \mathbb{1}(\bar{z}_{n,j} = \bar{x}_{n,k})\big)}{n-l+1}
\end{aligned}$$

P'_n-almost surely implying that

$$\mathbb{E}'_n\big[\zeta_{n,l}^{(\dot{u},l)} \mid \mathcal{F}'_{n,(l-1)/n}\big] = \sum_{j=1}^{m_n} \bar{z}_{n,j}^{(\dot{u},l)} \frac{\big(\nu_{n,j} - \sum_{k=1}^{l-1} \mathbb{1}(\bar{z}_{n,j} = \zeta_{n,k})\big)}{n-l+1} = \frac{\sum_{j=l}^{n} \zeta_{n,j}^{(\dot{u},l)}}{n-l+1}$$

P'_n-almost surely, $u = 1, \ldots, r$. $\square$

Before we can apply the Martingale Limit Theorem 2.1.2, we have to calculate some quantities.

6.2.5 Lemma. Under Assumption 6.2.1 let $\omega = (\omega_1, \omega_2, \dots) \in \Omega$ be fixed. Using the notation provided in Definition 6.2.2, let $\varepsilon \geq 0$ and consider the processes

$$J^{\varepsilon}[c^{\mathrm{T}}\widehat{\widehat{U}}_n] = \big\{J^{\varepsilon}[c^{\mathrm{T}}\widehat{\widehat{U}}_n](t) \mid t \in [0,1]\big\}, \quad J^{\varepsilon}[c^{\mathrm{T}}\widehat{\widehat{U}}_n](t) = \sum_{l=1}^{\lfloor nt \rfloor} W_{n,l}\, \mathbb{1}\big(|W_{n,l}| \geq \varepsilon\big),$$

where

$$W_{n,l} = \frac{\bar{\delta}_{n,l}}{\sqrt{n}} \sum_{u=1}^{r} c^{(u)}\, \bar{\gamma}_{n,l}^{(\dot{u},\ddot{u})} \left(\zeta_{n,l}^{(\dot{u},l)} - \frac{\sum_{j=l}^{n} \zeta_{n,j}^{(\dot{u},l)}}{n-l+1} \right), \qquad l = 1, \dots, n,$$

and

$$A_n^{\varepsilon}[c^{\mathrm{T}}\widehat{\widehat{U}}_n] = \big\{A_n^{\varepsilon}[c^{\mathrm{T}}\widehat{\widehat{U}}_n](t) \mid t \in [0,1]\big\}, \qquad A_n^{\varepsilon}[c^{\mathrm{T}}\widehat{\widehat{U}}_n](t) = \sum_{l=1}^{\lfloor nt \rfloor} A_{n,l}^{\varepsilon}$$

where

$$\begin{aligned} A_{n,l}^{\varepsilon} = \frac{\bar{\delta}_{n,l}}{\sqrt{n}} \sum_{j=1}^{m_n} &\left(\sum_{u=1}^{r} c^{(u)} \bar{\gamma}_{n,l}^{(\dot{u},\ddot{u})} \left(\bar{z}_{n,j}^{(\dot{u},l)} - \frac{\sum_{k=l}^{n} \zeta_{n,k}^{(\dot{u},l)}}{n-l+1} \right) \right) \\ &\times \mathbb{1}\left(\frac{\bar{\delta}_{n,l}}{\sqrt{n}} \left| \sum_{u=1}^{r} c^{(u)} \bar{\gamma}_{n,l}^{(\dot{u},\ddot{u})} \left(\bar{z}_{n,j}^{(\dot{u},l)} - \frac{\sum_{k=l}^{n} \zeta_{n,k}^{(\dot{u},l)}}{n-l+1} \right) \right| \geq \varepsilon \right) \\ &\times \frac{\nu_{n,j} - \sum_{k=1}^{l-1} \mathbb{1}(\bar{z}_{n,j} = \zeta_{n,k})}{n-l+1}, \end{aligned}$$

$l = 1, \dots, n$. Then the following assertions hold true.

a) The processes

$$M_{n,1}^{\varepsilon} = \big\{M_{n,1}^{\varepsilon}(t) \mid t \in [0,1]\big\}, \quad M_{n,1}^{\varepsilon}(t) = J_n^{\varepsilon}[c^{\mathrm{T}}\widehat{\widehat{U}}_n](t) - A_n^{\varepsilon}[c^{\mathrm{T}}\widehat{\widehat{U}}_n](t),$$

and

$$M_{n,2}^{\varepsilon} = \big\{M_{n,2}^{\varepsilon}(t) \mid t \in [0,1]\big\}, \quad M_{n,2}^{\varepsilon}(t) = c^{\mathrm{T}}\widehat{\widehat{U}}_n(t) - M_{n,1}^{\varepsilon}(t),$$

are $\mathbb{F}_n'$-martingales, where $\mathbb{F}_n'$ is defined in Proposition 6.2.4. Moreover, the process $A_n^{\varepsilon}[c^{\mathrm{T}}\widehat{\widehat{U}}_n]$ is predictable with respect to $\mathbb{F}_n'$.

b) Let $\langle M_{n,1}^{\varepsilon} \rangle$ denote the predictable quadratic variation of $M_{n,1}^{\varepsilon}$, which is given by

$$\langle M_{n,1}^{\varepsilon} \rangle(t) = \sum_{l=1}^{\lfloor nt \rfloor} \big(K_{n,l}^{\varepsilon} - \big(A_{n,l}^{\varepsilon}\big)^2 \big), \quad t \in [0,1],$$

where

$$K_{n,l}^{\varepsilon} = \frac{\bar{\delta}_{n,l}}{n} \sum_{j=1}^{m_n} \left(\sum_{u=1}^{r} c^{(u)}\ \bar{\gamma}_{n,l}^{(\dot{u},\ddot{u})} \left(\bar{z}_{n,j}^{(\dot{u},l)} - \frac{\sum_{k=l}^{n} \zeta_{n,k}^{(\dot{u},l)}}{n-l+1} \right) \right)^2$$
$$\times\ \mathbb{1}\left(\frac{\bar{\delta}_{n,l}}{\sqrt{n}} \left| \sum_{u=1}^{r} c^{(u)} \bar{\gamma}_{n,l}^{(\dot{u},\ddot{u})} \left(\bar{z}_{n,j}^{(u,l)} - \frac{\sum_{k=l}^{n} \zeta_{n,k}^{(\dot{u},l)}}{n-l+1} \right) \right| \geq \varepsilon \right)$$
$$\times\ \frac{\nu_{n,j} - \sum_{k=1}^{l-1} \mathbb{1}(\bar{z}_{n,j} = \zeta_{n,k})}{n-l+1},$$

$l = 1, \ldots, n$.

c) Let $\langle M_{n,1}^{\varepsilon}, M_{n,2}^{\varepsilon} \rangle$ denote the predictable covariation of $M_{n,1}^{\varepsilon} \cdot M_{n,2}^{\varepsilon}$ which satisfies P_n'-almost surely the equality

$$\langle M_{n,1}^{\varepsilon}, M_{n,2}^{\varepsilon} \rangle(t) = \sum_{l=1}^{\lfloor nt \rfloor} \big(A_{n,l}^{\varepsilon}\big)^2, \qquad t \in [0,1].$$

d) The predictable quadratic variation of the process $\{ c^{\mathrm{T}} \widehat{\widehat{U}}_n(t) \mid 0 \leq t \leq 1 \}$ is given by

$$\langle c^{\mathrm{T}} \widehat{\widehat{U}}_n \rangle(t) = \sum_{l=1}^{\lfloor nt \rfloor} K_{n,l}^{0}, \qquad t \in [0,1].$$

Proof. Observe that $\bar{\zeta}_n = \sum_{l=0}^{n} \zeta_{n,l} = \sum_{l=1}^{n} Z_{n,\uparrow,l}(\omega_n)$ is non-random and that all considered processes are basically processes in discrete time. As $\mathcal{F}_{n,0}' = \mathcal{Z}_n'$, it holds that either $P_n'(F) = 0$ or $P_n'(F) = 1$ for $F \in \mathcal{F}_{n,0}'$. Therefore, we have P_n'-almost surely the following chain of equalities

$$\mathbb{E}_n'\big[W_{n,1}\ \mathbb{1}\big(|W_{n,1}| \geq \varepsilon\big) \mid \mathcal{F}_{n,0}\big] = \mathbb{E}_n'\Big(W_{n,1}\ \mathbb{1}\big(|W_{n,1}| \geq \varepsilon\big)\Big)$$

and

$$\mathbb{E}'_n\big(W_{n,1}\ \mathbb{1}(|W_{n,1}| \geq \varepsilon)\big) = \frac{\bar{\delta}_{n,1}}{\sqrt{n}} \sum_{j=1}^{m_n} \left(\sum_{u=1}^{r} c^{(u)}\ \bar{\gamma}_{n,1}^{(\dot{u},\ddot{u})} \left(\bar{z}_{n,j}^{(\dot{u},1)} - \frac{\bar{\zeta}_n^{(\dot{u},1)}}{n} \right) \right)$$
$$\times\ \mathbb{1}\left(\frac{\bar{\delta}_{n,1}}{\sqrt{n}} \left| \sum_{u=1}^{r} c^{(u)} \bar{\gamma}_{n,1}^{(\dot{u},\ddot{u})} \left(\bar{z}_{n,j}^{(\dot{u},1)} - \frac{\bar{\zeta}_n^{(\dot{u},1)}}{n} \right) \right| \geq \varepsilon \right) \frac{\nu_{n,j}}{n}$$

where we use Lemma 6.2.3.a. Applying Lemma 6.2.3.c gives that

$$\begin{aligned}
&\mathbb{E}'_n\big[W_{n,l}\ \mathbb{1}(|W_{n,l}| \geq \varepsilon) \mid (\zeta_{n,1}, \dots, \zeta_{n,l-1}) = (\bar{x}_{n,1}, \dots, \bar{x}_{n,l-1})\big] \\
&\quad = \frac{\bar{\delta}_{n,l}}{\sqrt{n}} \sum_{j=1}^{m_n} \left(\sum_{u=1}^{r} c^{(u)} \bar{\gamma}_{n,l}^{(\dot{u},\ddot{u})} \left(\bar{z}_{n,j}^{(\dot{u},l)} - \frac{\bar{\zeta}_n^{(\dot{u},l)} - \sum_{k=1}^{l-1} \bar{x}_{n,k}^{(\dot{u},l)}}{n-l+1} \right) \right) \\
&\qquad \times\ \mathbb{1}\left(\frac{\bar{\delta}_{n,l}}{\sqrt{n}} \left| \sum_{u=1}^{r} c^{(u)} \bar{\gamma}_{n,l}^{(\dot{u},\ddot{u})} \left(\bar{z}_{n,j}^{(\dot{u},l)} - \frac{\bar{\zeta}_n^{(\dot{u},l)} - \sum_{k=1}^{l-1} \bar{x}_{n,k}^{(\dot{u},l)}}{n-l+1} \right) \right| \geq \varepsilon \right) \\
&\qquad \times\ \frac{\nu_{n,j} - \sum_{k=1}^{l-1} \mathbb{1}(\bar{z}_{n,j} = \bar{x}_{n,k})}{n-l+1},
\end{aligned}$$

$l = 2, \dots, n$, implying that

$$\mathbb{E}'_n\big[W_{n,l}\ \mathbb{1}(|W_{n,l}| \geq \varepsilon) \mid \mathcal{F}'_{n,(l-1)/n}\big] = A^{\varepsilon}_{n,l}\ \ P'_n\text{-almost surely}, \qquad l = 1, \dots, n.$$

This is the first part of assertion a). Proposition 6.2.4 gives the second part of the assertion. The predictability of $A^{\varepsilon}_n[c^{\mathrm{T}}\widehat{\widehat{U}}_n]$ is straightforward.

Proof of b). It holds that $\big(M^{\varepsilon}_{n,1}(t)\big)^2 = 0$, $t \in [0, 1/n)$, and

$$\begin{aligned}
\big(M^{\varepsilon}_{n,1}(t)\big)^2 &= \big(M^{\varepsilon}_{n,1}(t-1/n)\big)^2 + \Big(W_{n,\lfloor nt \rfloor}\ \mathbb{1}\big(|W_{n,\lfloor nt \rfloor}| \geq \varepsilon\big) - A^{\varepsilon}_{n,\lfloor nt \rfloor}\Big)^2 \\
&\quad + 2M^{\varepsilon}_{n,1}(t-1/n)\Big(W_{n,\lfloor nt \rfloor}\ \mathbb{1}\big(|W_{n,\lfloor nt \rfloor}| \geq \varepsilon\big) - A^{\varepsilon}_{n,\lfloor nt \rfloor}\Big),
\end{aligned}$$

$t \in [1/n, 1]$. We also have that

$$\mathbb{E}'_n\Big[M^{\varepsilon}_{n,1}(t-1/n)\Big(W_{n,\lfloor nt \rfloor}\ \mathbb{1}\big(|W_{n,\lfloor nt \rfloor}| \geq \varepsilon\big) - A^{\varepsilon}_{n,\lfloor nt \rfloor}\Big)\ \Big|\ \mathcal{F}'_{n,(\lfloor nt \rfloor-1)/n}\Big] = 0$$

P'_n-almost surely, where we use that $M^{\varepsilon}_{n,1}(t-1/n)$ is almost surely bounded and $\mathcal{F}'_{n,(\lfloor nt \rfloor-1)/n}$–$\mathbb{B}$ measurable. Furthermore, using the $\mathcal{F}'_{n,(\lfloor nt \rfloor-1)/n}$–$\mathbb{B}$ measurability of $A^{\varepsilon}_{n,\lfloor nt\rfloor}$, it holds that

$$\begin{aligned}\mathbb{E}'_n\Big[\Big(W_{n,\lfloor nt\rfloor}\,\mathbb{1}\big(|W_{n,\lfloor nt\rfloor}|\geq\varepsilon\big)-A^{\varepsilon}_{n,\lfloor nt\rfloor}\Big)^2,\ \Big|\ \mathcal{F}'_{n,(\lfloor nt\rfloor-1)/n}\Big]=\\ \mathbb{E}'_n\Big[\Big(W_{n,\lfloor nt\rfloor}\Big)^2\,\mathbb{1}\big(|W_{n,\lfloor nt\rfloor}|\geq\varepsilon\big)\ \Big|\ \mathcal{F}'_{n,(\lfloor nt\rfloor-1)/n}\Big]-\Big(A^{\varepsilon}_{n,\lfloor nt\rfloor}\Big)^2,\end{aligned}$$

Analog to the previous considerations we get that

$$\mathbb{E}'_n\big[W^2_{n,1},\mathbb{1}\big(|W_{n,1}|\geq\varepsilon\big)\mid\mathcal{F}'_{n,0}\big]=\mathbb{E}'_n\Big(W^2_{n,1},\mathbb{1}\big(|W_{n,1}\geq\varepsilon|\big)\Big)$$

and

$$\begin{aligned}\mathbb{E}'_n\Big(W^2_{n,1},\mathbb{1}\big(|W_{n,1}\geq\varepsilon|\big)\Big)&=\frac{\bar{\delta}_{n,1}}{n}\sum_{j=1}^{m_n}\left(\sum_{u=1}^{r}c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,1}\left(\bar{z}^{(\dot{u},1)}_{n,j}-\frac{\bar{\zeta}^{(\dot{u},i)}_n}{n}\right)\right)^2\\&\quad\times\ \mathbb{1}\left(\frac{\bar{\delta}_{n,1}}{\sqrt{n}}\left|\sum_{u=1}^{r}c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,1}\left(\bar{z}^{(\dot{u},1)}_{n,j}-\frac{\bar{\zeta}^{(\dot{u},1)}_n}{n}\right)\right|\geq\varepsilon\right)\frac{\nu_{n,j}}{n}\end{aligned}$$

as well as

$$\begin{aligned}&\mathbb{E}'_n\big[W^2_{n,l}\ \mathbb{1}(|W_{n,l}|\geq\varepsilon)\mid(\zeta_{n,1},\ldots,\zeta_{n,l-1})=(\bar{x}_{n,1},\ldots,\bar{x}_{n,l-1})\big]\\&\quad=\frac{\bar{\delta}_{n,l}}{n}\sum_{j=1}^{m_n}\left(\sum_{u=1}^{r}c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,l}\left(\bar{z}^{(\dot{u},l)}_{n,j}-\frac{\bar{\zeta}^{(\dot{u},l)}_n-\sum_{k=1}^{l-1}\bar{x}^{(\dot{u},l)}_{n,k}}{n-l+1}\right)\right)^2\\&\qquad\times\ \mathbb{1}\left(\frac{\bar{\delta}_{n,l}}{\sqrt{n}}\left|\sum_{u=1}^{r}c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,l}\left(\bar{z}^{(\dot{u},l)}_{n,j}-\frac{\bar{\zeta}^{(\dot{u},l)}_n-\sum_{k=1}^{l-1}\bar{x}^{(\dot{u},l)}_{n,k}}{n-l+1}\right)\right|\geq\varepsilon\right)\\&\qquad\times\ \frac{\nu_{n,j}-\sum_{k=1}^{l-1}\mathbb{1}(\bar{z}_{n,j}=\bar{x}_{n,k})}{n-l+1},\end{aligned}$$

$l=2,\ldots,n$. On the whole, we showed that

$$\mathbb{E}'_n\big[W^2_{n,l}\ \mathbb{1}\big(|W_{n,l}|\geq\varepsilon\big)\mid\mathcal{F}'_{n,(l-1)/n}\big]=K^{\varepsilon}_{n,l}\quad P'_n\text{-almost surely.}$$

Proof of c). Note that

$$M^{\varepsilon}_{n,2}(t)=\sum_{l=1}^{\lfloor nt\rfloor}\Big(W_{n,l}\mathbb{1}\big(|W_{n,l}|<\varepsilon\big)+A^{\varepsilon}_{n,l}\Big).$$

Moreover, it holds that $M^{\varepsilon}_{n,1}\, M^{\varepsilon}_{n,2}(t) = 0$, $t \in [0, 1/n)$, and

$$\begin{aligned}
&\mathbb{E}'_n\big[M^{\varepsilon}_{n,1}(t)\, M^{\varepsilon}_{n,2}(t) \mid \mathcal{F}'_{n,(\lfloor nt\rfloor-1)}/n\big] = M^{\varepsilon}_{n,1}(t-1/n)\, M^{\varepsilon}_{n,2}(t-1/n)\\
&\quad + M^{\varepsilon}_{n,1}(t-1/n)\, \underbrace{\mathbb{E}'_n\Big[W_{n,\lfloor nt\rfloor}\, \mathbb{1}\big(|W_{n,\lfloor nt\rfloor}| < \varepsilon\big) + A^{\varepsilon}_{n,\lfloor nt\rfloor} \;\Big|\; \mathcal{F}'_{n,(\lfloor nt\rfloor-1)/n}\Big]}_{=0\ P'_n\text{-almost surely}}\\
&\quad + M^{\varepsilon}_{n,2}(t-1/n)\, \underbrace{\mathbb{E}'_n\Big[W_{n,\lfloor nt\rfloor}\, \mathbb{1}\big(|W_{n,\lfloor nt\rfloor}| \geq \varepsilon\big) - A^{\varepsilon}_{n,\lfloor nt\rfloor} \;\Big|\; \mathcal{F}'_{n,(\lfloor nt\rfloor-1)/n}\Big]}_{=0\ P'_n\text{-almost surely}}\\
&\quad + \mathbb{E}'_n\big[C_n(t) \mid \mathcal{F}'_{n,(\lfloor nt\rfloor-1)/n}\big]
\end{aligned}$$

$t \in [1/n, 1]$, where $C_n(t)$ is given by

$$\Big(W_{n,\lfloor nt\rfloor}\, \mathbb{1}\big(|W_{n,\lfloor nt\rfloor}| < \varepsilon\big) + A^{\varepsilon}_{n,\lfloor nt\rfloor}\Big)\Big(W_{n,\lfloor nt\rfloor}\, \mathbb{1}\big(|W_{n,\lfloor nt\rfloor}| \geq \varepsilon\big) - A^{\varepsilon}_{n,\lfloor nt\rfloor}\Big)$$

and a) is applied. Using the previous considerations and observing that

$$W^2_{n,\lfloor nt\rfloor}\, \mathbb{1}\big(|W_{n,\lfloor nt\rfloor}| < \varepsilon\big)\mathbb{1}\big(|W_{n,\lfloor nt\rfloor}| \geq \varepsilon\big) = 0 \quad P'_n\text{-almost surely},$$

one easily sees that $\mathbb{E}'_n\big[C_n(t) \mid \mathcal{F}'_{n,(\lfloor nt\rfloor-1)/n}\big] = \big(A^{\varepsilon}_{n,\lfloor nt\rfloor}\big)^2$ P'_n-almost surely.

The assertion of d) is an immediate consequence of the fact that $c^{\mathrm{T}}\widehat{\widehat{U}}_n = M^0_{n,1}$ and that $A^0_{n,l} = 0$, $l = 1, \ldots, n$, as an easy calculation shows. □

Now, we state the assumptions needed to prove the convergence in distribution of the statistic $\widehat{\widehat{U}}(t)$ for fixed $\omega \in \Omega$

6.2.6 Assumption. i) The functions $\bar{\mu}^{(u)}_1 : [0,1] \longrightarrow \mathbb{R}$, $u = 1, \ldots, p$, and $\bar{\mu}^{(u,v)}_2 : [0,1] \longrightarrow \mathbb{R}$, $u, v = 1, \ldots, p$, are measurable functions that are bounded on every interval $[0,t]$.

ii) $\sup_{s\in[0,t]}\big|\widehat{\widehat{\mu}}^{(u)}_{k_n,1}(s,\omega) - (1-s)\,\bar{\mu}^{(u)}_1(s)\big| \longrightarrow_{P'_{k_n}} 0$, as $n \to \infty$, $u = 1, \ldots, p$, for all $t \in [0,1)$.

iii) $\sup_{s\in[0,t]}\big|\widehat{\widehat{\mu}}^{(u,v)}_{k_n,2}(s,\omega) - (1-s)\,\bar{\mu}^{(u,v)}_2(s)\big| \longrightarrow_{P'_{k_n}} 0$, as $n \to \infty$, $u = 1, \ldots, p$, $v = 1, \ldots, p$, for all $t \in [0,1)$.

iv) $\lim_{n\to\infty}\int_{[0,t]}\big(\widehat{\gamma}^{(\dot{u},\ddot{u})}_{k_n:\lfloor k_n s\rfloor}(\omega_{k_n}) - \bar{\gamma}^{(\dot{u},\ddot{u})}(s)\big)^2\,\mathrm{d}s = 0$, where $\bar{\gamma}^{(\dot{u},\ddot{u})}$ is some square integrable function, $u = 1, \ldots, r$, for all $t \in [0,1)$.

v) $\int_{[0,t]} h_{k_n}(s,\omega)\,\mathrm{d}s \longrightarrow \bar{H}(t)$, as $n \to \infty$, for all $t \in [0,1]$, where $\bar{H}$ is a monotone non-decreasing function with $\bar{H}(0) = 0$.

vi) $\lim_{n\to\infty} \frac{1}{\sqrt{k_n}} \max_{1\le i\le k_n} \max\{|Z^{(u,v)}_{k_n,\uparrow,i}| \mid u = 1,\dots,p,\ v = 1,\dots,k_n\} = 0$.

Later we choose the functions $\bar{\mu}_1^{(u)}$, $\bar{\mu}_2^{(u,v)}$, $u,v = 1,\dots,p$, as well as $\bar{\gamma}^{(\dot{u},\ddot{u})}$, $u = 1,\dots,r$, and $\bar{H}(t)$ independently of $\omega \in \Omega$. Moreover, we need that the measure defined by $\bar{H}$ has a Lebesgue density.

6.2.7 Lemma. Under Assumption 6.2.1, let $\omega = (\omega_1,\omega_2,\dots) \in \Omega$ be fixed and use the notation provided in Definition 6.2.2. Suppose that Assumption 6.2.6.v holds for fixed $\omega \in \Omega$, where k_n, $n \in \mathbb{N}$, is some sub-sequence of natural numbers. Then the function $\bar{H} : [0,1] \to \mathbb{R}$ defines a Lebesgue-continuous measure, *i.e.* $\int_{[0,t]} 1\,\mathrm{d}\bar{H}(s) = \int_{[0,t]} h(s)\,\mathrm{d}s$. Additionally, one can choose a version of the density h, such that $0 \le h \le 1$.

Proof. By construction it holds that

$$\underbrace{\left|\frac{\int_{[0,t]} \bar{\delta}_{k_n,\lfloor k_n s\rfloor}\,\mathrm{d}s - \int_{[0,u]} \bar{\delta}_{k_n,\lfloor k_n s\rfloor}\,\mathrm{d}s}{t-u}\right|}_{\le 1} \longrightarrow \underbrace{\left|\frac{\bar{H}(t) - \bar{H}(u)}{t-u}\right|}_{\le 1}, \qquad \text{as } n \to \infty,$$

for all $t,u \in [0,1]$, which implies that the function $\bar{H}$ is Lipschitz-continuous, *i.e.* $\big|\bar{H}(t) - \bar{H}(u)\big| \le \big|t-u\big|$. Consequently, for $\varepsilon > 0$ and all $(a_i,b_i] \subset [0,1]$, $i = 1,\dots,k$, such that $\sum_{i=1}^k (b_i - a_i) \le \varepsilon$ we have $\sum_{i=1}^k \big|\bar{H}(b_i) - \bar{H}(a_i)\big| \le \varepsilon$. This means that the function $\bar{H}$ is absolute continuous on the interval $[0,1]$. Witting and Müller-Funk [72, Theorem B1.21] yield the existence of a Lebesgue-density h. Because $0 \le h(s) = \lim_{t\to s} \frac{\bar{H}(s)-\bar{H}(t)}{s-t} \le 1$ for almost all $s \in [0,1]$ the last assertion holds true. □

After these preparatory efforts, we can state a central limit theorem for the sequence of statistics $\hat{\hat{U}}_n(t)$, $n \in \mathbb{N}$, where $\omega \in \Omega$ is fixed.

6.2.8 Theorem. Suppose that Assumption 6.2.1 holds, let k_n, $n \in \mathbb{N}$, be some sub-sequence of natural numbers and let $\omega = (\omega_1,\omega_2,\dots) \in \Omega$ be fixed.

Using the notation provided in Definition 6.2.2 assume that Assumption 6.2.6.i – Assumption 6.2.6.vi hold for ω and k_n, $n \in \mathbb{N}$. Furthermore, let h denote a Lebesgue-density of $\bar{H}$, cf. Lemma 6.2.7. Then it holds that

$$\widehat{\widehat{U}}_{k_n}(t) \xrightarrow{\mathfrak{D}}_{P'_{k_n}} \mathcal{N}\big(0, \bar{\mathscr{I}}(t)\big), \quad \text{where } \bar{\mathscr{I}}(t) = \big(\bar{\mathscr{I}}^{(u,v)}(t)\big) \in \mathbb{R}^{r \times r}$$

and

$$\bar{\mathscr{I}}^{(u,v)}(t) = \int_{[0,t]} \bar{\gamma}^{(\dot{u},\ddot{u})}(s)\, \bar{\gamma}^{(\dot{v},\ddot{v})}(s)\, \big(\bar{\mu}_2^{(\dot{u},\dot{v})}(s) - \bar{\mu}_1^{(\dot{u})}(s)\, \bar{\mu}_1^{(\dot{u})}(s)\big)\, h(s)\, \mathrm{d}s$$

$u, v = 1, \ldots, r$, for all $t \in [0,1)$. Moreover, one gets that $\widehat{\widehat{V}}_{k_n}(t) - \bar{\mathscr{I}}(t) \longrightarrow_{P'_n} 0$, as $n \to \infty$.

Proof. Without loss of generality, we can suppose that $k_n = n$, $n \in \mathbb{N}$. We intend to apply Rebolledo's Central Limit Theorem. Therefore, we check the conditions stated in Theorem 2.1.2 in the following paragraphs.

As a first step we want to show that $\big\langle M_{n,1}^{\varepsilon}\big\rangle(t) \longrightarrow_{P'_n} 0$, as $n \to \infty$, for all $t \in [0,1)$. Let us introduce some more abbreviations

$$\begin{aligned}
w_{n,j}(s) &= \sum_{u=1}^{r} c^{(u)} \bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{u},\ddot{u})} \Bigg(\bar{z}_{n,j}^{(\dot{u},\lfloor ns \rfloor)} - \frac{\sum_{k=\lfloor ns \rfloor}^{n} \zeta_{n,k}^{(\dot{u},\lfloor ns \rfloor)}}{n - \lfloor ns \rfloor + 1} \Bigg), && j = 1, \ldots, m_n, \\
p_{n,j}(s) &= \frac{\nu_{n,j} - \sum_{k=1}^{\lfloor ns \rfloor - 1} \mathbb{1}\big(\bar{z}_{n,j} = \zeta_{n,k}\big)}{n - \lfloor ns \rfloor + 1}, && j = 1, \ldots, m_n, \\
\phi_n(s) &= \sum_{j=1}^{m_n} w_{n,j}^2(s)\, \mathbb{1}\bigg(\frac{h_n(s)}{\sqrt{n}} \big|w_{n,j}(s)\big| \geq \varepsilon \bigg) p_{n,j}(s), \\
\psi_n(s) &= \sum_{j=1}^{m_n} w_{n,j}(s)\, \mathbb{1}\bigg(\frac{h_n(s)}{\sqrt{n}} \big|w_{n,j}(s)\big| \geq \varepsilon \bigg) p_{n,j}(s),
\end{aligned}$$

where $s \in [0,1]$ and we remember that $\bar{\delta}_{n,0} = 0$, $\bar{\gamma}_{n,0}^{(\dot{u},\ddot{u})} = 0$, $u = 1, \ldots, r$, $n \in \mathbb{N}$, and $\zeta_{n,0}^{(\dot{u},l)} = 0$, $u = 1, \ldots, r$, $l = 1, \ldots, n$, $n \in \mathbb{N}$. It holds that

$$\big\langle M_{n,1}^{\varepsilon}\big\rangle(t) = \int_{[0,\lfloor nt+1 \rfloor/n]} \phi_n(s)\, h_n(s)\, \mathrm{d}s - \int_{[0,\lfloor nt+1 \rfloor/n]} \psi_n^2(s)\, h_n(s)\, \mathrm{d}s.$$

Choose $t^* \in (t,1)$. For all sufficiently large $n \in \mathbb{N}$ the estimates

$$0 \le \int_{[0,\lfloor nt+1\rfloor/n]} \phi_n(s)\, h_n(s)\,\mathrm{d}s \le \int_{[0,t^*]} \phi_n(s)\, h_n(s)\,\mathrm{d}s$$

and

$$0 \le \int_{[0,\lfloor nt+1\rfloor/n]} \psi_n^2(s)\, h_n(s)\,\mathrm{d}s \le \int_{[0,t^*]} \psi_n^2(s)\, h_n(s)\,\mathrm{d}s$$

hold. Therefore, it suffices to show that

$$\int_{[0,t^*]} \phi_n(s)\, h_n(s)\,\mathrm{d}s \longrightarrow_{P_n'} 0 \quad \text{and} \quad \int_{[0,t^*]} \psi_n^2(s)\, h_n(s)\,\mathrm{d}s \longrightarrow_{P_n'} 0,$$

as $n \to \infty$, for all $t^* \in [0,1)$.

Using Assumption 6.2.6.i, one can choose $C_{t^*} \in \mathbb{R}_+$, such that $\left|\bar{\mu}_1^{(u)}(s)\right| \le C_{t^*}$ and $\left|\bar{\mu}_2^{(u,v)}(s)\right| \le C_{t^*}$ for all $s \in [0,t^*]$, $u,v = 1,\ldots,p$. Moreover, note that $0 \le h_n \le 1$. It holds the estimate

$$\begin{aligned} 0 \le \phi_n(s)\, h_n(s) &\le \mathbb{1}\left(\frac{1}{\sqrt{n}} \max_{1\le i\le m_n} \left|w_{n,i}\right| \ge \varepsilon\right) r \sum_{u=1}^{r} \left(c^{(u)} \bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{u},\ddot{u})}\right)^2 \\ &\quad \times \sum_{i=1}^{m_n} \left(\bar{z}_{n,i}^{(\dot{u},\lfloor ns\rfloor)} - \frac{\sum_{j=\lfloor ns\rfloor}^{n} \zeta_{n,j}^{(\dot{u},\lfloor ns\rfloor)}}{n-\lfloor ns\rfloor+1}\right)^2 p_{n,i}(s) \\ &= \mathbb{1}\left(\frac{1}{\sqrt{n}} \max_{1\le i\le m_n} \left|w_{n,i}\right| \ge \varepsilon\right) r \sum_{u=1}^{r} \left(c^{(u)} \bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{u},\ddot{u})}\right)^2 \\ &\quad \times \left(\frac{\sum_{i=\lfloor ns\rfloor}^{n} \left(\zeta_{n,i}^{(\dot{u},\lfloor ns\rfloor)}\right)^2}{n-\lfloor ns\rfloor+1} - \left(\frac{\sum_{i=\lfloor ns\rfloor}^{n} \zeta_{n,i}^{(\dot{u},\lfloor ns\rfloor)}}{n-\lfloor ns\rfloor+1}\right)^2\right) \end{aligned} \tag{6.3}$$

for all $s \in [0,t^*]$, where we use the estimate $\left(\sum_{i=1}^r a_i\right)^2 \le r\sum_{i=1}^r a_i^2$. Obviously, it holds that

$$\begin{aligned} \widehat{\bar{\mu}}_{n,2}^{(u,u)}(s) &\le \sup_{0\le s\le t^*} \left|\widehat{\bar{\mu}}_{n,2}^{(u,u)}(s) - (1-s)\,\bar{\mu}_2^{(u,u)}(s)\right| + C_{t^*} \\ &= I_{n,1}^{(u)} + C_{t^*} \end{aligned}$$

and

$$\begin{aligned} \left|\widehat{\bar{\mu}}_{n,1}^{(u)}(s)\right|^2 &\le 2 \sup_{0\le s\le t^*} \left|\widehat{\bar{\mu}}_{n,1}^{(u)}(s) - (1-s)\,\bar{\mu}_1^{(u)}(s)\right|^2 + 2C_{t^*}^2 \\ &= I_{n,2}^{(u)} + 2C_{t^*}^2, \end{aligned}$$

$u = 1, \ldots, p$. Exploiting the estimates

$$\frac{n - \lfloor ns \rfloor + 1}{n} \geq 1 - \frac{ns}{n} + \frac{1}{n} \geq 1 - s \geq 1 - t^*$$

and expanding the fractions with $1/n$, we get that

$$\begin{aligned} &\left| \frac{\sum_{i=\lfloor ns \rfloor}^{n} \big(\zeta_{n,i}^{(u,\lfloor ns \rfloor)}\big)^2}{n - \lfloor ns \rfloor + 1} - \left(\frac{\sum_{i=\lfloor ns \rfloor}^{n} \zeta_{n,i}^{(u,\lfloor ns \rfloor)}}{n - \lfloor ns \rfloor + 1} \right)^2 \right| \\ &\qquad\qquad \leq (1 - t^*)^{-2} \Big(\widehat{\mu}_{n,2}^{(u,u)}(s) + \big(\widehat{\mu}_{n,1}^{(u)}(s)\big)^2 \Big) \leq I_n^{(u)} + \widetilde{C}_{t^*}, \end{aligned}$$

$u = 1, \ldots, p$, where

$$I_n^{(u)} = (1 - t^*)^{-2} \big(I_{n,1}^{(u)} + I_{n,2}^{(u)} \big) \quad \text{and} \quad \widetilde{C}_{t^*} = (1 - t^*)^{-2} \big(C_{t^*} + 2C_{t^*}^2 \big).$$

The estimate (6.3) implies that

$$\begin{aligned} \int_{[0,t^*]} \phi_n(s) h_n(s) \, \mathrm{d}s &\leq r \sum_{u=1}^{r} I_n^{(\dot{u})} \int_{[0,t^*]} \big(c^{(u)} \bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{u},\ddot{u})} \big)^2 \, \mathrm{d}s \\ &\quad + 2r \sum_{u=1}^{r} \widetilde{C}_{t^*} \int_{[0,t^*]} \big(c^{(u)} \big)^2 \big(\bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{u},\ddot{u})} - \bar{\gamma}^{(\dot{u},\ddot{u})}(s) \big)^2 \, \mathrm{d}s \\ &\quad + 2r \sum_{u=1}^{r} \widetilde{C}_{t^*} \int_{[0,t^*]} \mathbb{1} \left(\frac{1}{\sqrt{n}} \max_{1 \leq i \leq m_n} \big| w_{n,i} \big| \geq \varepsilon \right) \big(c^{(u)} \bar{\gamma}^{(\dot{u},\ddot{u})}(s) \big)^2 \, \mathrm{d}s. \end{aligned} \tag{6.4}$$

Because of Assumption 6.2.6.iv and Vitali's Theorem, see *e.g.* Witting [71, Satz 1.181], we get that

$$\int_{[0,t^*]} \big(c^{(u)} \bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{u},\ddot{u})} \big)^2 \, \mathrm{d}s \longrightarrow \int_{[0,t^*]} \big(c^{(u)} \bar{\gamma}^{(\dot{u},\ddot{u})}(s) \big)^2 \, \mathrm{d}s, \quad \text{as } n \to \infty, \tag{6.5}$$

implying

$$r \sum_{u=1}^{r} I_n^{(\dot{u})} \int_{[0,t^*]} \big(c^{(u)} \bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{u},\ddot{u})} \big)^2 \, \mathrm{d}s \longrightarrow_{P_n'} 0, \quad \text{as } n \to \infty,$$

where we also use Assumption 6.2.6.ii and Assumption 6.2.6.iii. Clearly, it holds that

$$\sum_{u=1}^{r} \widetilde{C}_{t^*} \int_{[0,t^*]} \big(c^{(u)} \big)^2 \big(\bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{u},\ddot{u})} - \bar{\gamma}^{(\dot{u},\ddot{u})}(s) \big)^2 \, \mathrm{d}s \longrightarrow_{P_n'} 0,$$

because of Assumption 6.2.6.iv. Finally, we show that the third summand of the right hand side of (6.4) converges to 0 in P_n'-probability. For this we show that $\mathbb{1}\big(\frac{1}{\sqrt{n}}\max_{1\le i\le m_n}|w_{n,i}|\ge\varepsilon\big)\longrightarrow_{\mu_{t^*}} 0$,where $\mu_{t^*}(B)=\frac{1}{t^*}\int_B 1\,\mathrm{d}s$, $B\in\mathbb{B}[0,t^*]$, as a first step. It holds the estimate

$$\begin{aligned}\frac{1}{\sqrt{n}}\max_{1\le i\le m_n}\big|w_{n,i}\big| \le{}& \left(\sum_{u=1}^{r}\big|c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,\lfloor ns\rfloor}\big|\right)\cdot\frac{1}{\sqrt{n}}\max_{1\le i\le m_n}\max_{1\le u\le p}\max_{1\le v\le n}\big|\bar{z}^{(u,v)}_{n,i}\big| \\ &+\frac{1}{\sqrt{n}(1-s)}\sum_{u=1}^{r}\big|c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,\lfloor ns\rfloor}\big|\sup_{v\in[0,t]}\Big|\widehat{\bar{\mu}}^{(\dot{u})}_{n,1}(v)-(1-v)\bar{\mu}^{(\dot{u})}_{1}(v)\Big| \\ &+\frac{1}{\sqrt{n}}\sum_{u=1}^{r}\big|c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,\lfloor ns\rfloor}\big|\cdot C_{t^*}.\end{aligned}$$

implying the estimate

$$\begin{aligned}\frac{1}{\sqrt{n}}\max_{1\le i\le m_n}\big|w_{n,i}\big| \le{}& \left(\sum_{u=1}^{r}\big|c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,\lfloor ns\rfloor}\big|\right)\cdot\frac{1}{\sqrt{n}}\max_{1\le i\le m_n}\max_{1\le u\le p}\max_{1\le v\le n}\big|\bar{z}^{(u,v)}_{n,i}\big| \\ +\frac{1}{\sqrt{n}(1-t^*)}&\max_{1\le u\le r}\max_{1\le i\le\lfloor nt^*\rfloor+1}\big|c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,i}\big|\sup_{v\in[0,t]}\Big|\widehat{\bar{\mu}}^{(\dot{u})}_{n,1}(v)-(1-v)\bar{\mu}^{(\dot{u})}_{1}(v)\Big| \\ +C_{t^*}\cdot\frac{1}{\sqrt{n}}&\max_{1\le u\le r}\max_{1\le i\le\lfloor nt^*\rfloor+1}\big|c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,i}\big|.\end{aligned}$$

Again, Assumption 6.2.6.iv and Vitali's Theorem imply that $c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,\lfloor n\cdot\rfloor}\longrightarrow c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}(\cdot)$ in μ_{t^*}-probability. As $c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}(\cdot)$ are integrable with respect to μ_{t^*}, cf. Assumption 6.2.6.iv, we can immediately conclude that

$$\left(\sum_{u=1}^{r}\big|c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,\lfloor ns\rfloor}\big|\right)\cdot\frac{1}{\sqrt{n}}\max_{1\le i\le m_n}\max_{1\le u\le p}\max_{1\le v\le n}\big|\bar{z}^{(u,v)}_{n,i}\big|\longrightarrow_{\mu_{t^*}} 0,\quad\text{as } n\to\infty,$$

where we use Assumption 6.2.6.vi. Assumption 6.2.6.iv and Neuhaus [57, Proof of Theorem 5.2] give that

$$\frac{1}{\sqrt{n}}\max_{1\le u\le r}\max_{1\le i\le\lfloor nt^*\rfloor+1}\big|c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,i}\big|\longrightarrow 0,\quad\text{as } n\to\infty,\qquad u=1,\dots,r.$$

Thus, it holds that

$$\frac{1}{\sqrt{n}}\max_{1\le i\le m_n}\big|w_{n,i}\big|\longrightarrow_{\mu_{t^*}} 0,\qquad\text{as } n\to\infty.$$

Finally, Lebesgue's Theorem yields that

$$\sum_{u=1}^{r} \widetilde{C}_t \int_{[0,t]} \mathbb{1}\left(\frac{1}{\sqrt{n}} \max_{1\le i\le m_n} \left|w_{n,i}\right| \ge \varepsilon \right) \left(c^{(u)}\bar{\gamma}^{(\dot{u},\ddot{u})}(s)\right)^2 \mathrm{d}s \longrightarrow_{P_n'} 0.$$

On the whole we showed that $\int_{[0,t^*]} \phi_n(s)h_n(s)\,\mathrm{d}s \longrightarrow_{P_n'} 0$, as $n\to\infty$. Applying the Cauchy-Schwarz inequality yields that $\left(\psi_n(s)\right)^2 \le \phi_n(s)$. Therefore, we get that $\int_{[0,t^*]} \left(\psi_n(s)\right)^2 h_n(s)\,\mathrm{d}s \longrightarrow_{P_n'} 0$, as $n\to\infty$. Recapitulating, we proved that $\left\langle M_{n,1}^{\varepsilon}\right\rangle(t) \longrightarrow_{P_n'} 0$ for all $t\in[0,1)$. Fortunately, it holds that

$$\sup_{0\le s\le t}\left|\left\langle M_{n,1}^{\varepsilon}, M_{n,2}^{\varepsilon}\right\rangle(s)\right| = \int_{[0,\lfloor nt+1\rfloor/n]} \left(\psi_n(s)\right)^2 h_n(s)\,\mathrm{d}s \longrightarrow_{P_n'} 0,$$

for all $t\in[0,1)$, which is an easy consequence of Lemma 6.2.5.c and the previous calculations.

In the next step we show that $\left\langle c^{\mathrm{T}}\widehat{\widehat{U}}_n\right\rangle(t) - A_c(t) \longrightarrow_{P_n} 0$, for all $t\in[0,1)$, where

$$A_c(t) = \sum_{u=1}^{r}\sum_{v=1}^{r} c^{(u)}\, c^{(v)}\ \mathscr{J}^{(u,v)}(t). \tag{6.6}$$

Lemma 6.2.5.d gives that

$$\begin{aligned} \left\langle c^{\mathrm{T}}\widehat{\widehat{U}}_n\right\rangle(t) &= \int_{[0,\lfloor nt+1\rfloor/n]} \left(\sum\nolimits_{i=1}^{m_n} w_{n,i}^2(s)\, p_{n,i}(s)\right) h_n(s)\,\mathrm{d}s \\ &= \sum_{u=1}^{r}\sum_{v=1}^{r} c^{(u)}c^{(v)}\ \widehat{\widehat{V}}_n^{(u,v)}(\lfloor nt+1\rfloor/n). \end{aligned} \tag{6.7}$$

As $\sum_{u=1}^{r}\sum_{v=1}^{r} c^{(u)}c^{(v)}\ \widehat{\widehat{V}}_n^{(u,v)}(t)$ is non-decreasing in t and $t\mapsto A_c(t)$ is continuous, if suffices to show that

$$\sum_{u=1}^{r}\sum_{v=1}^{r} c^{(u)}c^{(v)}\ \widehat{\widehat{V}}_n^{(u,v)}(t) - A_c(t) \longrightarrow_{P_n'} 0, \quad \text{as } n\to\infty,$$

for all $t\in[0,1)$. Using Assumption 6.2.6.i, one can choose $C_t\in\mathbb{R}_+$, such that $\left|\bar{\mu}_1^{(u)}(s)\right| \le C_t$ and $\left|\bar{\mu}_2^{(u,v)}(s)\right| \le C_t$ for all $s\in[0,t]$, $u,v=1,\dots,p$. As a first step we show that $\widehat{\widehat{V}}_n^{(u,v)}(t) - \widetilde{\widetilde{V}}_n^{(u,v)}(t) \longrightarrow_{P_n} 0$, as $n\to\infty$, where

$$\widetilde{\widetilde{V}}_n^{(u,v)}(t) = \int_{[0,t]} \bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{u},\ddot{u})}\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{v},\ddot{v})}\left(\bar{\mu}_2^{(\dot{u},\dot{v})}(s) - \bar{\mu}_1^{(\dot{u})}(s)\,\bar{\mu}_1^{(\dot{v})}(s)\right) h_n(s)\,\mathrm{d}s,$$

$u, v = 1, \ldots, r$. Some easy calculations yield that

$$\left| \left(1 - \frac{\lfloor ns \rfloor}{n} + \frac{1}{n}\right)^{-1} - (1-s)^{-1} \right| \le \frac{2}{n(1-t)^2}$$

and

$$\begin{aligned}
&\left| \left(1 - \frac{\lfloor ns \rfloor}{n} + \frac{1}{n}\right)^{-2} - (1-s)^{-2} \right| \\
&\quad = \left| \left(1 - \frac{\lfloor ns \rfloor}{n} + \frac{1}{n}\right)^{-1} - (1-s)^{-1} \right| \cdot \left| \left(1 - \frac{\lfloor ns \rfloor}{n} + \frac{1}{n}\right)^{-1} + (1-s)^{-1} \right| \\
&\quad \le \frac{4}{n(1-t)^3}
\end{aligned}$$

for all $0 \le s \le t$. Using these estimates and abbreviating

$$R_{n,1}^{(u,v)}(s) = \frac{\widehat{\bar{\mu}}_{n,2}^{(u,v)}(s)}{1 - \frac{\lfloor ns \rfloor}{n} + \frac{1}{n}} - \frac{\widehat{\bar{\mu}}_{n,2}^{(u,v)}(s)}{1-s} + \frac{\widehat{\bar{\mu}}_{n,1}^{(u)}(s)\widehat{\bar{\mu}}_{n,1}^{(v)}(s)}{\left(1-s\right)^2} - \frac{\widehat{\bar{\mu}}_{n,1}^{(u)}(s)\widehat{\bar{\mu}}_{n,1}^{(v)}(s)}{\left(1 - \frac{\lfloor ns \rfloor}{n} + \frac{1}{n}\right)^2},$$

$u, v = 1, \ldots, p$, one gets that

$$\left| R_{n,1}^{(u,v)}(s) \right| \le \frac{4}{n(1-t)^3} \Big(\left| \widehat{\bar{\mu}}_{n,2}^{(u,v)}(s) \right| + \left| \widehat{\bar{\mu}}_{n,1}^{(u)}(s) \widehat{\bar{\mu}}_{n,1}^{(v)}(s) \right| \Big).$$

Therefore, we have that

$$\begin{aligned}
&\left| \int_{[0,t]} \bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{u},\ddot{v})} \bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{v},\ddot{v})} R_{n,1}^{(\dot{u},\dot{v})}(s)\, h_n(s)\, \mathrm{d}s \right| \\
&\quad \le \frac{4}{n(1-t)^3} \int_{[0,t]} \left| \bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{u},\ddot{u})} \bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{v},\ddot{v})} \right| \cdot \Big(\left| \widehat{\bar{\mu}}_{n,2}^{(\dot{u},\dot{v})}(s) \right| + \left| \widehat{\bar{\mu}}_{n,1}^{(\dot{u})}(s) \widehat{\bar{\mu}}_{n,1}^{(\dot{v})}(s) \right| \Big)\, \mathrm{d}s \\
&\quad \le \frac{4}{n(1-t)^3}\, Q_n^{(u,v)}(t) \cdot \sup_{0 \le s \le t} \left| \widehat{\bar{\mu}}_{n,2}^{(\dot{u},\dot{v})}(s) - (1-s)\bar{\mu}_2^{(\dot{u},\dot{v})} \right| \\
&\qquad + \frac{4}{n(1-t)^3}\, Q_n^{(u,v)}(t) \cdot \sup_{0 \le s \le t} \left| \widehat{\bar{\mu}}_{n,1}^{(\dot{u})} \widehat{\bar{\mu}}_{n,1}^{(\dot{v})} - (1-s)^2 \bar{\mu}_1^{(\dot{u})}(s)\, \bar{\mu}_1^{(\dot{v})}(s) \right| \\
&\qquad + \frac{4}{n(1-t)^3}\, Q_n^{(u,v)}(t) \cdot (C_t + C_t^2),
\end{aligned}$$

$u, v = 1, \ldots, r$, where we set

$$Q_n^{(u,v)}(t) = \sqrt{\int_{[0,t]} \left(\bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{u},\ddot{u})}\right)^2 \mathrm{d}s \int_{[0,t]} \left(\bar{\gamma}_{n,\lfloor ns \rfloor}^{(\dot{v},\ddot{v})}\right)^2 \mathrm{d}s}, \qquad u, v = 1, \ldots, r,$$

and use the Cauchy-Schwarz inequality. Obviously, Assumption 6.2.6.i and Assumption 6.2.6.ii yield that

$$\sup_{0\leq s\leq t}\left|\widehat{\bar{\mu}}_{n,1}^{(u)}(s)\,\widehat{\bar{\mu}}_{n,1}^{(v)}(s)-(1-s)^2\,\bar{\mu}_1^{(u)}(s)\,\bar{\mu}_1^{(v)}(s)\right|\longrightarrow_{P_n'}0,\qquad\text{as } n\to\infty,$$

$u,v=1,\dots,p$. Consequently, Assumption 6.2.6.ii, Assumption 6.2.6.iii and equation (6.5) give that

$$\left|\int_{[0,t]}\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{u},\ddot{v})}\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{v},\ddot{v})}R_{n,1}^{(\dot{u},\dot{v})}(s)\,h_n(s)\,\mathrm{d}s\right|\longrightarrow_{P_n'}0,\quad\text{as } n\to\infty. \tag{6.8}$$

With the same arguments we receive that

$$\begin{aligned}&\left|\int_{[0,t]}\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{u},\dot{u})}\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{v},\ddot{v})}R_{n,2}^{(\dot{u},\dot{v})}(s)\,h_n(s)\,\mathrm{d}s\right|\\&\quad\leq(1-t)^{-2}\,Q_n^{(u,v)}(t)\cdot\sup_{0\leq s\leq t}\left|\widehat{\bar{\mu}}_{n,2}^{(u,v)}(s)-(1-s)\bar{\mu}_2^{(u,v)}(s)\right|\\&\qquad+(1-t)^{-2}\,Q_n^{(u,v)}(t)\cdot\sup_{0\leq s\leq t}\left|\widehat{\bar{\mu}}_{n,1}^{(u)}(s)\widehat{\bar{\mu}}_{n,1}^{(v)}(s)-(1-s)^2\bar{\mu}_1^{(u)}(s)\bar{\mu}_1^{(v)}(s)\right|,\end{aligned}$$

$u,v=1,\dots,r$, where

$$R_{n,2}^{(u,v)}(s)=\frac{\widehat{\bar{\mu}}_{n,2}^{(u,v)}(s)}{1-s}-\bar{\mu}_2^{(u,v)}(s)+\bar{\mu}_1^{(u)}(s)\,\bar{\mu}_1^{(v)}(s)-\frac{\widehat{\bar{\mu}}_{n,1}^{(u)}(s)\widehat{\bar{\mu}}_{n,1}^{(v)}(s)}{\big(1-s\big)^2}$$

implying that

$$\left|\int_{[0,t]}\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{u},\dot{u})}\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{v},\ddot{v})}R_{n,2}^{(\dot{u},\dot{v})}(s)\,h_n(s)\,\mathrm{d}s\right|\longrightarrow_{P_n'}0,\quad\text{as } n\to\infty. \tag{6.9}$$

Equation (6.8) and equation (6.9) yield that $\widehat{\bar{V}}_n^{(u,v)}(t)-\widetilde{\bar{V}}_n^{(u,v)}(t)\longrightarrow_{P_n'}0$, as $n\to\infty$, because of the estimate

$$\begin{aligned}\left|\widehat{\bar{V}}_n^{(u,v)}(t)-\widetilde{\bar{V}}_n^{(u,v)}(t)\right|\leq&\left|\int_{[0,t]}\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{u},\ddot{v})}\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{v},\ddot{v})}R_{n,1}^{(\dot{u},\dot{v})}(s)\,h_n(s)\,\mathrm{d}s\right|\\&+\left|\int_{[0,t]}\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{u},\dot{u})}\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{v},\ddot{v})}R_{n,2}^{(\dot{u},\dot{v})}(s)\,h_n(s)\,\mathrm{d}s\right|\end{aligned}$$

Clearly, $\widetilde{\bar{V}}_n^{(u,v)}(t) - \bar{V}_n^{(u,v)}(t) \longrightarrow_{P'_n} 0$, as $n \to \infty$, $u, v = 1, \dots, r$, where

$$\bar{V}_n^{(u,v)}(t) = \int_{[0,t]} \bar{\gamma}^{(\dot{u},\ddot{u})}(s)\bar{\gamma}^{(\dot{v},\ddot{v})}(s)\,\left(\bar{\mu}_2^{(u,v)}(s) - \bar{\mu}_1^{(u)}(s)\,\bar{\mu}_1^{(v)}(s)\right) h_n(s)\,\mathrm{d}s,$$

$u, v = 1, \dots, r$, is implied by Assumption 6.2.6.i and the estimate

$$\begin{aligned}
&\left|\widetilde{\bar{V}}_n^{(u,v)}(t) - \bar{V}_n^{(u,v)}(t)\right| \leq \\
&\quad (C_t + C_t^2)\sqrt{\int_{[0,t]} \left(\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{u},\ddot{u})} - \bar{\gamma}^{(\dot{u},\ddot{u})}(s)\right)^2 \mathrm{d}s \int_{[0,t]} \left(\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{v},\ddot{v})} - \bar{\gamma}^{(\dot{v},\ddot{v})}(s)\right)^2 \mathrm{d}s} \\
&\qquad + (C_t + C_t^2)\sqrt{\int_{[0,t]} \left(\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{u},\ddot{u})} - \bar{\gamma}^{(\dot{u},\ddot{u})}(s)\right)^2 \mathrm{d}s \int_{[0,t]} \left(\bar{\gamma}^{(\dot{v},\ddot{v})}(s)\right)^2 \mathrm{d}s} \\
&\qquad + (C_t + C_t^2)\sqrt{\int_{[0,t]} \left(\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{v},\ddot{v})} - \bar{\gamma}^{(\dot{v},\ddot{v})}(s)\right)^2 \mathrm{d}s \int_{[0,t]} \left(\bar{\gamma}^{(\dot{u},\ddot{u})}(s)\right)^2 \mathrm{d}s},
\end{aligned}$$

which is derived with the Cauchy-Schwarz inequality. Therefore, for proving $\left\langle c^{\mathrm{T}}\widehat{\bar{U}}_n\right\rangle(t) - A_c(t) \longrightarrow_{P'_n} 0$, as $n \to \infty$, it remains to be shown that

$$\int_{[0,t]} f^{(u,v)}(s)h_n(s)\,\mathrm{d}s - \int_{[0,t]} f^{(u,v)}(s)h(s)\,\mathrm{d}s \longrightarrow 0,$$

as $n \to \infty$, where

$$f^{(u,v)}(s) = \bar{\gamma}^{(\dot{u},\ddot{u})}(s)\bar{\gamma}^{(\dot{v},\ddot{v})}(s)\,\left(\bar{\mu}_2^{(u,v)}(s) - \bar{\mu}_1^{(u)}(s)\,\bar{\mu}_1^{(v)}(s)\right), \quad u, v = 1, \dots, r,$$

and h is the Lebesgue-density of the measure defined by the function $\bar{H}$, cf. in Lemma 6.2.7.

It is well known that for every $\varepsilon > 0$, there exists a continuous function $g_\varepsilon^{(u,v)}$: $[0,t] \longrightarrow \mathbb{R}$, such that $\int_{[0,t]} \left|g_\varepsilon^{(u,v)}(s) - f^{(u,v)}(s)\right| \mathrm{d}s \leq \varepsilon$. Lemma 6.2.7 and Witting and Müller-Funk [72, Satz 5.55 and Korollar 5.56] yield that

$$\int_{[0,t]} g_\varepsilon(s)\,h_n(s)\,\mathrm{d}s \longrightarrow \int_{[0,t]} g_\varepsilon(s)\,h(s)\,\mathrm{d}s.$$

Hence, it holds that

$$\begin{aligned}\limsup_{n\to\infty}\Big|\int_{[0,t]} f^{(u,v)}(s)h_n(s)\,\mathrm{d}s - \int_{[0,t]} f^{(u,v)}(s)h(s)\,\mathrm{d}s\Big| \leq \\ \limsup_{n\to\infty}\int_{[0,t]}\big|f^{(u,v)}(s)-g_\varepsilon(s)\big|h_n(s)\,\mathrm{d}s \\ +\limsup_{n\to\infty}\Big|\int_{[0,t]} g_\varepsilon(s)\,h_n(s)\,\mathrm{d}s - \int_{[0,t]} g_\varepsilon(s)\,h(s)\,\mathrm{d}s\Big| \\ +\limsup_{n\to\infty}\int_{[0,t]}\big|f^{(u,v)}(s)-g_\varepsilon(s)\big|h(s)\,\mathrm{d}s \leq 2\varepsilon,\end{aligned}$$

where one also uses the boundedness of h_n and h, see Lemma 6.2.7. Since $\varepsilon > 0$ was chosen arbitrarily, it follows the assertion. In the previous paragraphs we showed that

$$\sum_{u=1}^{r}\sum_{v=1}^{r} c^{(u)}c^{(v)}\ \widehat{\widehat{V}}_n^{(u,v)}(t) - \sum_{u=1}^{r}\sum_{v=1}^{r} c^{(u)}c^{(v)}\ \bar{\mathscr{J}}^{(u,v)}(t) \longrightarrow_{P_n'} 0, \quad \text{as } n\to\infty,$$

for all $c\in\mathbb{R}^r$, which is equivalent with $\widehat{\widehat{V}}_n(t) - \bar{\mathscr{J}}(t) \longrightarrow_{P_n'} 0$, as $n\to\infty$, $t\in[0,1)$.

Let $t\in(0,1)$ and consider the process $\{c^{\mathrm{T}}\widehat{\widehat{U}}_n(t\wedge s)\mid s\in\mathbb{R}_+\}$ and the filtration $\{\mathcal{F}'_{n,t\wedge s}\mid s\in\mathbb{R}_+\}$, where $\mathcal{F}'_{n,t\wedge s}$ is defined in Proposition 6.2.4. As this process is a martingale, where Proposition 6.2.4 and the Optional Stopping Theorem, cf. Fleming and Harrington [19, Theorem 2.2.2], are used, the previous calculations give that we can apply Rebolledo's Central Limit Theorem, see Theorem 2.1.2. Thus, it holds that

$$\{c^{\mathrm{T}}\widehat{\widehat{U}}_n(t\wedge s)\mid s\in\mathbb{R}_+\} \xrightarrow{\mathfrak{D}}_{P_n'} \{\mathbb{W}\circ A_c(t\wedge s)\mid s\in\mathbb{R}_+\}, \quad \text{as } n\to\infty, \tag{6.10}$$

in $D(\mathbb{R}_+,\mathbb{R})$. In particular this means that $c^{\mathrm{T}}\widehat{\widehat{U}}_n(t) \xrightarrow{\mathfrak{D}}_{P_n'} \mathcal{N}(0,A_c(t))$. As $c\in\mathbb{R}^r$ was arbitrary, applying the Crámer-Wold-device, cf. Witting and Müller-Funk [72, Korollar 5.69] completes the proof. □

The next result extends the assertions of the last Theorem to $t=1$.

6.2.9 Corollary. In the situation of Theorem 6.2.8 assume that $A_c(1) < \infty$, for all $c \in \mathbb{R}^r$, where $A_c(t)$ is defined in equation (6.6). Suppose that

$$\lim_{t\uparrow 1}\lim_{n\to\infty} P_n'\Big(\big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_{k_n}(\omega)\big\rangle(1) - \big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_{k_n}(\omega)\big\rangle(t) \geq \varepsilon\Big) = 0 \tag{6.11}$$

for all $\varepsilon > 0$ and for all $c \in \mathbb{R}^r$, where $\big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_{k_n}(\omega)\big\rangle(t)$ is defined in Lemma 6.2.5.d and a different representation of $\big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_{k_n}(\omega)\big\rangle(t)$ can be found in equation (6.7). Then the assertions of Theorem 6.2.8 also hold for $t = 1$.

Proof. Without loss of generality, we can assume that $k_n = n$, $n \in \mathbb{N}$. Consider the metric space $\big(D([0,1],\mathbb{R}_+),\mathcal{D}([0,1],\mathbb{R}_+),d\big)$, where d denotes the Skorokhod metric. Note that the Skorokhod metric is dominated by the supremum metric, *i.e.* $d(x,y) \leq \sup_{0\leq t\leq 1}\big|x(t) - y(t)\big|$ for all $x,y \in D[0,1]$. Moreover, let t_k, $k \in \mathbb{N}$, be a strictly increasing sequence satisfying $t_k < 1$ and $\lim_{k\to\infty} t_k = 1$. Once again, we aim to apply Theorem 2.1.1. Therefore, we define the following processes

$$\begin{aligned}
X_{n,k} &= \big\{X_{n,k}(t) \mid t \in [0,1]\big\}, & X_{n,k}(t) &= c^{\mathrm{T}}\widehat{\widehat{U}}_n(t \wedge t_k),\\
X_k &= \big\{X_{n,k}(t) \mid t \in [0,1]\big\}, & X_k(t) &= \mathbb{W}\circ A_c(t \wedge t_k),\\
X &= \big\{X(t) \mid t \in [0,1]\big\}, & X(t) &= \mathbb{W}\circ A_c(t),\\
Y_n &= \big\{Y_n(t) \mid t \in [0,1]\big\}, & Y_n(t) &= c^{\mathrm{T}}\widehat{\widehat{U}}_n(t),
\end{aligned}$$

where $A_c(t)$ is defined in equation (6.6) and $\mathbb{W}$ denotes a standard Wiener (Brownian) motion. In the proof of Theorem 6.2.8 we showed that $X_{n,k} \xrightarrow{\mathfrak{D}}_{P_n'} X_k$ in $D\big([0,1],\mathbb{R}_+\big)$, cf. equation (6.10) and note the remarks on page 26. Moreover, we have that

$$\sup_{0\leq t\leq 1}\big|X_k(t) - X(t)\big| = \sup_{t_k\leq t\leq 1}\big|\mathbb{B}\circ A_c(t_k) - \mathbb{B}\circ A_c(t)\big| \to 0 \text{ almost-surely,}$$

as $k \to \infty$, since the paths of a Wiener motion are almost surely continuous and A_c is continuous and non-decreasing. In the last step we show that

$$\lim_{k\to\infty}\limsup_{n\to\infty} P_n'\bigg(\sup_{0,\leq t\leq 1}\big|Y_n(t) - X_{n,k}(t)\big| \geq \varepsilon\bigg) = 0 \qquad \text{for all } \varepsilon > 0, \tag{6.12}$$

with the Lenglart Domination property, see Theorem 2.1.5.

Let us consider the $\{\mathcal{F}'_{n,t\wedge 1} \mid t \in \mathbb{R}_+\}$-sub-martingale

$$\big(Y_n - X_{n,k}\big)^2 = \Big\{\big(Y_n(t \wedge 1) - X_{n,k}(t \wedge 1)\big)^2 \;\big|\; t \in \mathbb{R}_+\Big\},$$

where $\mathcal{F}'_{n,t\wedge 1}$ is defined in Proposition 6.2.4. The predictable quadratic variation of $\big(Y_n - X_{n,k}\big)^2$ is given by

$$\big\langle Y_n - X_{n,k}\big\rangle(t) = \big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_n\big\rangle(t \wedge 1) - \big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_n\big\rangle(t \wedge t_k \wedge 1), \qquad t \in \mathbb{R}_+,$$

see also equation (6.7). For any bounded stopping time T, the process

$$\Big\{\big(Y_n - X_{n,k}\big)^2(T \wedge t \wedge 1) - \big\langle Y_n - X_{n,k}\big\rangle(T \wedge t \wedge 1) \;\big|\; t \in \mathbb{R}_+\Big\}$$

is a martingale because of the Optional Stopping Theorem, see Fleming and Harrington [19, Theorem 2.2.2]. Furthermore, it holds that

$$\begin{aligned}\mathbb{E}'_n\Big(\big(Y_n - X_{n,k}\big)^2(T \wedge t \wedge 1) - \big\langle Y_n - X_{n,k}\big\rangle(T \wedge t \wedge 1)\Big) = \\ \mathbb{E}'_n\Big(\big(Y_n - X_{n,k}\big)^2(0) - \big\langle Y_n - X_{n,k}\big\rangle(0)\Big) = 0\end{aligned}$$

implying that $\big(Y_n - X_{n,k}\big)^2$ is Lenglart dominated by $\big\langle Y_n - X_{n,k}\big\rangle$. Applying Theorem 2.1.5 (with the stopping time $T \equiv 1$) yields that

$$\begin{aligned}P'_n\Big(\sup_{0,\leq t \leq 1}\big|Y_n(t) - X_{n,k}(t)\big| \geq \varepsilon\Big) &= P'_n\Big(\sup_{0,\leq t \leq 1}\big|Y_n(t) - X_{n,k}(t)\big|^2 \geq \varepsilon^2\Big) \\ &\leq \frac{\eta}{\varepsilon^2} + P'_n\Big(\big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_n\big\rangle(1) - \big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_n\big\rangle(t_k) \geq \eta\Big).\end{aligned}$$

Consequently, we get that

$$\limsup_{k\to\infty}\limsup_{n\to\infty} P'_n\Big(\sup_{0,\leq t \leq 1}\big|Y_n(t) - X_{n,k}(t)\big| \geq \varepsilon\Big) \leq \frac{\eta}{\varepsilon^2}.$$

As $\eta > 0$ was arbitrary (6.12) holds. Theorem 2.1.1 gives that $Y_n \xrightarrow{\mathfrak{D}}_{P'_n} \big\{\mathbb{W} \circ A_c(t) \mid t \in [0,1]\big\}$ on $D[0,1]$.

In particular this means that $c^{\mathrm{T}}\widehat{\widehat{U}}_n(1) \xrightarrow{\mathfrak{D}}_{P'_n} \mathcal{N}(0, A_c(1))$. As $c \in \mathbb{R}^r$ was arbitrary, applying the Crámer-Wold-device, cf. Witting and Müller-Funk [72, Korollar 5.69] yields the first part of the result.

The second part of the assertion is straightforward. Let $\eta, \varepsilon > 0$ be arbitrary and choose $\tau \in (0,1)$, such that

$$\lim_{n\to\infty} P_n'\Big(\big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_n\big\rangle(1) - \big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_n\big\rangle(\tau) \geq \varepsilon/3\Big) \leq \eta \tag{6.13}$$

and $A_c(1) - A_c(\tau) \leq \varepsilon/3$. Using the estimate

$$\begin{aligned} P_n'\Big(\big|\big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_n\big\rangle(1) - A_c(1)\big| \geq \varepsilon\Big) &\leq P_n'\Big(\big|\big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_n\big\rangle(1) - \big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_n\big\rangle(\tau)\big| \geq \varepsilon/3\Big) \\ &\quad + P_n'\Big(\big|\big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_n\big\rangle(\tau) - A_c(\tau)\big| \geq \varepsilon/3\Big) + P_n'\Big(\big|A_c(\tau) - A_c(1)\big| \geq \varepsilon/3\Big), \end{aligned}$$

we receive that

$$\limsup_{n\to\infty} P_n'\Big(\Big|\big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_n\big\rangle(1) - A_c(1)\Big| \geq \varepsilon\Big) \leq \eta.$$

As $\eta > 0$ was chosen arbitrarily and using equation (6.6) and (6.7), we get that

$$\sum_{u=1}^{r}\sum_{v=1}^{r} c^{(u)}c^{(v)}\ \widehat{\widehat{V}}_n^{(u,v)}(1) - \sum_{u=1}^{r}\sum_{v=1}^{r} c^{(u)}c^{(v)}\ \bar{\mathscr{I}}^{(u,v)}(1) \longrightarrow_{P_n'} 0, \quad \text{as } n\to\infty,$$

for all $c \in \mathbb{R}^r$, which is equivalent to $\widehat{\widehat{V}}_n(1) - \bar{\mathscr{I}}(1) \longrightarrow_{P_n'} 0$, as $n \to \infty$. $\square$

The next result finally enables us to characterize the asymptotic properties of our conditional permutation tests.

6.2.10 Theorem. Abbreviating $W_{n,\uparrow} = (X_{n,\uparrow}, \Delta_{n,\uparrow}, Z_{n,\uparrow})$, let us assume that Assumption 6.2.1 is satisfied. Moreover, suppose that in every sub-sequence of natural numbers, we can find a sub-sub-sequence k_n, $n \in \mathbb{N}$, and a set $\Omega_0 \in \mathcal{F}$ with $P(\Omega_0) = 1$, such that the following premises are satisfied.

i) Assumption 6.2.6.i – Assumption 6.2.6.vi hold with the sub-sequence k_n, $n \in \mathbb{N}$, for all $\omega \in \Omega_0$.

ii) The condition (6.11) holds with the sub-sequence k_n, $n \in \mathbb{N}$, for all $\varepsilon > 0$, $c \in \mathbb{R}^r$ and $\omega \in \Omega_0$.

Then the following assertions are valid.

a) It holds that

$$F^{\star,1,\alpha}_{n,\mathfrak{J},W_{n,\uparrow}}(t) - F_{\mathfrak{J},\bar{\mathscr{I}}(1)}\big(t + F^{-1}_{\mathfrak{J},\bar{\mathscr{I}}(1)}(1-\alpha)\big) \longrightarrow_{P_{n,\xi}} 0, \quad \text{as } n \to \infty,$$

for all $t > -F^{-1}_{\mathfrak{J},\bar{\mathscr{I}}(1)}(1-\alpha)$, and $\alpha \in (0, 1/2)$, where the cumulative distribution function $F_{\mathfrak{J},\bar{\mathscr{I}}(1)}$ is defined in Theorem 4.1.14.

b) It holds that

$$\sup_{t\in\mathbb{R}_+}\Big|F^{\star,2,\alpha}_{n,\mathfrak{L}_0,\mathfrak{L}_1,W_{n,\uparrow}}(t) - F_{\mathfrak{L}_0,\mathfrak{L}_1,\bar{\mathscr{I}}(1)}\big(t + F^{-1}_{\mathfrak{L}_0,\mathfrak{L}_1,\bar{\mathscr{I}}(1)}(1-\alpha)\big)\Big| \longrightarrow_{P_{n,\xi}} 0,$$

as $n \to \infty$, for all $\alpha \in (0,1)$, where $F_{\mathfrak{L}_0,\mathfrak{L}_1,\bar{\mathscr{I}}(1)}$ denotes the distribution function of a χ^2-distribution with l degrees of freedom,

$$l = \operatorname{rank}\big(\bar{\mathscr{I}}(1)\mathscr{L}_1\big) - \operatorname{rank}\big(\bar{\mathscr{I}}(1)\mathscr{L}_0\big), \quad \operatorname{Im}(\mathscr{L})_i = \mathfrak{L}_i,\ i = 0, 1,$$

see Section 4.2, in particular Corollary 4.2.5.

Proof. First, we remember the equalities in (6.2) and note that Corollary 6.2.9 is applicable for fixed sub-sequence k_n, $n \in \mathbb{N}$, and fixed $\omega = (\omega_1, \omega_2, \ldots) \in \Omega_0$. Keeping the sub-sequence k_n, $n \in \mathbb{N}$, and $\omega \in \Omega_0$ fixed, we get that

$$\widehat{U}_{k_n,\star}\big(D'_{k_n}, W_{k_n,\uparrow}(\omega_{k_n})\big) \xrightarrow{\mathfrak{D}}_{P_{k'_n}} \mathcal{N}\big(0, \bar{\mathscr{I}}(1)\big), \quad \text{as } n \to \infty$$

and

$$\widehat{V}_{k_n,\star}(D_{k_n}, W_{k_n,\uparrow})(\omega_{k_n}) - \bar{\mathscr{I}}(1) \longrightarrow_{P'_{k_n}} 0, \qquad \text{as } n \to \infty,$$

by applying Corollary 6.2.9. We readily see that Assumption 4.1.12 holds. Theorem 4.1.13, Corollary 4.1.15 and Slutsky's Lemma, cf. Witting and Müller-Funk [72, Satz 5.83], yield that

$$T^{\star,1,\alpha}_{\mathfrak{J},1}\big(D'_{k_n}, W_{k_n,\uparrow}(\omega_{k_n})\big) \xrightarrow{\mathfrak{D}}_{P_{k'_n}} L_{\mathfrak{J},1}\big(X, \bar{\mathscr{I}}(1)\big) - F^{-1}_{\mathfrak{J},\bar{\mathscr{I}}(1)}(1-\alpha), \text{ as } n \to \infty,$$

where $X \sim \mathcal{N}\big(0, \bar{\mathscr{I}}(1)\big)$. Theorem 4.1.14 and Witting and Müller-Funk [72, Satz 5.58] give that

$$F^{\star,1,\alpha}_{k_n,\mathfrak{J},W_{n,\uparrow}(\omega_{k_n})}(t) \longrightarrow F_{\mathfrak{J},\bar{\mathscr{I}}(1)}\big(t + F^{-1}_{\mathfrak{J},\bar{\mathscr{I}}(1)}(1-\alpha)\big), \qquad \text{as } n \to \infty,$$

for all $t > -F^{-1}_{\mathfrak{J},\bar{\mathscr{J}}(1)}(1-\alpha)$. The sub-sub-sequence principle for random variables that converge in probability, see Proposition B.4.8, gives that

$$F^{\star,1,\alpha}_{n,\mathfrak{J},W_{n,\uparrow}}(t) - F_{\mathfrak{J},\bar{\mathscr{J}}(1)}\big(t + F^{-1}_{\mathfrak{J},\bar{\mathscr{J}}(1)}(1-\alpha)\big) \longrightarrow_{P_0} 0, \quad \text{as } n \to \infty$$

for all $t > -F^{-1}_{\mathfrak{J},\bar{\mathscr{J}}(1)}(1-\alpha)$, which implies the convergence in $P_{n,0}$-probability. Remark 2.2.3.c completes the proof of a).

b) is shown completely analogously, instead of Theorem 4.1.13, Corollary 4.1.15 and Theorem 4.1.14, Theorem 4.2.4 and Corollary 4.2.5 are used. Since the distribution function $F_{\mathfrak{L}_0,\mathfrak{L}_1,\bar{\mathscr{J}}(1)}$ is continuous, Witting and Müller-Funk [72, Satz 5.75] give that the cumulative distribution functions converge uniformly. □

Now, we can state the main result of this section, namely the asymptotic equivalence of the permutation tests introduced in Section 6.1 and the tests derived in Chapter 4.

6.2.11 Corollary. In the situation of Theorem 6.2.10, it holds that

$$\varphi_{n,1} - \varphi^\star_{n,1} \longrightarrow_{P_{n,\xi}} 0 \quad \text{and} \quad \varphi_{n,2} - \varphi^\star_{n,2} \longrightarrow_{P_{n,\xi}} 0, \qquad \text{as } n \to \infty.$$

This means in particular that the assertions of Corollary 4.3.6 also hold for $\varphi^\star_{n,1}$, $n \in \mathbb{N}$, and $\varphi^\star_{n,2}$, $n \in \mathbb{N}$.

Proof. We show that $\varphi_{n,1}$, $n \in \mathbb{N}$, and $\varphi^\star_{n,1}$, $n \in \mathbb{N}$, are asymptotically equivalent. The proof for the other sequence is exactly the same. Remark 2.2.3.c implies that it suffices to show the assertion under $P_{n,0}$, $n \in \mathbb{N}$. Because of Theorem 4.1.14.d, $F_{\mathfrak{J},\bar{\mathscr{J}}(1)}\big(\cdot + F^{-1}_{\mathfrak{J},\bar{\mathscr{J}}(1)}(1-\alpha)\big)$ is continuous and strictly increasing on the interval $\big(-F^{-1}_{\mathfrak{J},\bar{\mathscr{J}}(1)}(1-\alpha), \infty\big)$. Theorem 6.2.10 and Witting and Müller-Funk [72, Satz 5.76] give that $k^{\star,1}_{n,\mathfrak{J}}(\alpha, W_{n,\uparrow}) \longrightarrow_{P_{n,0}} 0$, as $n \to \infty$. Setting

$$Q_n = L_{\mathfrak{J},1}\big(\widehat{U}_n(\infty), \widehat{V}_n(\infty)\big) - c_{\mathfrak{J},1}\big(\alpha, \widehat{V}_n(\infty)\big)$$

and

$$\widehat{Q}_n = T^{\star,1,\alpha}_{\mathfrak{J},1}(D_n, W_{n,\uparrow}) - k^{\star,1}_{n,\mathfrak{J}}(\alpha, W_{n,\uparrow}),$$

we get that $Q_n - \widehat{Q}_n = k_{n,\mathcal{J}}^{\star,1}(\alpha, W_{n,\uparrow}) \longrightarrow_{P_{n,0}} 0$, as $n \to \infty$, because of $Q_n = T_{\mathcal{J},1}^{\star,1,\alpha}(D_n, W_{n,\uparrow})$. Moreover, we readily check that

$$\begin{aligned}\varphi_{n,1} - \varphi_{n,1}^{\star} = \mathbb{1}\big(Q_n \geq 0\big) - \mathbb{1}\big(\widehat{Q}_n \geq 0\big) \\ + \big(1 - r_{n,\mathcal{J}}^{\star,1}(\alpha, W_{n,\uparrow})\big) \cdot \mathbb{1}\big(\widehat{Q}_n = 0\big) - \mathbb{1}\big(Q_n = 0\big).\end{aligned}$$

Proposition B.4.7 yields that $\mathbb{1}\big(Q_n \geq 0\big) - \mathbb{1}\big(\widehat{Q}_n \geq 0\big) \longrightarrow_{P_{n,0}} 0$, as $n \to \infty$. For all $\varepsilon > 0$ it holds that

$$P_{n,0}\Big(\big|\big(1 - r_{n,\mathcal{J}}^{\star,1}(\alpha, W_{n,\uparrow})\big) \cdot \mathbb{1}\big(\widehat{Q}_n = 0\big)\big| \geq \varepsilon\Big) \leq P_{n,0}\big(\widehat{Q}_n = 0\big).$$

Because of Theorem 4.1.13, Corollary 4.1.15 and Slutsky's Lemma, cf. Witting and Müller-Funk [72, Satz 5.83], we get that

$$\widehat{Q}_n \longrightarrow_{P_{n,0}} L_{\mathcal{J},1}\big(X, \mathscr{I}^{\star}(\infty)\big) - F_{\mathcal{J},\mathscr{I}^{\star}(\infty)}^{-1}(1 - \alpha), \quad \text{as } n \to \infty,$$

where $X \sim \mathcal{N}\big(0, \mathscr{I}^{\star}(\infty)\big)$, as $n \to \infty$. Theorem 4.1.14 and the Portmanteau Theorem, cf. Witting and Müller-Funk [72, Satz 5.40], finally yield that $\lim_{n\to\infty} P_{n,0}\big(\widehat{Q}_n = 0\big) = 0$. With the same considerations one receives that $\mathbb{1}(Q_n = 0) \longrightarrow_{P_{n,0}} 0$, as $n \to \infty$, completing the proof. □

6.3 Checking Assumptions

In analogy to Section 5.1 and Section 5.2 it is shown that Assumption 6.2.6 is satisfied for an important class of examples in this Section. Note that the assumptions of Theorem 6.2.10 and Corollary 6.2.11 are based on the sub-sub-sequence principle for random variables that converge in probability, see Proposition B.4.8. Therefore, we merely have to show that the quantities in question converge in probability, as the sub-sub-sequence principle implies the assertion for fixed sub-sequences k_n and $\omega \in \Omega$.

First, it is intended to discuss Assumption 6.2.6.v, but before we can prove conditions implying this assumption we have to introduce the notion of a pseudo-inverse.

6.3.1 Definition (Pseudo-Inverse). Assume that $F : \mathbb{R} \longrightarrow [0,1]$ is some cumulative distribution function, *i.e.* F is non-decreasing, right continuous and normed in the sense that $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$. The function

$$F^{-1} : [0,1] \longrightarrow \mathbb{R} \cup \{\pm\infty\}, \qquad F^{-1}(t) = \begin{cases} \sup\bigl\{s \mid F(s) = 0\bigr\}, & t = 0, \\ \inf\bigl\{s \mid F(s) \geq t\bigr\}, & t \in (0,1), \\ \inf\bigl\{s \mid F(s) = 1\bigr\}, & t = 1, \end{cases}$$

where we define $\sup\emptyset = -\infty$ and $\inf\emptyset = \infty$, is called pseudo-inverse of F.

6.3.2 Proposition. Suppose that Assumption 6.2.1 holds and that $G, \widetilde{G} : \mathbb{R}_+ \longrightarrow [0,1]$ are continuous non-decreasing, functions. In particular assume that G is a cumulative distribution function. Set

$$G_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X_{n,i} \leq t) \;\text{ and }\; \widetilde{G}_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X_{n,i} \leq t) \cdot \Delta_{n,i},$$

$t \in \mathbb{R}_+$, $n \in \mathbb{N}$. If the conditions

$$\sup_{t\in\mathbb{R}}\bigl|G_n(t) - G(t)\bigr| \longrightarrow_{P_{n,0}} 0 \quad\text{and}\quad \sup_{t\in\mathbb{R}}\bigl|\widetilde{G}_n(t) - \widetilde{G}(t)\bigr| \longrightarrow_{P_{n,0}} 0, \tag{6.14}$$

as $n \to \infty$, hold, then

$$\sup_{t\in[0,1]}\left|\int_{[0,t]} \Delta_{n:\lfloor ns\rfloor}\, \mathrm{d}s - \widetilde{G} \circ G^{-1}(t)\right| \longrightarrow_{P_{n,0}} 0, \quad \text{as } n \to \infty,$$

where G^{-1} denotes the pseudo-inverse of G, see Definition 6.3.1. Clearly, $\widetilde{G} \circ G^{-1}$ is a non-decreasing function.

Proof. In the following we always use the pseudo-inverse, see Definition 6.3.1. One readily checks that

$$\sup_{t\in[0,1]}\left|\int_{[0,t]} \Delta_{n:\lfloor ns\rfloor}\, \mathrm{d}s - \widetilde{G}_n \circ G_n^{-1}(t)\right| \leq \frac{1}{n} \tag{6.15}$$

for all $n \in \mathbb{N}$. Therefore, we have merely to show that

$$\sup_{t\in[0,1]}\bigl|\widetilde{G}_n \circ G_n^{-1}(t) - \widetilde{G} \circ G^{-1}(t)\bigr| \longrightarrow_{P_n} 0, \quad \text{as } n \to \infty.$$

Without loss of generality, we can assume that all random variables are defined on the same probability space. Using the sub-sub-sequence principle for random variables that converge in probability, cf. Proposition B.4.8, we receive that in every sub-sequence of natural numbers we can find a sub-sub-sequence k_n, $n \in \mathbb{N}$, and a set $\Omega_0 \in \mathcal{F}$, such that $P_0(\Omega_0) = 1$ and

$$\sup_{t \in \mathbb{R}} \big| G_{k_n}(t,\omega) - G(t) \big| \longrightarrow 0 \quad \text{and} \quad \sup_{t \in \mathbb{R}} \big| \widetilde{G}_{k_n}(t,\omega) - \widetilde{G}(t) \big| \longrightarrow 0,$$

for all $\omega \in \Omega_0$. Keeping $\omega \in \Omega_0$ fixed, the functions G_{k_n}, $n \in \mathbb{N}$, are cumulative distribution functions, so that applying Witting and Müller-Funk [72, Satz 5.76] gives $G_{k_n}^{-1}(t,\omega) \to G^{-1}(t)$ for all $t \in \mathrm{Con}(G^{-1})$, where $\mathrm{Con}(G^{-1})$ denotes the set of a continuity points of G^{-1}. Because of the estimate

$$\begin{aligned}
\big| \widetilde{G}_{k_n} \circ G_{k_n}^{-1}(t,\omega) - \widetilde{G} \circ G^{-1}(t) \big| \\
&\leq \big| \widetilde{G}_{k_n} \circ G_{k_n}^{-1}(t,\omega) - \widetilde{G} \circ G_{k_n}^{-1}(t,\omega) \big| \\
&\qquad + \big| \widetilde{G} \circ G_{k_n}^{-1}(t,\omega) - \widetilde{G} \circ G^{-1}(t) \big| \\
&\leq \sup_{t \in \mathbb{R}} \big| \widetilde{G}_{k_n}(t,\omega) - \widetilde{G}(t) \big| + \big| \widetilde{G} \circ G_{k_n}^{-1}(t,\omega) - \widetilde{G} \circ G^{-1}(t) \big|,
\end{aligned}$$

we have that

$$\big| \widetilde{G}_{k_n} \circ G_{k_n}^{-1}(t,\omega) - \widetilde{G} \circ G^{-1}(t) \big| \longrightarrow 0, \quad \text{as } n \to \infty, \;\; t \in \mathrm{Con}(G^{-1}), \tag{6.16}$$

where we use (6.14) and the continuity of $\widetilde{G}$. Moreover, note that $\mathrm{Con}(G^{-1})$ is a dense set in $(0,1)$, since G^{-1} is a non-decreasing function. Additionally, one shows that (6.16) also holds for $t = 0$ and $t = 1$. As an immediate consequence we receive that $\big| \widetilde{G}_{k_n} \circ G_{k_n}^{-1}(t,\omega) - \widetilde{G} \circ G^{-1}(t) \big|$ converges uniformly to 0 on $[0,1]$. The sub-sub-sequence principle for random variables that converge in probability and (6.15) give the assertion. □

In the next step, conditions implying Assumption 6.2.6.iv are stated.

6.3.3 Proposition. Suppose that Assumption 6.2.1 holds, that H is a continuous function and that

$$\gamma^{(\dot{u},\ddot{u})} = \gamma_0^{(\dot{u},\ddot{u})} \circ H \quad \text{and} \quad \widehat{\gamma}_n^{(\dot{u},\ddot{u})} = \gamma_0^{(\dot{u},\ddot{u})} \circ \widehat{H}_n, \qquad n \in \mathbb{N},\; u = 1, \ldots, r$$

where $\gamma_0^{(\dot{u},\ddot{u})} : [0,1] \longrightarrow \mathbb{R}$, $u = 1,\dots,r$, are some continuous functions. If

$$\sup_{t\in I(\tau_0^c)} \big|\widehat{H}_n(t) - H(t)\big| \longrightarrow_{P_{n,0}} 0 \quad \text{and} \quad \sup_{t\in I(\tau_0^c)} \big|G_n(t) - G(t)\big| \longrightarrow_{P_{n,0}} 0,$$

as $n \to \infty$, $G_n(t) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}(X_{n,i} \leq t)$, $t \in \mathbb{R}_+$. Then it holds that

$$\int_{[0,1]} \big(\widehat{\gamma}_{n:\lfloor ns\rfloor}^{(\dot{u},\ddot{u})} - \bar{\gamma}^{(\dot{u},\ddot{u})}(s)\big)^2 \,\mathrm{d}s \longrightarrow_{P_{n,0}} 0, \quad \text{as } n \to \infty,$$

where $\widehat{\gamma}_{n:0}^{(\dot{u},\ddot{u})} = 0$, $\widehat{\gamma}_{n:i}^{(\dot{u},\ddot{u})} = \widehat{\gamma}_n^{(\dot{u},\ddot{u})}(X_{n:i})$, $i = 1,\dots,n$, as well as $\bar{\gamma}^{(\dot{u},\ddot{u})} = \gamma_0^{(\dot{u},\ddot{u})} \circ H \circ G^{-1}$, $u = 1,\dots,r$ and G^{-1} denotes the pseudo-inverse of G.

Proof. Let λ denote the Lebesgue measure on $\mathbb{B}[0,1]$. As $G^{-1}(u) = X_{n:i}$, for all $u \in ((i-1)/n, i/n]$, it holds that

$$\widehat{\gamma}_{n:\lfloor n\cdot\rfloor}^{(\dot{u},\ddot{u})} = \gamma_0^{(\dot{u},\ddot{u})} \circ \widehat{H}_n \circ G_n^{-1}(\cdot - 1/n)\cdot \mathbb{1}(\cdot \geq 1/n) \quad \lambda\text{-almost surely.}$$

By applying the sub-sub-sequence principle for random variables that converge in probability, cf. Proposition B.4.8, we can assume that in every sub-sequence of natural numbers we can find a sub-sub-sequence k_n, $n \in \mathbb{N}$, such that

$$\sup_{t\in I(\tau_0^c)} \big|\widehat{H}_{k_n}(t) - H(t)\big| \longrightarrow 0 \quad \text{and} \quad \sup_{t\in I(\tau_0^c)} \big|G_{k_n}(t) - G(t)\big| \longrightarrow 0,$$

as $n \to \infty$, converge P_0-almost surely. Note that $\sup_{t\in I(\tau_0^c)}\big|G_{k_n}(t) - G(t)\big| \longrightarrow 0$, as $n \to \infty$, implies that $G_{k_n}^{-1}(s - 1/k_n) \to G^{-1}(s)$, as $n \to \infty$, for all $s \in \mathrm{Con}(G^{-1})$. This is an immediate consequence of Witting and Müller-Funk [72, Satz 5.76], the monotonicity of $G_{k_n}^{-1}$, $n \in \mathbb{N}$, and G^1 as well as the left continuity of G^{-1}. As $[0,1]\backslash\mathrm{Con}(G^{-1})$ is countable the previous convergence holds for λ-almost all $s \in [0,1]$. Consequently, it holds that

$$\begin{aligned}
&\big|\widehat{H}_{k_n} \circ G_{k_n}^{-1}(s - 1/k_n) - H \circ G^{-1}(s)\big| \\
&\qquad = \big|\widehat{H}_{k_n} \circ G_{k_n}^{-1}(s - 1/k_n) - \widehat{H} \circ G_{k_n}^{-1}(s - 1/k_n)\big| \\
&\qquad\qquad + \big|\widehat{H} \circ G_{k_n}^{-1}(s - 1/k_n) - H \circ G^{-1}(s)\big| \\
&\qquad \leq \sup_{t\in I(\tau_0^c)} \big|\widehat{H}_{k_n}(t) - H(t)\big| + \big|H \circ G_{k_n}^{-1}(s - 1/k_n) - H \circ G^{-1}(s)\big| \longrightarrow 0,
\end{aligned}$$

as $n \to \infty$, for λ-almost all $s \in [0,1]$, where we also use that H is continuous on $I(\tau_0^c)$. Using the continuity of $\gamma_0^{(\dot{u},\ddot{u})}$ gives that

$$\gamma_0^{(\dot{u},\ddot{u})} \circ \widehat{H}_{k_n} \circ G_{k_n}^{-1}(\cdot - 1/k_n) \cdot \mathbb{1}(s \geq 1/k_n) \to \gamma_0^{(\dot{u},\ddot{u})} \circ H \circ G^{-1}(s) = \bar{\gamma}^{(\dot{u},\ddot{u})}(s)$$

for λ-almost all $s \in [0,1]$. Note that by construction

$$\big(\widehat{\gamma}_{k_n:\lfloor k_n s\rfloor}^{(\dot{u},\ddot{u})} - \bar{\gamma}^{(\dot{u},\ddot{u})}(s)\big)^2 \leq C, \quad s \in [0,1] \text{ and } n \in \mathbb{N},$$

for some suitable $C \in \mathbb{R}_+$. Thus, the Dominated Convergence Theorem yields that

$$\int_{[0,1]} \big(\widehat{\gamma}_{k_n:\lfloor k_n s\rfloor}^{(\dot{u},\ddot{u})} - \bar{\gamma}^{(\dot{u},\ddot{u})}(s)\big)^2 \,\mathrm{d}s \longrightarrow 0, \qquad \text{as } n \to \infty.$$

Again, applying the sub-sub-sequence principle for random variables that converge in probability yields that assertion. □

Before we finally discuss Assumption 6.2.6.ii and Assumption 6.2.6.iii in the special case of time-independent covariates, we state conditions implying the premises of Corollary 6.2.9 which are essential for proving Theorem 6.2.10 and Corollary 6.2.11.

6.3.4 Proposition. Under Assumption 6.2.1, define

$$M_{n,i}(\omega) = \max\Big\{\big|Z_{n,\uparrow,i}^{(u,v)}(\omega_n)\big|^2 \;\big|\; u = 1,\dots,p,\ v = 1,\dots,n\Big\}, \quad \omega \in \Omega,$$

$i = 1,\dots,n$, $n \in \mathbb{N}$. Suppose that in every sub-sequence of natural numbers we can find a sub-sub-sequence $k_n = k_n$, $n \in \mathbb{N}$, and a set $\Omega_0 \in \mathfrak{F}$, $P_0(\Omega_0) = 1$, such that for all $\omega = (\omega_1, \omega_2, \dots) \in \Omega_0$ the following conditions hold.

i) $\frac{1}{k_n}\sum_{i=1}^{k_n} M_{k_n,i}(\omega) \leq C(\omega) < \infty$, $n \in \mathbb{N}$.

ii) $\big(\widehat{\gamma}_{k_n:k_n}^{(\dot{u},\ddot{u})}(\omega_n)\big)^2/k_n \longrightarrow 0$, as $n \to \infty$, $u = 1,\dots,r$.

iii) $\int_{[0,1]} \big(\widehat{\gamma}_{k_n,\lfloor k_n s\rfloor}^{(\dot{u},\ddot{u})}(\omega_n) - \bar{\gamma}^{(\dot{u},\ddot{u})}\big)^2 \,\mathrm{d}s \longrightarrow 0$, as $n \to \infty$, $u = 1,\dots,r$, where $\bar{\gamma}^{(\dot{u},\ddot{u})} : [0,1] \longrightarrow \mathbb{R}$, $u = 1,\dots,r$ are square integrable functions.

Then the condition (6.11) in Corollary 6.2.9 holds with the sequence k_n, $n \in \mathbb{N}$, for all $\varepsilon > 0$, $c \in \mathbb{R}^r$ and $\omega \in \Omega_0$.

Proof. Let $\omega \in \Omega_0$ be fixed and let us use the notation provided in Definition 6.2.2. Without loss of generality we can assume that $k_n = n$, $n \in \mathbb{N}$. Remember that $\langle c^{\mathrm{T}} \widehat{\widehat{U}}_n(\omega) \rangle(t) = \sum_{l=1}^{\lfloor nt \rfloor} K^0_{n,l}$, see Lemma 6.2.5. Using the estimates $\left(\sum_{j=1}^{q} a_j\right)^2 \le q \sum_{j=1}^{q} a_j^2$, where $a_j \in \mathbb{R}$, $j = 1, \ldots, q$, and $q \in \mathbb{N}$, and $0 \le \bar{\delta}_{n,i} \le 1$, we get the estimate

$$K^0_{n,l} \le \frac{2r}{n} \sum_{j=1}^{m_n} \sum_{u=1}^{r} \left(c^{(u)}\, \bar{\gamma}^{(\dot{u},\ddot{u})}_{n,l}\right)^2 \left(\left(\bar{z}^{(\dot{u},l)}_{n,j}\right)^2 + \frac{\sum_{k=l}^{n} \left(\zeta^{(\dot{u},l)}_{n,k}\right)^2}{n+1-l} \right) \times \frac{\nu_{n,j} - \sum_{k=1}^{l-1} \mathbb{1}(\bar{z}_{n,j} = \zeta_{n,k})}{n+1-l}.$$

Lemma 6.2.3 gives that

$$\sum_{j=1}^{m_n} \left(\bar{z}^{(\dot{u},l)}_{n,j}\right)^2 \cdot \frac{\nu_{n,j} - \sum_{k=1}^{l-1} \mathbb{1}(\bar{z}_{n,j} = \zeta_{n,k})}{n+1-l} = \mathbb{E}'_n \left[(\zeta^{(\dot{u},l)}_{n,l})^2 \mid \zeta_{n,1}, \ldots, \zeta_{n,l-1} \right]$$

and

$$\sum_{j=1}^{m_n} \frac{\sum_{k=l}^{n} \left(\zeta^{(\dot{u},l)}_{n,k}\right)^2}{n+1-l} \cdot \frac{\nu_{n,j} - \sum_{k=1}^{l-1} \mathbb{1}(\bar{z}_{n,j} = \zeta_{n,k})}{n+1-l} = \frac{\sum_{k=l}^{n} \left(\zeta^{(\dot{u},l)}_{n,k}\right)^2}{n+1-l}.$$

Consequently, it holds that

$$K^0_{n,l} \le \frac{2r}{n} \sum_{u=1}^{r} \left(c^{(u)}\, \bar{\gamma}^{(\dot{u},\ddot{u})}_{n,l}\right)^2 \left(\mathbb{E}'_n \left[(\zeta^{(\dot{u},l)}_{n,l})^2 \mid \mathcal{F}'_{n,(l-1)/n} \right] + \frac{\sum_{k=l}^{n} \left(\zeta^{(\dot{u},l)}_{n,k}\right)^2}{n+1-l} \right).$$

Thus, we receive that

$$\begin{aligned}
\langle c^{\mathrm{T}} \widehat{\widehat{U}}_n(\omega) \rangle(1) - \langle c^{\mathrm{T}} \widehat{\widehat{U}}_n(\omega) \rangle(t) = \sum_{l=\lfloor nt \rfloor+1}^{n} K^0_{n,l} \le \\
2r \sum_{u=1}^{r} \left(c^{(u)}\right)^2 \int_{[t,1]} \left(\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,\lfloor ns \rfloor}\right)^2 \mathbb{E}'_n \left[(\zeta^{(\dot{u},\lfloor ns \rfloor)}_{n,\lfloor ns \rfloor})^2 \mid \mathcal{F}'_{n,(\lfloor ns \rfloor-1)/n} \right] \mathrm{d}s \\
+ 2r \sum_{u=1}^{r} \left(c^{(u)}\right)^2 \int_{[t,1]} \left(\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,\lfloor ns \rfloor}\right)^2 \frac{\sum_{k=\lfloor ns \rfloor}^{n} \left(\zeta^{(\dot{u},\lfloor ns \rfloor)}_{n,k}\right)^2}{n+1-\lfloor ns \rfloor} \mathrm{d}s \\
+ \frac{2r}{n} \sum_{u=1}^{r} \left(c^{(u)}\right)^2 \left(\bar{\gamma}^{(\dot{u},\ddot{u})}_{n,n}\right)^2 \left(\mathbb{E}'_n \left[(\zeta^{(\dot{u},n)}_{n,n})^2 \mid \mathcal{F}'_{n,(n-1)/n} \right] + \left(\zeta^{(\dot{u},n)}_{n,n}\right)^2 \right).
\end{aligned}$$

Moreover, we note that

$$\begin{aligned}\mathbb{E}_n'\Big(\mathbb{E}_n'\big[(\zeta_{n,\lfloor ns\rfloor}^{(\dot{u},\lfloor ns\rfloor)})^2 \mid \mathcal{F}_{n,(\lfloor ns\rfloor-1)/n}'\big]\Big) &= \mathbb{E}_n'\big(\zeta_{n,\lfloor ns\rfloor}^{(\dot{u},\lfloor ns\rfloor)}\big)^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}\big(Z_{n,\uparrow,i}^{(\dot{u},\lfloor ns\rfloor)}(\omega)\big)^2 \le \frac{1}{n}\sum_{i=1}^{n} M_{n,i}(\omega) \quad (6.17)\end{aligned}$$

and that

$$\mathbb{E}_n'\left(\frac{\sum_{k=\lfloor ns\rfloor}^{n}\big(\zeta_{n,k}^{(\dot{u},\lfloor ns\rfloor)}\big)^2}{n+1-\lfloor ns\rfloor}\right) = \frac{\sum_{k=\lfloor ns\rfloor}^{n}\mathbb{E}_n'\big(\zeta_{n,k}^{(\dot{u},\lfloor ns\rfloor)}\big)^2}{n+1-\lfloor ns\rfloor} \le \frac{1}{n}\sum_{i=1}^{n} M_{n,i}(\omega). \tag{6.18}$$

Applying the Markov-inequality, cf. Gänssler and Stute [20, Lemma 1.18.1], Fubini's Theorem, cf. Bauer [6, Korollar 23.7], as well as (6.17) and (6.18), we receive that

$$\begin{aligned}&P_n'\Big(\big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_{k_n}(\omega)\big\rangle(1) - \big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_{k_n}(\omega)\big\rangle(t) \ge \varepsilon\Big) \le \\ &\quad \frac{8r}{\varepsilon}\cdot\left(\frac{1}{n}\sum_{i=1}^{n} M_{n,i}(\omega)\right)\cdot\sum_{u=1}^{r}\big(c^{(u)}\big)^2\int_{[t,1]}\big(\bar{\gamma}_{n,\lfloor ns\rfloor}^{(\dot{u},\ddot{u})} - \bar{\gamma}^{(\dot{u},\ddot{u})}(s)\big)^2\,\mathrm{d}s \\ &\qquad + \frac{8r}{\varepsilon}\cdot\left(\frac{1}{n}\sum_{i=1}^{n} M_{n,i}(\omega)\right)\cdot\sum_{u=1}^{r}\big(c^{(u)}\big)^2\int_{[t,1]}\big(\bar{\gamma}^{(\dot{u},\ddot{u})}(s)\big)^2\,\mathrm{d}s \\ &\qquad + \frac{4r}{\varepsilon}\cdot\left(\frac{1}{n}\sum_{i=1}^{n} M_{n,i}(\omega)\right)\cdot\sum_{u=1}^{r}\big(c^{(u)}\big)^2\,\frac{1}{n}\big(\bar{\gamma}_{n,n}^{(\dot{u},\ddot{u})}\big)^2.\end{aligned}$$

Therefore, it holds that

$$\begin{aligned}&\limsup_{n\to\infty} P_n'\Big(\big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_{k_n}(\omega)\big\rangle(1) - \big\langle c^{\mathrm{T}}\widehat{\widehat{U}}_{k_n}(\omega)\big\rangle(t) \ge \varepsilon\Big) \le \\ &\qquad\qquad \frac{8r\cdot C(\omega)}{\varepsilon}\sum_{u=1}^{r}\big(c^{(u)}\big)^2\int_{[t,1]}\big(\bar{\gamma}^{(\dot{u},\ddot{u})}(s)\big)^2\,\mathrm{d}s \longrightarrow 0,\end{aligned}$$

as $t \to 1$, where we also use the Dominated Convergence Theorem. □

Remember that the theory developed in this chapter only applies to external covariates, for which time-independent covariates are major example. Therefore, we only consider time-independent covariates for the remaining conditions of Assumption 6.2.6.

6.3.5 Proposition (Time-Independent Covariates). Suppose that Assumption 6.2.1 holds, that $\widetilde{Z}_{n,i}$, $i = 1, \dots, n$, $n \in \mathbb{N}$, are time independent covariates and that the covariates processes $\{Z_{n,i}(t) \mid t \in \mathbb{R}_+\}$ are given by $Z_{n,i}(t) = \widetilde{Z}_{n,i} \cdot \mathbb{1}(t > 0)$, cf. Example 5.2.11. If

$$\frac{1}{n} \sum_{i=1}^{n} \widetilde{Z}_{n,i}^{(u)} \longrightarrow_{P_{n,0}} \widetilde{\mu}_1^{(u)} \in \mathbb{R} \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^{n} \widetilde{Z}_{n,i}^{(u)} \widetilde{Z}_{n,i}^{(v)} \longrightarrow_{P_{n,0}} \widetilde{\mu}_2^{(u,v)} \in \mathbb{R},$$

as $n \to \infty$, $u, v = 1, \dots, p$, and

$$\frac{1}{\sqrt{n}} \max_{1 \le i \le n} \max_{1 \le u \le p} |\widetilde{Z}_{n,i}^{(u)}| \longrightarrow_{P_{n,0}} 0, \qquad \text{as } n \to \infty,$$

then we can find in every sub-sequence of natural numbers a sub-sub-sequence k_n, $n \in \mathbb{N}$, and a set $\Omega_0 \in \mathcal{F}$, $P_0(\Omega_0) = 1$, such that Assumption 6.2.6.i, Assumption 6.2.6.ii and Assumption 6.2.6.iii hold with k_n, $n \in \mathbb{N}$, for all $\omega \in \Omega_0$.

The following Glivenko-Cantelli-type result is the key for the proof of Proposition 6.3.5.

6.3.6 Lemma. Under Assumption 6.2.1, let $a_{k_n,i}$, $i = 1, \dots, k_n$, $k_n \in \mathbb{N}$, $n \in \mathbb{N}$, be a triangular array of real numbers and assume that $\lim_{n\to\infty} k_n = \infty$ and that

$$\frac{1}{k_n} \sum_{i=1}^{k_n} a_{k_n,i} \longrightarrow a \in \mathbb{R} \quad \text{and} \quad \frac{1}{k_n} \max_{1 \le i \le k_n} |a_{k_n,i}| \longrightarrow 0, \tag{6.19}$$

as $n \to \infty$, and that the sequence $\frac{1}{k_n} \sum_{i=1}^{k_n} |a_{k_n,i}|$, $n \in \mathbb{N}$, is bounded. Then it holds that

$$\sup_{0 \le t \le 1} \left| \frac{1}{n} \sum_{i=1}^{\lfloor k_n t \rfloor} a_{k_n, D'_{k_n,i}} - t \cdot a \right| \longrightarrow_{P'_{k_n}} 0.$$

Proof. Without loss of generality, we can assume that $k_n = n$, $n \in \mathbb{N}$. Let us assume that $a_{n,i} \ge 0$, $i = 1, \dots, n$, $n \in \mathbb{N}$. As a first step we show that

$\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} a_{n,D'_{n,i}} \longrightarrow_{P'_n} t\cdot a$ for every fixed $t\in[0,1]$. As $\frac{\lfloor nt\rfloor}{n^2}\sum_{i=1}^{n} a_{n,i} \to t\cdot a$ and

$$P'_n\left(\left|\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} a_{n,D'_{n,i}} - \frac{\lfloor nt\rfloor}{n^2}\sum_{i=1}^{n} a_{n,i}\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2}\operatorname{Var}_{P'_n}\left(\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} a_{n,D'_{n,i}}\right),$$

where we used the Tchebychef-inequality, cf. Gänssler and Stute [20, Korollar 1.18.3], we simply need to show that

$$\begin{aligned}\operatorname{Var}_{P'_n}\left(\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} a_{n,D'_{n,i}}\right) &= \frac{\lfloor nt\rfloor}{n}\frac{1}{n}\operatorname{Var}_{P'_n}\big(a_{n,D'_{n,1}}\big)\\ &\quad + \frac{\lfloor nt\rfloor\big(\lfloor tn\rfloor - 1\big)}{n^2}\operatorname{Cov}\big(a_{n,D'_{n,1}}, a_{n,D'_{n,2}}\big) \longrightarrow_{P'_n} 0,\end{aligned}$$

as $n\to\infty$. We have that $\lfloor nt\rfloor/n \to t$ and

$$\begin{aligned}0 \leq \frac{1}{n}\operatorname{Var}_{P'_n}\big(a_{n,D'_{n,1}}\big) \leq \frac{1}{n}\mathbb{E}_{P'_n}\big(a^2_{n,D'_{n,1}}\big) &= \frac{1}{n^2}\sum_{i=1}^{n} a^2_{n,i}\\ &\leq \frac{1}{n}\max_{1\leq i\leq n}|a_{n,i}|\,\frac{1}{n}\sum_{i=1}^{n}|a_{n,i}| \longrightarrow 0,\end{aligned}$$

as $n\to\infty$, as well as

$$\begin{aligned}\mathbb{E}'_n\big(a_{n,D'_{n,1}}\,a_{n,D'_{n,2}}\big) &= \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,\,j\neq i}^{n} a_{n,i}a_{n,j}\\ &= \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n} a_{n,i}\right)^2 - \frac{n}{n-1}\frac{1}{n^2}\sum_{i=1}^{n} a^2_{n,i} \longrightarrow a^2,\end{aligned}$$

as $n\to\infty$, and

$$\mathbb{E}'_n\big(a_{n,D'_{n,i}}\big) = \frac{1}{n}\sum_{i=1}^{n} a_{n,i} \longrightarrow a, \qquad \text{as } n\to\infty.$$

Consequently, it holds that

$$\operatorname{Cov}\big(a_{n,D'_{n,1}}, a_{n,D'_{n,2}}\big) = \mathbb{E}'_n\big(a_{n,D'_{n,1}}\,a_{n,D'_{n,2}}\big) - \mathbb{E}'_n\big(a_{n,D'_{n,1}}\big)\,\mathbb{E}'_n\big(a_{n,D'_{n,2}}\big) \longrightarrow 0,$$

as $n\to\infty$, completing the proof of the first step.

As a second step we show that the convergence is uniform. For this purpose we use the same idea as for the proof of the Glivenko-Cantelli Theorem. It holds the estimate

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{\lfloor n(j-1)/m\rfloor} a_{n,D'_{n,i}} - \frac{j-1}{m}a - \frac{a}{m} &\leq \frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} a_{n,D'_{n,i}} - t\cdot a \\ &\leq \frac{1}{n}\sum_{i=1}^{\lfloor n\,j/m\rfloor} a_{n,D'_{n,i}} - \frac{j}{m}a + \frac{a}{m},\end{aligned}$$

whenever $\frac{j-1}{m} \leq t \leq \frac{j}{m}$, $m \in \mathbb{N}$. Since we can find for every $\varepsilon > 0$ an $m \in \mathbb{N}$, such that $\varepsilon - \frac{a}{m} > 0$, we get that

$$\begin{aligned}P'_n&\left(\sup_{0\leq t\leq 1}\left|\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} a_{n,D'_{n,i}} - t\cdot a\right| \geq \varepsilon\right) \\ &\leq P'_n\left(\max_{1\leq j\leq m}\left|\frac{1}{n}\sum_{i=1}^{\lfloor n\,j/m\rfloor} a_{n,D'_{n,i}} - \frac{j}{m}\cdot a\right| \geq \varepsilon - \frac{a}{m}\right),\end{aligned}$$

where the right hand side converges to 0 because of $\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} a_{n,D'_{n,i}} \longrightarrow_{P'_n} t\cdot a$, as $n \to \infty$. By now we have shown that the assertion of the Lemma holds for non-negative $a_{n,i}$, $i = 1,\dots,n$.

In the last step we consider arbitrary $a_{n,i}$, $i = 1,\dots,n$. We define $a^+_{n,i} = a_{n,i}\,\mathbb{1}\big(a_{n,i} \geq 0\big)$ and $a^-_{n,i} = -a_{n,i}\,\mathbb{1}\big(a_{n,i} \geq 0\big)$, $i = 1,\dots,n$, $n \in \mathbb{N}$. Obviously, we have that

$$\frac{1}{n}\max_{1\leq i\leq n}\left|a^+_{n,i}\right| \longrightarrow 0 \quad \text{and} \quad \frac{1}{n}\max_{1\leq i\leq n}\left|a^-_{n,i}\right| \longrightarrow 0,$$

as $n \to \infty$. Additionally, $\frac{1}{n}\sum_{i=1}^{n} a^+_{n,i}$ and $\frac{1}{n}\sum_{i=1}^{n} a^-_{n,i}$ are bounded. Therefore we can find in every sub-sequence of natural numbers a sub-sequence k_n, $n \in \mathbb{N}$, such that $\frac{1}{k_n}\sum_{i=1}^{k_n} a^+_{k_n,i} \to a^+$ and $\frac{1}{k_n}\sum_{i=1}^{k_n} a^-_{k_n,i} \to a^-$, as $n \to \infty$. Because of

$$\frac{1}{k_n}\sum_{i=1}^{k_n} a^+_{k_n,i} - \frac{1}{k_n}\sum_{i=1}^{k_n} a^-_{k_n,i} = \frac{1}{k_n}\sum_{i=1}^{k_n} a_{k_n,i},$$

we get that $a^+ - a^- = a$. As we have already proved that the Lemma holds for non-negative $a_{n,i}$, $i = 1, \dots, n$, it results that

$$P'_{k_n}\left(\sup_{0\le t\le 1}\left|\frac{1}{k_n}\sum_{i=1}^{\lfloor k_n t\rfloor} a_{n_k,D'_{k_n,i}} - t\cdot a\right| \ge \varepsilon\right) \le$$
$$P'_{k_n}\left(\sup_{0\le t\le 1}\left|\frac{1}{k_n}\sum_{i=1}^{\lfloor k_n t\rfloor} a^+_{n_k,D'_{k_n,i}} - t\cdot a^+\right| \ge \frac{\varepsilon}{2}\right) +$$
$$P'_{k_n}\left(\sup_{0\le t\le 1}\left|\frac{1}{k_n}\sum_{i=1}^{\lfloor k_n t\rfloor} a^-_{n_k,D'_{k_n,i}} - t\cdot a^-\right| \ge \frac{\varepsilon}{2}\right) \longrightarrow 0,$$

as $n \to \infty$. Applying the sub-sub-sequence principle yields that

$$\lim_{n\to\infty} P'_n\left(\sup_{0\le t\le 1}\left|\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} a_{n,D'_{n,i}} - t\cdot a\right| \ge \varepsilon\right) = 0.$$

□

Proof of Proposition 6.3.5. Because of the sub-sub-sequence principle for random variables that converge in probability, cf. Proposition B.4.8, for every sub-sequence of the natural number there exists a sub-sub-sequence k_n, $n \in \mathbb{N}$, and a set $\Omega_0 \in \mathcal{F}$, $P(\Omega_0) = 1$, such that

$$\frac{1}{k_n}\sum_{i=1}^{k_n} \widetilde{Z}^{(u)}_{k_n,i}(\omega) \longrightarrow \widetilde{\mu}^{(u)}_1 \quad \text{and} \quad \frac{1}{k_n}\sum_{i=1}^{k_n} \widetilde{Z}^{(u)}_{k_n,i}(\omega)\,\widetilde{Z}^{(v)}_{k_n,i}(\omega) \longrightarrow \widetilde{\mu}^{(u,v)}_2 \tag{6.20}$$

as well as

$$\frac{1}{k_n}\max_{1\le i\le k_n}\left|Z^{(u)}_{k_n,i}(\omega)\right| \longrightarrow 0, \quad \text{and} \quad \frac{1}{k_n}\max_{1\le i\le k_n}\left|Z^{(u)}_{k_n,i}(\omega)\,Z^{(v)}_{k_n,i}(\omega)\right| \longrightarrow 0,$$

as $n \to \infty$, $u, v = 1, \dots, n$, for all $\omega \in \Omega_0$.

Keeping $\omega \in \Omega_0$ fixed, we show that

$$\frac{1}{k_n}\sum_{i=1}^{k_n}\left|\widetilde{Z}^{(u)}_{k_n,i}(\omega)\right| \le C \quad \text{and} \quad \frac{1}{k_n}\sum_{i=1}^{k_n}\left|\widetilde{Z}^{(u)}_{k_n,i}(\omega)\,\widetilde{Z}^{(v)}_{k_n,i}(\omega)\right| \le C, \tag{6.21}$$

$n \in \mathbb{N}$, $u, v = 1, \ldots, p$, for some $C \in \mathbb{R}_+$. Because of (6.20) there exists a $C \in \mathbb{R}_+$, such that

$$\frac{1}{k_n} \sum_{i=1}^{k_n} \big| \widetilde{Z}_{k_n,i}^{(u)}(\omega) \big|^2 \leq C, \quad n \in \mathbb{N}, \qquad u = 1, \ldots, p. \tag{6.22}$$

Applying the Jensen inequality and the Cauchy-Schwarz inequality, cf. Gänssler and Stute [20, Satz 5.4.7, Satz 1.13.2] give that

$$\frac{1}{k_n} \sum_{i=1}^{k_n} \big| \widetilde{Z}_{k_n,i}^{(u)}(\omega) \big| \leq \sqrt{\frac{1}{k_n} \sum_{i=1}^{k_n} \big| \widetilde{Z}_{k_n,i}^{(u)}(\omega) \big|^2}$$

and

$$\frac{1}{k_n} \sum_{i=1}^{k_n} \big| \widetilde{Z}_{k_n,i}^{(u)}(\omega)\, \widetilde{Z}_{k_n,i}^{(v)}(\omega) \big| \leq \sqrt{\frac{1}{k_n} \sum_{i=1}^{k_n} \big| \widetilde{Z}_{k_n,i}^{(u)}(\omega) \big|^2} \cdot \sqrt{\frac{1}{k_n} \sum_{i=1}^{k_n} \big| \widetilde{Z}_{k_n,i}^{(v)}(\omega) \big|^2}.$$

Consequently, (6.22) implies (6.21). Now, the assertion is an immediate consequence of Lemma 6.3.6. □

6.3.7 Remark. Both in Proposition 6.3.2 and Proposition 6.3.3 we assume that

$$\sup_{t \in I(\tau_0^c)} \big| G_n(t) - G(t) \big| \longrightarrow_{P_{n,0}} 0, \qquad \text{as } n \to \infty, \tag{6.23}$$

where $G_n(t) = \frac{1}{n} \sum_{i=1}^{n} 1(X_{n,i} \leq t)$, $t \in \mathbb{R}$. In particular, it holds that $G_n(t) = 1 - \widehat{\mu}_{n,0}(t+)$, where $\widehat{\mu}_{n,0}(t+) = \lim_{h \downarrow 0} \widehat{\mu}_{n,0}(t+h)$. As in our setting it holds that $t \mapsto \mathbb{E}_{n,0}\big(\widehat{\mu}_{n,0}(t)\big)$ is a continuous function, we get that

$$\sup_{t \in I(\tau_0^c)} \big| \widehat{\mu}_{n,0}(t+) - \mathbb{E}_{n,0} \widehat{\mu}_{n,0}(t) \big| = \sup_{t \in I(\tau_0^c)} \big| \widehat{\mu}_{n,0}(t) - \mathbb{E}_{n,0} \widehat{\mu}_{n,0}(t) \big|,$$

$P_{n,0}$-almost surely, where we also use Assumption 3.2.1. Therefore, Proposition 5.2.2 and Corollary 5.2.3 can be used to verify condition (6.23). Moreover, one sees that the remaining assumptions of Proposition 6.3.3 are the same as in Example 5.2.7.

In the following discussion we summarize the previous results and state a setting in which previous premises are satisfied.

6.3.8 Discussion. Suppose that Assumption 4.3.3 holds and that $\widetilde{Z}_{n,i}$, $i = 1,\dots,n$, $n \in \mathbb{N}$, are time-independent covariates and that the covariates processes $\{Z_{n,i}(t) \mid t \in \mathbb{R}_+\}$, $i = 1,\dots,n$, are given by $Z_{n,i}(t) = \widetilde{Z}_{n,i} \cdot \mathbb{1}(t > 0)$, $t \in \mathbb{R}$, $i = 1,\dots,n$, cf. Example 5.2.11. Moreover, let us assume that $n > k$ and that $n_i = n_i(n)$, $i = 0,\dots,k$, are sequences of natural numbers, such that $n_0 = 0$ and $n_k = n$ and that the random variables

$$(\widetilde{Z}_{n,i}, X_{n,i}, \Delta_{n,i}) \sim (\widetilde{Z}_l, X_l, \Delta_l), \quad n_{l-1} < i \leq n_l,$$

under $P_{n,0}$, $n \in \mathbb{N}$, $l = 1,\dots,k$. In other words we consider a k-sample problem, see also Example 5.2.5. If $\big(n_l - n_{l-1}\big)/n \to \nu_l$, as $n \to \infty$, for $l = 1,\dots,k$ and if all covariates are square integrable it obviously holds that

$$\lim_{C\to\infty} \sup_{n\in\mathbb{N}} \sup_{1\leq i\leq n} \int \mathbb{1}\Big(\max_{1\leq u\leq p} (\widetilde{Z}^{(u)}_{n,i})^2 > C\Big) \max_{1\leq u\leq p} (\widetilde{Z}^{(u)}_{n,i})^2 \,\mathrm{d}P_{n,0} = 0.$$

Remark 5.2.9.a yields that

$$\frac{1}{\sqrt{n}} \max_{1\leq i\leq n} \max_{1\leq u\leq p} |\widetilde{Z}^{(u)}_{n,i}| \longrightarrow_{P_{n,0}} 0, \qquad \text{as } n \to \infty. \tag{6.24}$$

Chinchin's Weak Law of Large Numbers (WLLN) gives that

$$\frac{1}{n}\sum_{i=1}^{n} \widetilde{Z}^{(u)}_{n,i} \longrightarrow_{P_{n,0}} \widetilde{\mu}^{(u)}_1 \in \mathbb{R} \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n} \widetilde{Z}^{(u)}_{n,i}\widetilde{Z}^{(v)}_{n,i} \longrightarrow_{P_{n,0}} \widetilde{\mu}^{(u,v)}_2 \in \mathbb{R}, \tag{6.25}$$

in particular the Assumptions of Proposition 6.3.5 hold. Setting

$$G_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X_{n,i} \leq t) \quad \text{and} \quad \widetilde{G}_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X_{n,i} \leq t)\cdot \Delta_{n,i},$$

we get that $G_n(t) - G(t) \longrightarrow_{P_{n,0}} 0$ and $G_n(t) - G(t) \longrightarrow_{P_{n,0}} 0$, as $n \to \infty$, for all $t \in (-\infty, \infty)$, where

$$G(t) = \sum_{l=1}^{k} \nu_l \cdot \mathbb{P}(X_l \leq t) \quad \text{and} \quad \widetilde{G}(t) = \sum_{l=1}^{k} \nu \cdot \mathbb{E}\big(\mathbb{1}(X_{n,i} \leq t)\cdot \Delta_{n,i}\big),$$

and the WLLN is applied again. As an immediate consequence, we get that $\sup_{t\in\mathbb{R}_+}\left|G_n(t)-G(t)\right| \longrightarrow_{P_{n,0}} 0$ and $\sup_{t\in\mathbb{R}_+}\left|\widetilde{G}_n(t)-\widetilde{G}(t)\right| \longrightarrow_{P_{n,0}} 0$, as $n\to\infty$. Thus, we proved that the premises of Proposition 6.3.2 hold, *i.e.*

$$\sup_{t\in[0,1]}\left|\int_{[0,t]}\Delta_{n:\lfloor ns\rfloor}\,\mathrm{d}s-\widetilde{G}\circ G^{-1}(t)\right| \longrightarrow_{P_{n,0}} 0, \quad \text{as } n\to\infty. \tag{6.26}$$

Furthermore, let us suppose that we are in the situation of Proposition 6.3.3, see also Example 5.2.7. Note that the only assumption in Proposition 6.3.3 not concerning the weight functions, namely $\sup_{t\in I(\tau_0^c)}\left|G_n(t)-G(t)\right| \longrightarrow_{P_{n,0}} 0$, as $n\to\infty$, was already verified. Consequently, we get that

$$\int_{[0,1]}\left(\widehat{\gamma}^{(\dot{u},\ddot{u})}_{n:\lfloor ns\rfloor}-\bar{\gamma}^{(\dot{u},\ddot{u})}(s)\right)^2\mathrm{d}s \longrightarrow_{P_{n,0}} 0, \quad \text{as } n\to\infty. \tag{6.27}$$

Let m_n, $n\in\mathbb{N}$, be some sub-sequence of natural numbers. Using Proposition 6.3.5, cf. equation (6.25), we can find a sub-sub-sequence k_n, $n\in\mathbb{N}$, and a set $\Omega_{0,1}\in\mathcal{F}$, $P_0(\Omega_{0,1})=1$, such that Assumption 6.2.6.i – Assumption 6.2.6.iii hold for k_n, $n\in\mathbb{N}$, and all $\omega\in\Omega_{0,1}$. Because of the sub-sequence principle for random variables that converge in probability, see Proposition B.4.8 and (6.24), (6.26) as well as (6.27), we can find a set $\Omega_{0,2}\in\mathcal{F}$, $P_0(\Omega_{0,2})=1$, and a sub-sub-sequence of the sub-sequence k_n, $n\in\mathbb{N}$, which we call k'_n, $n\in\mathbb{N}$, such that Assumption 6.2.6.iv – Assumption 6.2.6.vi hold for the sub-sub-sequence k'_n, $n\in\mathbb{N}$, and for all $\omega\in\Omega_0=\Omega_{0,1}\cap\Omega_{0,2}$, where we note that $P_0(\Omega_0)=1$. In particular this means that Assumption 6.2.6.i – Assumption 6.2.6.vi hold with the sub-sequence k'_n, $n\in\mathbb{N}$, for all $\omega\in\Omega_0$.

Because of the estimate

$$0\le\frac{1}{k'_n}\sum_{i=1}^{k_n}M_{k'_n,i}(\omega)\le\sum_{u=1}^{p}\frac{1}{k'_n}\sum_{i=1}^{k'_n}\widehat{\mu}^{(u,u)}_{k'_n}(0,\omega)\longrightarrow\sum_{u=1}^{p}\widetilde{\mu}^{(u,u)}_2,$$

as $n\to\infty$, for all $\omega\in\Omega_0$. Assumption i) of Proposition 6.3.4 holds. Assumption ii) of Proposition 6.3.4 is valid because of the boundedness of the weight functions and assumption iii) of Proposition 6.3.4 is exactly Assumption 6.2.6.iv.

Recapitulating, we showed that in the setting of this discussion the premises of Theorem 6.2.10 and Corollary 6.2.11 hold, *i.e.* the permutation tests introduced in Section 6.1 and the tests of Section 4.3 are asymptotically equivalent.

A Omitted Proofs

In the text several proofs were omitted for various reason. Some of them are well known results or slight modifications of such results. Others were omitted to increase the readability. These proofs are collected in this Appendix.

A.1 Proof of Corollary 2.1.3

We want to apply Theorem 2.1.2. Setting $U_n(t) = \sum_{i=1}^{k_n} \int_{[0,t]} H_n^{(i)}(s)\, \mathrm{d}M_n^{(i)}(s)$, $t \in \mathbb{R}_+$, Fleming and Harrington [19, Theorem 2.4.3 and Theorem 2.5.2] give that $\langle U_n \rangle(t)$ is exactly the left hand side of (2.1). Obviously, it holds that

$$J^\varepsilon[U_n](t) = \sum_{i=1}^{k_n} \int_{[0,t]} H_n^{(i)}(s)\, \mathbb{1}\Big(\big|H_n^{(i)}(s)\big| \geq \varepsilon\Big)\, \mathrm{d}N_n^{(i)}$$

and

$$A^\varepsilon[U_n](t) = \sum_{i=1}^{k_n} \int_{[0,t]} H_n^{(i)}(s)\, \mathbb{1}\Big(\big|H_n^{(i)}(s)\big| \geq \varepsilon\Big)\, \mathrm{d}A_n^{(i)},$$

cf. Fleming and Harrington [19, Theorem 2.4.1]. Setting $U_{n,1}^\varepsilon = J^\varepsilon[U_n] - A^\varepsilon[U_n]$ and $U_{n,2}^\varepsilon = U_n - U_{n,1}^\varepsilon$, one sees with the same arguments as above that $\langle U_{n,1}^\varepsilon, U_{n,2}^\varepsilon \rangle(t) = 0$ for all $t \in \mathbb{R}_+$ and that $\langle U_{n,1}^\varepsilon \rangle(t)$ is the left hand side of (2.2).

A.2 Proof of Corollary 2.1.6

The proof is a slight modification of a proof given in Fleming and Harrington [19, Corollary 3.4.1]. Lemma 2.2.3 in Fleming and Harrington [19] enables us to choose a localization sequence $\{\tau_1, \tau_2, \ldots\}$, such that for any $k \in \mathbb{N}$ we have the processes $N^{(i)}(\cdot \wedge \tau_k)$, $A^{(i)}(\cdot \wedge \tau_k)$ and $H^{(i)}(\cdot \wedge \tau_k)$, $i = 1, \ldots, n$, are bounded by k. Note that the processes $A^{(i)}$ are always locally bounded. $M^{(i)}(\cdot \wedge \tau_k)$ is a square integrable martingale. Theorem 1.5.1 in Fleming and Harrington [19] yields that the processes $\int_{[0,t\wedge\tau_k]} H^{(i)}(s)\,\mathrm{d}M^{(i)}(s)$, $i = 1, \ldots, n$, are martingales. Because of the linearity of the conditional expectation we get that the process $\sum_{i=1}^n \int_{[0,t\wedge\tau_k]} H^{(i)}(s)\,\mathrm{d}M^{(i)}(s)$ is a martingale. Let T be a bounded stopping time then the Optional Stopping Theorem, Theorem 2.4.2 and Theorem 2.5.2 in Fleming and Harrington [19] give that

$$\mathbb{E}\big(X_k(t \wedge T) - Y_k(t \wedge T)\big) = 0, \qquad \text{for any } t > 0, \tag{A.1}$$

where

$$X_k(t) = \left(\sum_{i=1}^n \int_{[0,t\wedge\tau_k]} H^{(i)}(s)\,\mathrm{d}M^{(i)}(s)\right)^2$$

and

$$Y_k(t) = \sum_{i=1}^n \int_{[0,t\wedge\tau_k]} \big(H^{(i)}(s)\big)^2\,\mathrm{d}A^{(i)}(s).$$

It holds that $X_k(t \wedge T) \to X_k(T)$ and $Y_k(t \wedge T) \uparrow Y_k(T)$, as $t \to \infty$. Hence, the Dominated Convergence Theorem and the Monotone Convergence Theorem give that $\mathbb{E}\big(X_k(t \wedge T)\big) \to \mathbb{E}\big(X_k(T)\big)$ and $\mathbb{E}\big(Y_k(t \wedge T)\big) \to \mathbb{E}\big(Y_k(T)\big)$, where $\mathbb{E}\big(X_k(T)\big)$ and $\mathbb{E}\big(Y_k(T)\big)$ are finite. By (A.1) we get $\mathbb{E}\big(X_k(T)\big) = \mathbb{E}\big(Y_k(T)\big)$. Applying Theorem 2.1.5 yields that

$$p_{1,k} = \mathbb{P}\left(\sup_{0\le t\le T} X_k(t) \ge \varepsilon\right) \le \frac{\eta}{\varepsilon} + \mathbb{P}\big(Y_k(t) \ge \eta\big) = \frac{\eta}{\varepsilon} + p_{2,k}.$$

The Monotone Convergence Theorem gives that

$$p_{2,k} \uparrow \mathbb{P}\left(\sum_{i=1}^n \int_{[0,T]} \big(H^{(i)}(s)\big)^2\,\mathrm{d}A^{(i)}(s) \ge \eta\right) = p_2, \quad \text{as } k \to \infty,$$

so for every $k \in \mathbb{N}$ we have that $p_{1,k} \le \frac{\eta}{\varepsilon} + p_2$. The Monotone Convergence Theorem finally implies

$$p_{1,k} \to \mathbb{P}\left(\sup_{0\le t\le T}\left(\sum_{i=1}^n \int_{[0,t]} H^{(i)}(s)\,\mathrm{d}M^{(i)}(s)\right)^2 \ge \varepsilon\right), \quad \text{as } k \to \infty.$$

A.3 Proof of Theorem 2.2.7

Note that Proposition B.5.3 guarantees that the processes $\{f_{n,\xi}^{(i)}(t) \mid t \in \mathbb{R}\}$, $i = 1, \ldots, n$, are predictable and locally bounded. Proposition B.5.3 is used implicitly in the proof to come several times.

First, we prove the asymptotic expansion of $\log \Upsilon_{n,\xi}(t)$, $t \in [0, \tau]$. A Taylor-expansion gives that

$$\sum_{i=1}^{k_n} \int_{I(t)} \log\left(\frac{\alpha_{n,\xi}^{(i)}(s)}{\alpha_{n,0}^{(i)}(s)}\right) \mathrm{d}N_n^{(i)}(s) = \\ 2\sum_{i=1}^{k_n} \int_{I(t)} f_{n,\xi}^{(i)}(s)\, \mathrm{d}N_n^{(i)}(s) - \sum_{i=1}^{k_n} \int_{I(t)} r\big(f_{n,\xi}^{(i)}(s)\big)\, \big(f_{n,\xi}^{(i)}(s)\big)^2\, \mathrm{d}N_n^{(i)}(s),$$

where $r(x) = (1+\theta(x))^{-2}$ with $|\theta(x)| \in \big[0, |x|\big]$ is the remainder of the Taylor-expansion. Using this result, Jacod's Formula, cf. Proposition 2.2.5, gives that

$$\log \Upsilon_{n,\xi}(t) = -\sum_{i=1}^{k_n} \int_{I(t)} \left(\frac{\alpha_{n,\xi}^{(i)}(s)}{\alpha_{n,0}^{(i)}(s)} - 1\right) \lambda_{n,0}^{(i)}(s)\, \mathrm{d}s \\ + 2\sum_{i=1}^{k_n} \int_{I(t)} f_{n,\xi}^{(i)}(s)\, \mathrm{d}N_n^{(i)}(s) - \sum_{i=1}^{k_n} \int_{I(t)} r\big(f_{n,\xi}^{(i)}(s)\big)\, \big(f_{n,\xi}^{(i)}(s)\big)^2\, \mathrm{d}N_n^{(i)}(s).$$

Adding and subtracting the following terms $\sum_{i=1}^{k_n} \int_{I(t)} \big(f_{n,\xi}^{(i)}(s)\big)^2\, \mathrm{d}N_n^{(i)}(s)$ and $\sum_{i=1}^{k_n} \int_{I(t)} \big(f_{n,\xi}^{(i)}(s)\big)^2\, \lambda_{n,0}^{(i)}(s)\, \mathrm{d}s$ and $2\sum_{i=1}^{k_n} \int_{I(t)} f_{n,\xi}^{(i)}(s)\, \lambda_{n,0}^{(i)}(s)\, \mathrm{d}s$ yields

$$\log \Upsilon_{n,\xi}(t) = -2\sum_{i=1}^{k_n} \int_{I(t)} \big(f_{n,\xi}^{(i)}(s)\big)^2\, \lambda_{n,0}^{(i)}(s)\, \mathrm{d}s \\ +2\sum_{i=1}^{k_n} \int_{I(t)} \big(f_{n,\xi}^{(i)}(s)\big)\, \mathrm{d}M_{n,0}^{(i)}(s) - \sum_{i=1}^{k_n} \int_{I(t)} \big(f_{n,\xi}^{(i)}(s)\big)^2\, \mathrm{d}M_{n,0}^{(i)}(s) \\ +\sum_{i=1}^{k_n} \int_{I(t)} \Big(1 - r\big(f_{n,\xi}^{(i)}(s)\big)\Big)\, \big(f_{n,\xi}^{(i)}(s)\big)^2\, \mathrm{d}N_n^{(i)}(s) \\ = T_{n,\xi}^{(1)}(t) + T_{n,\xi}^{(2)}(t) - T_{n,\xi}^{(3)}(t) + T_{n,\xi}^{(4)}(t).$$

Note that $T^{(i)}_{n,\xi}(\infty) = \lim_{t\to\infty} T^{(i)}_{n,\xi}(t)$ almost surely, $i = 1, \ldots, 4$, which is mainly relevant for considering the case that $\tau = \infty$.

In the next step, (2.5) is proved. Let us consider $T^{(1)}_{n,\xi}$ and use the abbreviation $g_\xi(t) = 2^{-1}\mathscr{J}_\xi(t)$. For every $\varepsilon > 0$ there exists a $k \in \mathbb{N}$ and $0 = t_0 < \ldots < t_k = \tau$, such that $g_\xi(t_i) - g_\xi(t_{i-1}) < \varepsilon/2$, $i = 1, \ldots, k$. It holds the estimate

$$T^{(1)}_{n,\xi}(t_{i-1}) + g_\xi(t_{i-1}) + \frac{\varepsilon}{2} \geq T^{(1)}_{n,\xi}(t) + g_\xi(t) \geq T^{(1)}_{n,\xi}(t_i) + g_\xi(t_i) - \frac{\varepsilon}{2}$$

for all $t \in [t_{i-1}, t_i]$. Consequently, we receive that

$$\begin{aligned} P_{n,0}\left(\sup_{t\in I(\tau)}\left|T^{(1)}_{n,\xi}(t) + g_\xi(t)\right| \geq \varepsilon\right) &\leq P_{n,0}\left(\max_{0\leq i\leq k}\left|T^{(1)}_{n,\xi}(t_i) + g_\xi(t_i)\right| \geq \frac{\varepsilon}{2}\right) \\ &\leq \sum_{i=1}^{k} P_{n,0}\left(\left|T^{(1)}_{n,\xi}(t_i) + g_\xi(t_i)\right| \geq \frac{\varepsilon}{2}\right) \longrightarrow 0, \end{aligned}$$

as $n \to \infty$, where we use (2.3).

Now, we show that

$$\limsup_{n\to\infty} P_{n,0}\left(\sup_{t\in I(\tau)} |T^{(3)}_{n,\xi}(t)| \geq \varepsilon\right) \leq 4\eta\varepsilon^{-1} + 4\eta\varepsilon^{-2},$$

for all $\varepsilon, \eta > 0$. Then as η is arbitrary, it follows $P_{n,0}\big(\sup_{t\in[0,\tau]}|T^{(3)}_{n,\xi}(t)| \geq \varepsilon\big) \to 0$ as $n \to \infty$. Choose δ, such that $\eta\delta^{-2} > 2^{-1}g_\xi(\tau)$. It holds the estimate

$$\begin{aligned} &P_{n,0}\left(\sup_{t\in I(\tau)} |T^{(3)}_{n,\xi}| \geq \varepsilon\right) \\ &\leq P_{n,0}\left(\sup_{t\in I(\tau)}\left|\sum_{i=1}^{k_n}\int_{I(t)} \big(f^{(i)}_{n,\xi}(s)\big)^2 \mathbb{1}\Big(\left|f_{n,\xi}(s)\right| > \delta\Big)\,\mathrm{d}M^{(i)}_{n,0}(s)\right| \geq \frac{\varepsilon}{2}\right) \\ &\quad + P_{n,0}\left(\sup_{t\in I(\tau)}\left|\sum_{i=1}^{k_n}\int_{I(t)} \big(f^{(i)}_{n,\xi}(s)\big)^2 \mathbb{1}\Big(\left|f_{n,\xi}(s)\right| \leq \delta\Big)\,\mathrm{d}M^{(i)}_{n,0}(s)\right| \geq \frac{\varepsilon}{2}\right) \\ &= p_{n,1} + p_{n,2}. \end{aligned}$$

Moreover, one sees that

$$p_{n,1} \leq P_{n,0}\left(\sum_{i=1}^{k_n} \int_{I(\tau)} \big(f_{n,\xi}^{(i)}(s)\big)^2 \mathbb{1}\Big(\big|f_{n,\xi}(s)\big| > \delta\Big)\, \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s \geq \frac{\varepsilon}{4} \right)$$
$$+ P_{n,0}\left(\sup_{t\in I(\tau)} \sum_{i=1}^{k_n} \int_{I(t)} \big(f_{n,\xi}^{(i)}(s)\big)^2 \mathbb{1}\Big(\big|f_{n,\xi}(s)\big| > \delta\Big)\, \mathrm{d}N_n^{(i)}(s) \geq \frac{\varepsilon}{4} \right)$$

The first summand asymptotically vanishes because of (2.4). Note that

$$X_{n,\xi}(t\wedge\tau) = \sum_{i=1}^{k_n} \int_{I(t\wedge\tau)} \big(f_{n,\xi}^{(i)}(s)\big)^2 \mathbb{1}\Big(\big|f_{n,\xi}(s)\big| > \delta\Big)\, \mathrm{d}N_n^{(i)}(s), \quad t \geq 0,$$

is Lenglart-dominated by

$$Y_{n,\xi}(t\wedge\tau) = \sum_{i=1}^{k_n} \int_{I(t\wedge\tau)} \big(f_{n,\xi}^{(i)}(s)\big)^2 \mathbb{1}\Big(\big|f_{n,\xi}(s)\big| > \delta\Big)\, \lambda_{n,\xi}^{(i)}(s)\,\mathrm{d}s, \quad t \geq 0,$$

since $\big\{(X_{n,\xi} - Y_{n,\xi})(t\wedge\tau) \mid t \geq 0\big\}$ is a local martingale, cf. Jacod and Shiryaev [32, Theorem I.3.18]. A similar technique is used in the proof of Corollary 2.1.6. Therefore Theorem 2.1.5 gives that

$$P_{n,0}\left(\sup_{t\in I(\tau)} X_{n,\xi}(t\wedge\tau) \geq \frac{\varepsilon}{4} \right) \leq \frac{4\eta}{\varepsilon} + P_{n,0}\big(Y_{n,\xi}(\tau) \geq \eta\big) \longrightarrow \frac{4\eta}{\varepsilon}, \tag{A.2}$$

as $n \to \infty$, because of (2.4). Because of Corollary 2.1.6 and 2.1.7, it holds that

$$p_{n,2} \leq \frac{4\eta}{\varepsilon^2} + P_{n,0}\left(\sum_{i=1}^{k_n} \int_{I(\tau)} \big(f_{n,\xi}^{(i)}(s)\big)^2 \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s \geq \frac{\eta}{\delta^2} \right) \leq \frac{4\eta}{\varepsilon^2} +$$
$$P_{n,0}\left(\left| \sum_{i=1}^{k_n} \int_{I(\tau)} \big(f_{n,\xi}^{(i)}(s)\big)^2 \lambda_{n,0}^{(i)}(s)\,\mathrm{d}s - \frac{1}{2} g_\xi(\tau) \right| \geq \frac{\eta}{\delta^2} - \frac{1}{2} g_\xi(\tau) \right), \tag{A.3}$$

where the second term of the right hand side tends to 0 as $n \to \infty$, because of (2.3).

Let us consider $T_{n,\xi}^{(4)}(t)$. Obviously, it holds that

$$\begin{aligned}
T_{n,\xi}^{(4)}(t) &= \sum_{i=1}^{k_n} \int_{I(t)} \Big(1 - r\big(f_{n,\xi}^{(i)}(s)\big)\Big) \, \big(f_{n,\xi}^{(i)}(s)\big)^2 \mathbb{1}\Big(\big|f_{n,\xi}^{(i)}(s)\big| \le \delta\Big) \, \mathrm{d}N_n^{(i)}(s) \\
&\quad + \sum_{i=1}^{k_n} \int_{I(t)} \Big(1 - r\big(f_{n,\xi}^{(i)}(s)\big)\Big) \, \big(f_{n,\xi}^{(i)}(s)\big)^2 \mathbb{1}\Big(\big|f_{n,\xi}^{(i)}(s)\big| > \delta\Big) \, \mathrm{d}N_n^{(i)}(s) \\
&= T_{n,\xi}^{(4,1)}(t) + T_{n,\xi}^{(4,2)}(t)
\end{aligned}$$

for all $\delta > 0$. Because of (A.2), we get that

$$\begin{aligned}
&P_{n,0}\Big(\sup_{t\in I(\tau)} \big|T_{n,\xi}^{(4,2)}(t)\big| \ge \varepsilon\Big) \\
&\quad \le P_{n,0}\left(\sup_{t\in I(\tau)} \max_{1\le i\le k_n} \int_{I(t)} \big(f_{n,\xi}^{(i)}(s)\big)^2 \mathbb{1}\Big(\big|f_{n,\xi}^{(i)}(s)\big| > \delta\Big) \, \mathrm{d}N_n^{(i)}(s) > \delta^2\right) \\
&\quad \le P_{n,0}\left(\sup_{t\in I(\tau)} \sum_{i=1}^{k_n} \int_{I(t)} \big(f_{n,\xi}^{(i)}(s)\big)^2 \mathbb{1}\Big(\big|f_{n,\xi}^{(i)}(s)\big| > \delta\Big) \, \mathrm{d}N_n^{(i)}(s) > \delta^2\right) \to 0.
\end{aligned}$$

For all $|x| \le \delta < 1$ it holds that

$$\big|1 - r(x)\big| = \left|\frac{\theta^2(x) + 2\theta(x)}{(1+\theta(x))^2}\right| \le \frac{3\delta}{(1-\delta)^2} = \eta(\delta), \quad \big|\theta(x)\big| \in \big[0, |x|\big].$$

For sufficiently small $\delta > 0$, we get the estimate

$$\begin{aligned}
&P_{n,0}\big(\sup_{t\in I(\tau)} \big|T_{n,\xi}^{(4,1)}(t)\big| \ge \varepsilon\big) \\
&\quad \le P_{n,0}\left(\sup_{t\in I(\tau)} \sum_{i=1}^{k_n} \int_{I(t)} \big(f_{n,\xi}^{(i)}(s)\big)^2 \mathbb{1}\Big(\big|f_{n,\xi}^{(i)}(s)\big| \le \delta\Big) \, \mathrm{d}N_n^{(i)}(s) \ge \frac{\varepsilon}{\eta(\delta)}\right) \\
&\quad \le P_{n,0}\left(\sup_{t\in I(\tau)} \left|\sum_{i=1}^{k_n} \int_{I(t)} \big(f_{n,\xi}^{(i)}(s)\big)^2 \, \mathbb{1}\Big(\big|f_{n,\xi}^{(i)}(s)\big| \le \delta\Big) \, \mathrm{d}M_{n,0}^{(i)}(s)\right| \ge \frac{\varepsilon}{2\eta(\delta)}\right) \\
&\qquad + P_{n,0}\left(\left|\sum_{i=1}^{k_n} \int_{I(\tau)} \big(f_{n,\xi}^{(i)}(s)\big)^2 \, \lambda_{n,0}^{(i)}(s) \, \mathrm{d}s - \frac{1}{2}\, g_\xi(\tau)\right| \ge \frac{\varepsilon}{2\eta(\delta)} - \frac{1}{2}\, g_\xi(\tau)\right) \\
&\quad = p_{n,3} + p_{n,4}.
\end{aligned}$$

Note that one can always choose δ, such that $\varepsilon\big(2\eta(\delta)\big)^{-1} - \frac{1}{2}\, g_\xi(\tau) > 0$. $p_{n,4} \to 0$, because of (2.3). $p_{n,3} \to 0$ is proved completely analogously to $p_{n,2} \to 0$. Hence, the proof of (2.5) is complete.

Equation (2.8) is an immediate consequence of equation (2.5), where we use the fact that

$$\big|\log \Upsilon_{n,\xi}(\tau) + g_\xi(\tau) - T^{(2)}_{n,\xi}(\tau)\big| \le \sup_{t \in I(\tau)} \big|\log \Upsilon_{n,\xi}(t) + g_\xi(t) - T^{(2)}_{n,\xi}(t)\big|.$$

Let be $c \in \mathbb{R}^r$ and $\xi_j \in \mathbb{R}^m$, $j = 1, \ldots, r$. Consider the process

$$\big\{U_n(t) \mid t \in \mathbb{R}_+\big\}, \qquad \text{where } U_n(t) = -\sum_{j=1}^r c^{(j)}\, T^{(2)}_{n,\xi_j}(t).$$

We want to apply Corollary 2.1.3. Because of (2.3) and (2.4), it holds that

$$\sum_{i=1}^{k_n} \int_{I(t)} \left(2 \sum_{j=1}^r c^{(i)} f^{(i)}_{n,\xi_j}(s)\right)^2 \lambda^{(i)}_{n,\xi}(s)\,\mathrm{d}s \longrightarrow_{P_{n,0}} \sum_{i=1}^r \sum_{j=1}^r c^{(i)}\, c^{(j)}\, \xi_i^{\mathrm{T}} \mathscr{I}(t) \xi_j,$$

as $n \to \infty$, and

$$\begin{aligned}
&\sum_{i=1}^{k_n} \int_{I(t)} \left(2 \sum_{j=1}^r c^{(j)} f^{(i)}_{n,\xi_j}(s)\right)^2 \mathbb{1}\left(\left|\sum_{j=1}^r c^{(j)} f^{(i)}_{n,\xi_j}(s)\right| \ge \varepsilon/2\right) \lambda^{(i)}_{n,0}(s)\,\mathrm{d}s \\
&\le 4r \sum_{i=1}^{k_n} \sum_{j=1}^r \sum_{k=1}^r \int_{I(t)} \big(c^{(j)}\, f^{(i)}_{n,\xi_j}(s)\big)^2 \mathbb{1}\Big(\big|c^{(k)}\, f^{(i)}_{n,\xi_k}(s)\big| \ge \varepsilon/(2r)\Big)\, \lambda^{(i)}_{n,0}(s)\,\mathrm{d}s \\
&\le 8r^2 \sum_{j=1}^r \sum_{i=1}^{k_n} \int_{I(t)} \big(f^{(i)}_{n,\xi_j}(s)\big)^2 \mathbb{1}\Big(\big|f^{(i)}_{n,\xi_j}(s)\big| \ge \varepsilon/r\Big)\, \lambda^{(i)}_{n,0}(s)\,\mathrm{d}s \longrightarrow_{P_{n,0}} 0,
\end{aligned}$$

where the estimates $\big(\sum_{j=1}^r a_j\big)^2 \le r \sum_{j=1}^r a_j^2$ and

$$\begin{aligned}
&\big(c^{(j)}\, f^{(i)}_{n,\xi_j}\big)^2 \mathbb{1}\big(|c^{(k)}\, f^{(i)}_{n,\xi_k}| \ge \varepsilon/(2r)\big) \\
&\qquad = (c^{(j)}\, f^{(i)}_{n,\xi_j})^2 \mathbb{1}\big(|c^{(k)}\, f^{(i)}_{n,\xi_k}| \ge \varepsilon/(2r)\big)\, \mathbb{1}\big(|c^{(j)}\, f^{(i)}_{n,\xi_j}| \ge \varepsilon/(2r)\big) \\
&\qquad\qquad + (c^{(j)}\, f^{(i)}_{n,\xi_j})^2 \mathbb{1}\big(c^{(k)}\, |f^{(i)}_{n,\xi_k}| \ge \varepsilon/(2r)\big)\, \mathbb{1}\big(|c^{(j)}\, f^{(i)}_{n,\xi_j}| < \varepsilon/(2r)\big) \\
&\qquad \le (c^{(k)}\, f^{(i)}_{n,\xi_k})^2 \mathbb{1}\big(|c^{(k)}\, f^{(i)}_{n,\xi_k}| \ge \varepsilon/(2r)\big) \\
&\qquad\qquad + (c^{(j)}\, f^{(i)}_{n,\xi_j})^2 \mathbb{1}\big(|c^{(j)}\, f^{(i)}_{n,\xi_j}| \ge \varepsilon/(2r)\big)
\end{aligned}$$

were used. Thus, the conditions (2.1) and (2.2) hold. Corollary 2.1.3 yields

$$\{U_n(t\wedge\tau)\mid t\in\mathbb{R}_+\} \xrightarrow{\mathfrak{D}}_{P_{n,0}} \left\{\mathbb{W}\circ\sum_{i=1}^{r}\sum_{j=1}^{r} c^{(i)}c^{(j)}\xi_i^{\mathrm{T}}\mathscr{I}(t\wedge\tau)\xi_j \,\middle|\, t\in\mathbb{R}_+\right\}, \quad (\mathrm{A}.4)$$

as $n\to\infty$, on $D(\mathbb{R}_+,\mathbb{R})$. Choosing $r=1$ and $c=1$ one sees that (2.6) holds. Applying Jacod and Shiryaev [32, Proposition VI.3.17] yields that

$$X_n = \left\{2\sum_{i=1}^{k_n}\int_{I(t\wedge\tau)} f_{n,\xi}^{(i)}(s)\,\mathrm{d}M_{n,0}^{(i)} - \; g_\xi(t\wedge\tau) \,\middle|\, t\in\mathbb{R}_+\right\}$$
$$\xrightarrow{\mathfrak{D}}_{P_{n,0}} \{\mathbb{W}\circ\xi^{\mathrm{T}}\mathscr{I}(t\wedge\tau)\xi - g_\xi(t\wedge\tau)\mid t\in\mathbb{R}_+\}, \quad (\mathrm{A}.5)$$

as $n\to\infty$. Let d denote the metric on $D(\mathbb{R}_+,\mathbb{R})$ defined in Jacod and Shiryaev [32, Formula VI.1.26]. d generates the Skorokhod topology and makes $D(\mathbb{R}_+,\mathbb{R})$ a Polish space. (Note that there are metrics on $D(\mathbb{R}_+,\mathbb{R})$ that generate the Skorokhod topology, but fail to make $D(\mathbb{R}_+,\mathbb{R})$ a complete space, see Jacod and Shiryaev [32, Remark VI.1.27].) Looking at the definition of the metric, one sees that

$$d\Big(\{\log\Upsilon_{n,\xi}(t\wedge\tau)\mid t\in\mathbb{R}_+\},\, X_n\Big)$$
$$\le \sup_{t\in I(\tau)}\left|\log\Upsilon_{n,\xi}(t) + \; g_\xi(t\wedge\tau) - 2\sum_{i=1}^{k_n}\int_{I(t)} f_{n,\xi}^{(i)}(s)\,\mathrm{d}M_{n,0}^{(i)}\right| \longrightarrow_{P_{n,0}} 0,$$

as $n\to\infty$, where we use (2.5). Applying (A.5) and Slutsky's Lemma, cf. Billingsley [9, Theorem 4.1], yield (2.7).

In the case $\tau<\infty$, equation (2.9) follows directly from equation (A.4) by setting $r=1$ and $c=1$ and using Proposition VI.3.14 in Jacod and Shiryaev [32]. Analogously, one sees that

$$U_n(\tau) - \sum_{i=1}^{r} c^{(i)}g_{\xi_i}(\tau) \xrightarrow{\mathfrak{D}}_{P_{n,0}} \mathcal{N}\left(-\sum_{i=1}^{r} c^{(i)}g_{\xi_i}(\tau),\;\; c^{\mathrm{T}}\mathscr{S}(\tau)c\right),$$

as $n\to\infty$, $c=(c^{(1)},\ldots,c^{(r)})$, where we use Witting and Müller-Funk [72, Satz 5.83] and (A.4). Applying (2.8) and the Cramér-Wold device, cf. Billingsley [9, Theorem 7.7], gives (2.10).

Now, let us consider the case $\tau = \infty$. We want to apply Theorem 2.1.1. Assume that τ_k, $k \in \mathbb{N}$, is a sequence of real numbers, such that $\lim_{k\to\infty} \tau_k = \infty$. Defining the following processes

$$\begin{aligned}
X_{n,k} &= \big\{ U_n(t \wedge \tau_k) \mid t \in [0,\infty] \big\}, \\
X_k &= \left\{ \mathbb{W} \circ \sum_{i=1}^{r} \sum_{j=1}^{r} c^{(i)}\, c^{(j)} \xi_i^{\mathrm{T}} \mathscr{I}(t \wedge \tau_k) \xi_j \;\middle|\; t \in [0,\infty] \right\}, \\
X &= \left\{ \mathbb{W} \circ \sum_{i=1}^{r} \sum_{j=1}^{r} c^{(i)}\, c^{(j)} \xi_i^{\mathrm{T}} \mathscr{I}(t) \xi_j \;\middle|\; t \in [0,\infty] \right\}, \\
\widetilde{X}_n &= \big\{ U_n(t) \mid t \in [0,\infty] \big\},
\end{aligned}$$

it holds that $X_{n,k} \xrightarrow{\mathfrak{D}}_{P_{n,0}} X_k$, as $n \to \infty$, and $X_k \xrightarrow{\mathfrak{D}}_{P_{n,0}} X$, as $n \to \infty$, on $D\big([0,\infty],\mathbb{R}\big)$, see also the remarks on page 26. Therefore, it remains to be proven that

$$\lim_{k\to\infty} \limsup_{n\to\infty} P_{n,0}\big(\widetilde{d}(\widetilde{X}_n, X_{n,k}) \geq \varepsilon\big) = 0 \quad \text{for all } \varepsilon > 0, \tag{A.6}$$

where $\widetilde{d}$ denotes a metric that generates the Skorokhod topology and such that ensures that $D\big([0,\infty],\mathbb{R}\big)$ is a Polish space. For example choose $\widetilde{d}(x,y) = d_0(x \circ T^{-1}, y \circ T^{-1})$, where d_0 is the metric generating the Skorokhod topology on $D\big([0,1],\mathbb{R}_+\big)$ and making $D\big([0,1],\mathbb{R}_+\big)$ a Polish space, cf. Billingsley [9, pp. 112], and $T : [0,1] \longrightarrow [0,\infty]$, $T(t) = t(1-t)^{-1}$, $t \in [0,1)$, and $T(1) = \infty$.

Note that

$$\widetilde{d}(\widetilde{X}_n, X_{n,k}) \leq \sup_{t\in[0,\infty)} \big|\widetilde{X}_n(t) - X_{n,k}(t)\big|.$$

Let $\eta > 0$ be arbitrary and set $h_{k,i} = \eta - \frac{1}{4}\big(c^{(i)}\big)^2 \xi_i^{\mathrm{T}} \big(\mathscr{I}(\infty) - \mathscr{I}(\tau_k)\big)\xi_i$. There exists a $k_0 \in \mathbb{N}$, such that $h_{k,i} > 0$ for all $i = 1,\ldots,r$ and all $k \geq k_0$.

$$P_{n,0}\big(\widetilde{d}(\widetilde{X}_n, X_{n,k}) \geq \varepsilon\big) \leq P_{n,0}\left(\sup_{t\in[0,\infty)} \big|\widetilde{X}_n(t) - X_{n,k}(t)\big| \geq \varepsilon \right) \leq$$

$$\sum_{i=1}^{r} P_{n,0}\Bigg(\big|c^{(i)}\big| \sup_{t\in[0,\infty)} \big|T^{(2)}_{n,\xi_i}(t) - T^{(2)}_{n,\xi_i}(\tau_k \wedge t)\big| \ge \frac{\varepsilon}{r}\Bigg)$$
$$\le K + \sum_{i=1}^{r} P_{n,0}\Bigg(\big(c^{(i)}\big)^2 \sum_{j=1}^{k_n} \int_{(\tau_k,\infty)} \big(f^{(j)}_{n,\xi_i}(s)\big)^2\, \lambda^{(j)}_{n,0}(s)\,\mathrm{d}s \ge \eta\Bigg)$$
$$\le K + \sum_{i=1}^{r} P_{n,0}\Bigg(\big(c^{(i)}\big)^2 \Big|\sum_{j=1}^{k_n} \int_{(\tau_k,\infty)} \big(f^{(j)}_{n,\xi_i}(s)\big)^2\, \lambda^{(j)}_{n,0}(s)\,\mathrm{d}s + h_{k,i} - \eta\Big| \ge h_{k,i}\Bigg),$$

where we set $K = r^2\eta/(\varepsilon^2)$ and use Corollary 2.1.7. If $k \ge k_0$ we have that

$$P_{n,0}\Bigg(\big(c^{(i)}\big)^2 \Big|\sum_{j=1}^{n} \int_{(\tau_k,\infty)} \big(f^{(j)}_{n,\xi_i}(s)\big)^2\, \lambda^{(j)}_{n,0}(s)\,\mathrm{d}s + h_{k,i} - \eta\Big| \ge h_{k,i}\Bigg) \le$$
$$P_{n,0}\Bigg(\big(c^{(i)}\big)^2 \Big|\sum_{j=1}^{k_n} \int_{I(\infty)} \big(f^{(j)}_{n,\xi_i}(s)\big)^2\, \lambda^{(j)}_{n,0}(s)\,\mathrm{d}s - \frac{1}{4}\xi_i^{\mathrm{T}} \mathscr{I}(\infty)\xi_i\Big| \ge \frac{h_{k,i}}{2}\Bigg)$$
$$+ P_{n,0}\Bigg(\big(c^{(i)}\big)^2 \Big|\sum_{j=1}^{k_n} \int_{I(\tau_k)} \big(f^{(j)}_{n,\xi_i}(s)\big)^2\, \lambda^{(j)}_{n,0}(s)\,\mathrm{d}s - \frac{1}{4}\xi_i^{\mathrm{T}} \mathscr{I}(\tau_k)\xi_i\Big| \ge \frac{h_{k,i}}{2}\Bigg) \to 0,$$

as $n \to \infty$, for all $i = 1, \ldots, r$, see (2.3). A $\eta > 0$ was arbitrary, assertion (A.6) holds. We showed $\widetilde{X}_n \xrightarrow{\mathfrak{D}}_{P_{n,0}} X$ on $D\big([0,\infty], \mathbb{R}_+\big)$. Thus, it holds that

$$U_n(\infty) \xrightarrow{\mathfrak{D}}_{P_{n,0}} \mathcal{N}\Bigg(0,\ \sum_{i=1}^{r}\sum_{j=1}^{r} c^{(i)}\, c^{(j)} \xi_i^{\mathrm{T}} \mathscr{I}(\infty)\xi_j\Bigg).$$

(2.9) results, if one chooses $r = 1$ and $c = 1$. Applying (2.8), Witting and Müller Funk [72, Theorem 5.83] and the Cramér-Wold device, cf. Billingsley [9, Theorem 7.7], yield (2.10).

A.4 Proof of Theorem 4.2.1

The statistic S is sufficient for the distribution family $\big\{P_\xi \mid \xi \in \mathbb{R}^{q+r}\big\}$, see Witting [71, Satz 3.19]. Consequently, we only need to consider the induced distribution family $\big\{\mathcal{N}(\mathscr{I}\xi, \mathscr{I}) \mid \xi \in \mathbb{R}^{q+r}\big\}$, cf. Witting [71, Satz 3.30]. Therefore, without loss of generality, we can assume that $P_0 = \mathcal{N}(0, \mathscr{I})$ and S is the identity.

Proof of a). Let $\mathscr{I}^{1/2}$ be a positive semi-definite, symmetric matrix, such that $\mathscr{I}^{\frac{1}{2}}\mathscr{I}^{\frac{1}{2}} = \mathscr{I}$, see Proposition B.3.3.a. Abbreviating

$$\Pi = \mathscr{V}_1\big(\mathscr{V}_1^{\mathrm{T}}\mathscr{I}\mathscr{V}_1\big)^{-}\mathscr{V}_1^{\mathrm{T}} - \mathscr{V}_0\big(\mathscr{V}_0^{\mathrm{T}}\mathscr{I}\mathscr{V}_0\big)^{-}\mathscr{V}_0^{\mathrm{T}},$$

we see that

$$\mathscr{I}\Pi\mathscr{I}\Pi x = \mathscr{I}\Pi x \qquad \text{for all } x \in \mathrm{Im}(\mathscr{I}), \tag{A.7}$$

where we use Proposition B.2.4.c. Moreover, it holds that the matrix $\mathscr{A} = (\mathscr{I}\Pi)^{\mathrm{T}}\mathscr{I}^{-}(\mathscr{I}\Pi)$ is symmetric and self-adjoint with respect to the Euclidean inner product and that

$$\begin{aligned}
\mathscr{I}^{\frac{1}{2}}\mathscr{A}\mathscr{I}^{\frac{1}{2}}\mathscr{I}^{\frac{1}{2}}\mathscr{A}\mathscr{I}^{\frac{1}{2}} &= \mathscr{I}^{\frac{1}{2}}(\mathscr{I}\Pi)^{\mathrm{T}}\mathscr{I}^{-}(\mathscr{I}\Pi)\mathscr{I}(\mathscr{I}\Pi)^{\mathrm{T}}\mathscr{I}^{-}(\mathscr{I}\Pi)\mathscr{I}^{\frac{1}{2}}\\
&= \mathscr{I}^{\frac{1}{2}}(\mathscr{I}\Pi)^{\mathrm{T}}\mathscr{I}^{-}(\mathscr{I}\Pi)(\mathscr{I}\Pi)(\mathscr{I}\Pi)\mathscr{I}^{\frac{1}{2}}\\
&= \mathscr{I}^{\frac{1}{2}}(\mathscr{I}\Pi)^{\mathrm{T}}\mathscr{I}^{-}(\mathscr{I}\Pi)\mathscr{I}^{\frac{1}{2}}\\
&= \mathscr{I}^{\frac{1}{2}}\mathscr{A}\mathscr{I}^{\frac{1}{2}},
\end{aligned}$$

where we use equation (A.7) and that Π and $\mathscr{I}$ are symmetric. Hence, $\mathscr{I}^{\frac{1}{2}}\mathscr{A}\mathscr{I}^{\frac{1}{2}}$ is an orthogonal projection of rank l, see Eaton [18, Proposition 1.17]. Eaton [18, Proposition 3.8] yields that the statistic $T(S)$ is $\chi^2_l(\delta)$-distributed. We show that $l = \dim\big(\mathrm{Im}(\mathscr{I}\mathscr{V}_1)\big) - \dim\big(\mathrm{Im}(\mathscr{I}\mathscr{V}_0)\big)$. Using Proposition B.3.3.a we get that

$$\begin{aligned}
l &= \mathrm{rank}(\mathscr{I}^{\frac{1}{2}}\mathscr{A}\mathscr{I}^{\frac{1}{2}}) = \dim \mathrm{Im}(\mathscr{I}^{\frac{1}{2}}\mathscr{A}\mathscr{I}^{\frac{1}{2}}) = \dim \mathrm{Im}(\mathscr{I}\mathscr{A}\mathscr{I})\\
&= \dim \mathrm{Im}(\mathscr{I}\Pi\mathscr{I}) = \dim\big\{\mathscr{I}\Pi x \mid x \in \mathrm{Im}(\mathscr{I})\big\}\\
&= \dim\big\{\Pi_{\mathscr{V}_1}(x) \mid x \in \mathrm{Im}(\mathscr{I})\big\} - \big\{\Pi_{\mathscr{V}_0}(x) \mid x \in \mathrm{Im}(\mathscr{I})\big\}\\
&= \dim\big(\mathrm{Im}(\mathscr{I}\mathscr{V}_1)\big) - \dim\big(\mathrm{Im}(\mathscr{I}\mathscr{V}_0)\big).
\end{aligned}$$

The equivalence is an easy consequence of the projection properties of the statistic. Using the concept presented in Section B.2, especially Proposition B.2.5, one sees that

$$0 = T(\mathscr{I}\xi) = \big|\big|\Pi_{\mathcal{V}_1}(\mathscr{I}\xi) - \Pi_{\mathcal{V}_0}(\mathscr{I}\xi)\big|\big|^2_{\mathscr{I}^{-}} \iff \Pi_{\mathcal{V}_1}(\mathscr{I}\xi) = \Pi_{\mathcal{V}_0}(\mathscr{I}\xi).$$

As $\Pi_{\mathcal{V}_1}(\mathscr{J}\xi) = \mathscr{J}\xi$, the assertion of a) is proved.

Proof of b). One readily checks that $\mathfrak{Q}_0$ is indeed a group. The invariance of the testing problem is more or less obvious. The invariance of T can be seen as follows. Using the decomposition

$$s = \big(s - \Pi_{\mathcal{V}_1}(s)\big) + \big(\Pi_{\mathcal{V}_1}(s) - \Pi_{\mathcal{V}_0}(s)\big) + \Pi_{\mathcal{V}_0}(s), \qquad s \in \mathrm{Im}(\mathscr{J}),$$

we get for any $\pi = \mathscr{Q}(\cdot) + u \in \mathfrak{Q}_0$ that

$$\begin{aligned} &\Pi_{\mathcal{V}_1}(\pi\, s) - \Pi_{\mathcal{V}_0}(\pi\, s) \\ &\qquad = \Pi_{\mathcal{V}_1}\big(\mathscr{Q}\Pi_{\mathcal{V}_1}(s) - \mathscr{Q}\Pi_{\mathcal{V}_0}(s)\big) + \Pi_{\mathcal{V}_1}\big(\mathscr{Q}\Pi_{\mathcal{V}_0}(s)\big) - \Pi_{\mathcal{V}_0}\big(\mathscr{Q}\Pi_{\mathcal{V}_0}(s)\big) \\ &\qquad = \mathscr{Q}\Pi_{\mathcal{V}_1}(s) - \mathscr{Q}\Pi_{\mathcal{V}_0}(s), \end{aligned}$$

since $\mathscr{Q}\big(s - \Pi_{\mathcal{V}_1}(s)\big)$ is orthogonal on $\mathrm{Im}(\mathscr{J}\mathscr{V}_1)$ and $\mathscr{Q}\big(\Pi_{\mathcal{V}_1}(s) - \Pi_{\mathcal{V}_0}(s)\big)$ is orthogonal on $\mathrm{Im}(\mathscr{J}\mathscr{V}_0)$ with respect to $<\cdot,\cdot>_{\mathscr{J}^-}$, see also Proposition B.2.4. Applying $\mathscr{Q}^{\mathrm{T}}\mathscr{J}^-\mathscr{Q} = \mathscr{J}^-$, gives the first assertion $T(s) = T(\pi s)$.

Let $\mathscr{B}$ and $\mathscr{C}$ the matrices defined in Proposition B.3.3.b, *i.e.* it holds that

$$\mathscr{B}\mathscr{B}^{\mathrm{T}} = \mathscr{J}, \quad \mathscr{C}\mathscr{C}^{\mathrm{T}} = \mathscr{J}^-, \quad \text{and} \quad \mathscr{B}^{\mathrm{T}}\mathscr{C} = \mathscr{C}^{\mathrm{T}}\mathscr{B} = \mathscr{E}_{\mathrm{rank}(\mathscr{J})}$$

where $\mathscr{E}_l$ denotes the $(l \times l)$ unity-matrix. Assume that $x, y \in \mathrm{Im}\big(\mathscr{J}\big)$, such that $x^{\mathrm{T}}\mathscr{J}^-x = y^{\mathrm{T}}\mathscr{J}^-y$ and $x \neq y$. Set $w = \mathscr{C}^{\mathrm{T}}(x-y)/\sqrt{<x-y, x-y>_{\mathscr{J}^-}}$ and

$$\mathscr{H} = \mathscr{B}\big(\mathscr{E}_{\mathrm{rank}(\mathscr{J})} - 2ww^{\mathrm{T}}\big)\mathscr{C}^{\mathrm{T}}. \tag{A.8}$$

$\mathscr{E}_{\mathrm{rank}(\mathscr{J})} - 2ww^{\mathrm{T}}$ is a so-called Householder-Matrix, as $w^{\mathrm{T}}w = 1$. Using the basic properties of these matrices, cf. Stoer [67, pp. 181], one easily shows that $\mathscr{H}x = y$, $\mathscr{H}^{\mathrm{T}}\mathscr{J}^-\mathscr{H} = \mathscr{J}^-$, $\mathscr{H}\mathscr{J}\mathscr{H}^{\mathrm{T}} = \mathscr{J}$ and $z = \mathscr{H}z$, whenever $<z, x-y>_{\mathscr{J}^-} = 0$, $z \in \mathrm{Im}(\mathscr{J})$.

We have to show that $T(x) = T(y)$ implies that there exists a $\pi \in \mathfrak{Q}_0$, such that $\Pi_{\mathcal{V}_1}(y) = \Pi_{\mathcal{V}_1}(\pi x) = \pi\Pi_{\mathcal{V}_1}(x)$. Obviously, the matrices defined in equation (A.8) are helpful to construct such elements of the group. Using the decomposition

$$x = \big(x - \Pi_{\mathcal{V}_1}(x)\big) + \big(\Pi_{\mathcal{V}_1}(x) - \Pi_{\mathcal{V}_0}(x)\big) + \Pi_{\mathcal{V}_0}(x) = x_3 + x_2 + x_1$$

and

$$y = \big(y - \Pi_{\mathcal{V}_1}(y)\big) + \big(\Pi_{\mathcal{V}_1}(y) - \Pi_{\mathcal{V}_0}(y)\big) + \Pi_{\mathcal{V}_0}(y) = y_3 + y_2 + y_1.$$

$T(x) = T(y)$ means $x_2^{\mathrm{T}} \mathscr{I}^{-} x_2 = y_2^{\mathrm{T}} \mathscr{I}^{-} y_2$. By the previous considerations we know that there exists a matrix $\mathscr{H}$, such that $\mathscr{H} x_2 = y_2$ and $\pi(\cdot) = \mathscr{H}(\cdot) + (y_1 - x_1) \in \mathfrak{Q}_0$. Easy calculations give that

$$\Pi_{\mathcal{V}_1}(\pi\, x) = \Pi_{\mathcal{V}_1}\big(\mathscr{H}(x_3 + x_2 + x_1)\big) + (y_1 - x_1) = y_2 + y_1 = \Pi_{\mathcal{V}_1}(y)$$

and

$$\pi\, \Pi_{\mathcal{V}_1}(x) = \mathscr{H}(x_2 + x_1) + (y_1 - x_1) = y_2 + y_1 = \Pi_{\mathcal{V}_1}(y).$$

The proof of c) is straightforward. It holds that

$$\log \frac{\mathrm{d}P_\xi^{(\tau)}}{\mathrm{d}P_0^{(\tau)}} = \big\langle S - \Pi_{\mathcal{V}_1}(S), \mathscr{I}\xi \big\rangle_{\mathscr{I}^-} + \big\langle \Pi_{\mathcal{V}_1}(S), \mathscr{I}\xi \big\rangle_{\mathscr{I}^-} - \frac{1}{2}\big\langle \mathscr{I}\xi, \mathscr{I}\xi \big\rangle_{\mathscr{I}^-}.$$

As $\big\langle S - \Pi_{\mathcal{V}_1}(S), \mathscr{I}\xi \big\rangle_{\mathscr{I}^-} = 0$ for all $\xi \in \mathcal{H}_2^{\mathcal{L}_1}$, the statistic $\Pi_{\mathcal{V}_1}(S)$ is sufficient for the distribution family $\mathcal{N}\big(\mathscr{I}\mathscr{V}_1\kappa, \mathscr{I}\big)$, $\kappa \in \mathbb{R}^k$, cf. Witting [71, Satz 3.19]. Therefore, we can assume that $\mathcal{V}_1 = \mathrm{Im}\big(\mathscr{I}\big)$ without loss of generality. T is a maximal invariant statistic in the conventional sense. Choose $x \in \mathrm{Im}(\mathscr{I})$, such that $<x, x>_{\mathscr{I}^-} = 1$ and $<x, z>_{\mathscr{I}^-} = 0$ for all $z \in \mathcal{V}_0$ The mapping $T^-(s) = \sqrt{s}x$, satisfies $T \circ T^-(s) = s$ for all $s \in [0, \infty)$. Witting [71, Satz 3.91, Satz 3.92] yields that every invariant test φ is of the form $\varphi = \psi \circ T$. The fact that the class of $\chi_l^2(\delta)$-distributions, $\delta \geq 0$, has a monotone likelihood ratio in the identity implies the assertion, cf. Witting [71, Satz 2.24, Satz 2.36].

B Supplementary Results

B.1 Generalized Inverse

This section contains a small summary of the facts on generalized inverses used in the previous chapters.

B.1.1 Definition (Generalized Inverse). Let $\mathscr{A}$ be a real $(m \times r)$ matrix. Any real $(r \times m)$ matrix $\mathscr{B}$ that satisfies the conditions

i) $\mathscr{A}\mathscr{B}$ and $\mathscr{B}\mathscr{A}$ are symmetric,

ii) $\mathscr{A}\mathscr{B}\mathscr{A} = \mathscr{A}$,

iii) $\mathscr{B}\mathscr{A}\mathscr{B} = \mathscr{B}$,

is called the generalized inverse of $\mathscr{A}$.

B.1.2 Proposition (Existence and Uniqueness of the Generalized Inverse). For any real $(m \times r)$ matrix $\mathscr{A}$, there exists a uniquely determined matrix $\mathscr{B}$ satisfying the conditions of Definition B.1.1. This matrix is abbreviated $\mathscr{A}^-$.

Proof. Cf. Graybill [24, Theorem 6.2.1, Theorem 6.2.4]. □

B.1.3 Proposition (Relation to Inverse). Let $\mathscr{A}$ be a real $(m \times m)$ matrix with full rank. Then it holds that $\mathscr{A}^- = \mathscr{A}^{-1}$.

Proof. Cf. Graybill [24, Theorem 6.2.13]. □

B.1.4 Proposition (Inverse of Transposed Matrix). For any real $(m \times r)$ matrix $\mathscr{A}$ it holds that $(\mathscr{A}^{\mathrm{T}})^{-} = (\mathscr{A}^{-})^{\mathrm{T}}$. Especially, the generalized inverse of a symmetric matrix is also symmetric.

Proof. Cf. Graybill [24, Theorem 6.2.5] □

B.1.5 Proposition (Consistency of Linear Equations). The system of linear equations $\mathscr{A}x = b$ is consistent, if and only if $\mathscr{A}\mathscr{A}^{-}b = b$.

Proof. Cf. Graybill [24, Theorem 6.3.1]. □

B.1.6 Proposition (Generalized Inverse and Orthogonal Matrices). Let $\mathscr{A}$ be a real $(m \times m)$ matrix and let $\mathscr{F}$ be a real orthogonal $(m \times m)$ matrix, *i.e.* $\mathscr{F}^{\mathrm{T}}\mathscr{F} = \mathscr{F}\mathscr{F}^{\mathrm{T}} = \mathscr{E}_m$, where $\mathscr{E}_m$ denotes the $(m \times m)$ unity-matrix. It holds that $(\mathscr{F}^{\mathrm{T}}\mathscr{A}\mathscr{F})^{-} = \mathscr{F}^{\mathrm{T}}\mathscr{A}^{-}\mathscr{F}$.

Proof. Cf. Graybill [24, Theorem 6.2.10]. □

B.2 Projections in Hilbert Spaces

This section contains some results on projections in Hilbert spaces used in previous chapters.

B.2.1 Definition (Hilbert Space). Let $\big(\mathcal{V}, ||\cdot||\big)$ denote a real, complete, normed vector space. If there exists an inner product $\big\langle \cdot, \cdot \big\rangle : \mathcal{V} \times \mathcal{V} \longrightarrow \mathbb{R}$, *i.e.* a positive definite, symmetric, bilinear mapping, satisfying $||v|| = \sqrt{\langle v, v\rangle}$, $v \in \mathcal{V}$, then we call the tuple $\big(\mathcal{V}, \langle \cdot, \cdot \rangle\big)$ a real Hilbert space.

B.2.2 Definition (Closed, Convex Cone). Let $\big(\mathcal{V}, \langle \cdot, \cdot \rangle\big)$ be some real Hilbert space. A set $\mathcal{V}_0 \subset \mathcal{V}$ is called closed, convex cone, if

(i) $\mathcal{V}_0$ is closed in the $||\cdot||$-topology,

(ii) $v_1, v_2 \in \mathcal{V}_0$ and $\alpha \in (0,1)$ imply that $\alpha v_1 + (1-\alpha)v_2 \in \mathcal{V}_0$,

(iii) $v \in \mathcal{V}_0$ and $\alpha \geq 0$ imply that $\alpha v \in \mathcal{V}_0$.

B.2.3 Proposition (Characterization of Projections). Let $\big(\mathcal{V}, \langle\cdot,\cdot\rangle\big)$ be some real Hilbert space and $\mathcal{V}_0$ be a closed, convex cone.

a) For every $v \in \mathcal{V}$ there exists a unique element $\Pi_{\mathcal{V}_0}(v) \in \mathcal{V}_0$, such that

$$||v - \Pi_{\mathcal{V}_0}(v)|| = \inf_{\widetilde{v} \in \mathcal{V}_0} ||v - \widetilde{v}||.$$

$\Pi_{\mathcal{V}_0}(v)$ is called the projection of v on $\mathcal{V}_0$.

b) The projection $\Pi_{\mathcal{V}_0}(v)$ is uniquely determined by the conditions

$$\big\langle \Pi_{\mathcal{V}_0}(v), v \big\rangle = ||\Pi_{\mathcal{V}_0}(v)||^2$$

and

$$\big\langle v, \widetilde{v} \big\rangle \leq \big\langle \Pi_{\mathcal{V}_0}(v), \widetilde{v} \big\rangle, \qquad \widetilde{v} \in \mathcal{V}_0.$$

Proof. Cf. Behnen and Neuhaus [7, Section 7.2]. □

B.2.4 Proposition (Properties of Projections). Let $\big(\mathcal{V}, \langle\cdot,\cdot\rangle\big)$ be some real Hilbert space and $\mathcal{V}_0$ and $\mathcal{V}_1$ be closed, convex cones, such that $\mathcal{V}_0 \subset \mathcal{V}_1$.

a) For all $\alpha \geq 0$ it holds that $\Pi_{\mathcal{V}_0}(\alpha v) = \alpha \Pi_{\mathcal{V}_0}(v)$.

b) $\Pi_{\mathcal{V}_0}\big(v - \Pi_{\mathcal{V}_0}(v)\big) = 0$.

c) $\Pi_{\mathcal{V}_0}(v) = v$ for all $v \in \mathcal{V}_0$.

d) If $\mathcal{V}_0$ is a linear subspace then $\big\langle v - \Pi_{\mathcal{V}_0}(v), \widetilde{v} \big\rangle = 0$,for all $\widetilde{v} \in \mathcal{V}_0$.

e) If $\mathcal{V}_0$ is a linear subspace then $\Pi_{\mathcal{V}_0}(v_1 + v_2) = \Pi_{\mathcal{V}_0}(v_1) + \Pi_{\mathcal{V}_0}(v_2)$.

f) If $\mathcal{V}_1$ is a linear subspace then $\big\langle \Pi_{\mathcal{V}_1}(v), \Pi_{\mathcal{V}_0}(v) \big\rangle = ||\Pi_{\mathcal{V}_0}(v)||^2$.

g) It holds that $||v - \Pi_{\mathcal{V}_1}(v)|| \leq ||v - \Pi_{\mathcal{V}_0}(v)||$ and $||\Pi_{\mathcal{V}_1}(v)|| \geq ||\Pi_{\mathcal{V}_0}(v)||$ for all $v \in \mathcal{V}$. Equality in one of the inequalities implies $\Pi_{\mathcal{V}_0}(v) = \Pi_{\mathcal{V}_1}(v)$.

h) Assume that $\Pi_{\mathcal{V}_0}(v)$ is an inner point of $\mathcal{V}_0$ in the sense that for every $\widetilde{v} \in \mathcal{V}_1$ there exists a $\varepsilon > 0$, such that

$$(1 - \alpha)\Pi_{\mathcal{V}_0}(v) + \alpha\widetilde{v} \in \mathcal{V}_0 \qquad \text{for all } \alpha < \varepsilon.$$

Then it holds that $\Pi_{\mathcal{V}_0}(v) = \Pi_{\mathcal{V}_1}(v)$.

Proof. a), b) and c) are proved by checking the conditions stated in Proposition B.2.3.b.

Proof of d). As $\mathcal{V}_0$ is a linear space, it holds that $\langle v, \widetilde{v}\rangle \leq \langle \Pi_{\mathcal{V}_0}(v), \widetilde{v}\rangle$ and $\langle v, -\widetilde{v}\rangle \leq \langle \Pi_{\mathcal{V}_0}(v), -\widetilde{v}\rangle$ for $\widetilde{v} \in \mathcal{V}_0$, cf. Proposition B.2.3.b. Combining these inequalities gives the assertion.

Proof of e). Because of d) it is straightforward to check the conditions of Proposition B.2.3.b.

Proof of f). Using e) and Proposition B.2.3.b gives

$$\langle \Pi_{\mathcal{V}_1}(v), \Pi_{\mathcal{V}_0}(v)\rangle = -\langle v - \Pi_{\mathcal{V}_1}(v), \Pi_{\mathcal{V}_0}(v)\rangle + \langle v, \Pi_{\mathcal{V}_0}(v)\rangle = ||\Pi_{\mathcal{V}_0}(v)||^2.$$

Proof of g). Using Proposition B.2.3.a we get that

$$||v - \Pi_{\mathcal{V}_0}(v)|| = \inf_{\widetilde{v} \in \mathcal{V}_0} ||v - \widetilde{v}|| \geq \inf_{\widetilde{v} \in \mathcal{V}_1} ||v - \widetilde{v}|| = ||v - \Pi_{\mathcal{V}_1}(v)||.$$

Using Proposition B.2.3.b one easily shows Pythagoras's equality

$$||v - \Pi_{\mathcal{V}_i}(v)||^2 = ||v||^2 - 2\langle v, \Pi_{\mathcal{V}_i}(v)\rangle + ||\Pi_{\mathcal{V}_i}(v)||^2 = ||v||^2 - ||\Pi_{\mathcal{V}_i}(v)||^2.$$

Therefore, the inequalities $||v - \Pi_{\mathcal{V}_1}(v)|| \leq ||v - \Pi_{\mathcal{V}_0}(v)||$ and $||\Pi_{\mathcal{V}_1}(v)|| \geq ||\Pi_{\mathcal{V}_0}(v)||$ are equivalent. Assume that $||v - \Pi_{\mathcal{V}_1}(v)|| = ||v - \Pi_{\mathcal{V}_0}(v)||$. Consequently, it holds that $||v - \Pi_{\mathcal{V}_0}(v)|| = \inf_{\widetilde{v} \in \mathcal{V}_1} ||v - \widetilde{v}||$. As the projection is unique, see Proposition B.2.3.a, it follows the second part of the assertion.

Proof of h). Assume that $\Pi_{\mathcal{V}_0}(v) \neq \Pi_{\mathcal{V}_1}(v)$. Define the function

$$\begin{aligned} g(\alpha) &= ||v - (1-\alpha)\Pi_{\mathcal{V}_0}(v) - \alpha\Pi_{\mathcal{V}_1}(v)||^2 \\ &= \alpha^2 ||\Pi_{\mathcal{V}_0}(v) - \Pi_{\mathcal{V}_1}(v)||^2 + 2\alpha\langle \Pi_{\mathcal{V}_0}(v) - \Pi_{\mathcal{V}_1}(v), v - \Pi_{\mathcal{V}_0}(v)\rangle \\ &\qquad + ||v - \Pi_{\mathcal{V}_0}(v)||^2. \end{aligned}$$

As $\mathrm{d}^2 g/\mathrm{d}^2\alpha = ||\Pi_{\mathcal{V}_0}(v) - \Pi_{\mathcal{V}_1}(v)||^2 > 0$ we get that g is a strictly convex function. g) gives that $g(0) > g(1)$. Consequently, it holds that $g(0) > g(\alpha)$ for all $\alpha \in (0, 1]$. All in all, we have that $(1-\alpha_0)\Pi_{\mathcal{V}_0}(v) + \alpha_0\Pi_{\mathcal{V}_1}(v) \in \mathcal{V}_0$ and

$$||v - (1-\alpha_0)\Pi_{\mathcal{V}_0}(v) - \alpha_0\Pi_{\mathcal{V}_1}(v)|| < ||v - \Pi_{\mathcal{V}_0}(v)|| = \inf_{\widetilde{v} \in \mathcal{V}_0} ||v - \widetilde{v}||$$

for sufficiently small $\alpha_0 > 0$, which is a contradiction. □

B.2.5 Proposition. Let $\mathscr{J}$ be some real, positive semi-definite, symmetric $(m \times m)$ matrix and set

$$\mathcal{V} = \mathrm{Im}\big(\mathscr{J}\big) \qquad \text{and} \qquad \big\langle v_1, v_2 \big\rangle_{\mathscr{J}^-} = v_1^{\mathrm{T}} \mathscr{J}^- v_2, \quad v_1, v_2 \in \mathcal{V}.$$

It holds that

a) $\big(\mathcal{V}, \langle \cdot, \cdot \rangle_{\mathscr{J}^-}\big)$ is a real Hilbert space.

b) Let $\widetilde{\mathcal{V}} \subset \mathcal{V}$ a closed, convex cone, then

$$\sup_{\widetilde{v} \in \widetilde{\mathcal{V}}} \Big(\big\langle v, \widetilde{v} \big\rangle_{\mathscr{J}^-} - \frac{1}{2} ||\widetilde{v}||^2_{\mathscr{J}^-} \Big) = \frac{1}{2} \big|\big| \Pi_{\widetilde{\mathcal{V}}}(v) \big|\big|^2_{\mathscr{J}^-}, \qquad v \in \mathcal{V}.$$

c) Let $\mathscr{L}$ be some real $(m \times q)$ matrix. The sets

$$\mathcal{V}_0 = \big\{ \mathscr{J}\mathscr{L}\xi \mid \xi \in \mathbb{R}^q \big\} \quad \text{and} \quad \mathcal{V}_0^+ = \big\{ \mathscr{J}\mathscr{L}\xi \mid \xi \in \mathbb{R}^q, \xi \geq 0 \big\}$$

are closed, convex cones.

d) $\Pi_{\mathcal{V}_0}(v) = \mathscr{J}\mathscr{L}\big(\mathscr{L}^{\mathrm{T}}\mathscr{J}\mathscr{L}\big)^- \mathscr{L}^{\mathrm{T}} v$, $v \in \mathcal{V}$.

e) It holds that

$$\begin{aligned} \big|\big|\Pi_{\mathcal{V}_0^+}(v)\big|\big|^2_{\mathscr{J}^-} = \max\Big\{ & \pi^q_{\mathfrak{J}}(\mathscr{L}^{\mathrm{T}} v)^{\mathrm{T}} \big(\rho^q_{\mathfrak{J}}(\mathscr{L}^{\mathrm{T}}\mathscr{J}\mathscr{L})\big)^- \pi^q_{\mathfrak{J}}(\mathscr{L}^{\mathrm{T}} v) \times \\ & \prod_{i \in \mathfrak{J}} \mathbb{1}\Big(\pi^{|\mathfrak{J}|}_{\{i\}} \Big(\big(\rho^q_{\mathfrak{J}}(\mathscr{L}^{\mathrm{T}}\mathscr{J}\mathscr{L})\big)^- \pi^q_{\mathfrak{J}}(\mathscr{L}^{\mathrm{T}} v) \Big) \geq 0 \Big) \\ & \Big| \; \emptyset \neq \mathfrak{J} \subset \{1, \ldots, q\} \Big\}, \end{aligned} \tag{B.1}$$

where we use the notation provided in Definition 4.1.3.

Proof. $\langle \cdot, \cdot \rangle_{\mathscr{J}^-}$ is clearly symmetric and bilinear, see Proposition B.1.4. It remains to be shown that $\langle \cdot, \cdot \rangle_{\mathscr{J}^-}$ is positive definite. As $v \in \mathrm{Im}(\mathscr{J})$, we have $v = \mathscr{J} v_0$ for some v_0. It holds that $\langle v, v \rangle_{\mathscr{J}^-} = v_0 \mathscr{J} v_0 \geq 0$. $\langle v, v \rangle_{\mathscr{J}^-} = 0$ implies that $v_0 \in \ker(\mathscr{J})$, see Proposition B.3.2.b. Consequently, we get $v =$

$\mathscr{J} v_0 = 0$. It is well known that $(\mathcal{V}, ||\cdot||_{\mathscr{J}^-})$, $||v||_{\mathscr{J}^-} = \sqrt{\langle v, v\rangle_{\mathscr{J}^-}}$, $v \in \mathcal{V}$, is a real, complete, normed vector space. Therefore, the proof of a) is complete.

Proof of b). Using Proposition B.2.3 we get the following chain of equations

$$\begin{aligned}
\frac{1}{2}||\Pi_{\widetilde{\mathcal{V}}}(v)||^2_{\mathscr{J}^-} &= \frac{1}{2}||v||^2_{\mathscr{J}^-} - \frac{1}{2}||v||^2_{\mathscr{J}^-} + \big\langle v, \Pi_{\widetilde{\mathcal{V}}}(v)\big\rangle_{\mathscr{J}^-} - \frac{1}{2}||\Pi_{\widetilde{\mathcal{V}}}(v)||^2_{\mathscr{J}^-} \\
&= \frac{1}{2}||v||^2_{\mathscr{J}^-} - \frac{1}{2}||v - \Pi_{\widetilde{\mathcal{V}}}(v)||^2_{\mathscr{J}^-} \\
&= \frac{1}{2}||v||^2_{\mathscr{J}^-} - \frac{1}{2}\inf_{\widetilde{v}\in\widetilde{\mathcal{V}}}\big(||v - \widetilde{v}||^2_{\mathscr{J}^-}\big) \\
&= \frac{1}{2}||v||^2_{\mathscr{J}^-} - \frac{1}{2}\inf_{\widetilde{v}\in\widetilde{\mathcal{V}}}\big(||v||^2_{\mathscr{J}^-} - 2\big\langle v, \widetilde{v}\big\rangle_{\mathscr{J}^-} - ||\widetilde{v}||^2_{\mathscr{J}^-}\big) \\
&= \sup_{\widetilde{v}\in\widetilde{\mathcal{V}}}\bigg(\big\langle v, \widetilde{v}\big\rangle_{\mathscr{J}^-} - \frac{1}{2}||\widetilde{v}||^2_{\mathscr{J}^-}\bigg),
\end{aligned}$$

see also Behnen and Neuhaus [7, Equation (3.2.10)]. c) is straightforward. d) is a consequence of the usual calculus to compute projections on linear sub-spaces and the fact that $v \in \mathrm{Im}(\mathscr{J})$.

Proof of e). For $\mathcal{I} \subset \{1, \ldots, q\}$ we define the following closed, convex cones

$$\mathcal{V}_{\mathcal{I}} = \big\{\mathscr{J}\mathscr{L}\mathscr{T}^q_{\mathcal{I}}\xi \mid \xi \in \mathbb{R}^{|\mathcal{I}|}\big\} \quad \text{and} \quad \mathcal{V}^+_{\mathcal{I}} = \big\{\mathscr{J}\mathscr{L}\mathscr{T}^q_{\mathcal{I}}\xi \mid \xi \geq 0,\, \xi \in \mathbb{R}^{|\mathcal{J}|}\big\},$$

where $\mathcal{V}_\emptyset = \mathcal{V}^+_\emptyset = \{0\}$ and $\mathscr{T}^q_{\mathcal{I}}$ are given in Definition 4.1.3. In the following we abbreviate $z_j = \mathscr{J}\mathscr{L}\mathscr{T}^q_{\{j\}}$, $j = 1, \ldots, q$.

In the next step it is shown that if $\kappa^{(j)} \geq 0$, $j \in \mathcal{J}$, and z_j, $j \in \mathcal{J}$ are linearly dependent, where $\mathcal{J} \subset \{1, \ldots, q\}$, then there exists $\mathcal{I} \subsetneq \mathcal{J}$ and $\widetilde{\kappa}^{(i)} \geq 0$, $i \in \mathcal{I}$, such that

$$\sum_{j\in\mathcal{J}} \kappa^{(j)} z_j = \sum_{i\in\mathcal{I}} \widetilde{\kappa}^{(i)} z_i.$$

Without loss of generality we can assume that $\kappa^{(j)} > 0$. Because of the linear dependency of the vectors z_j, $j \in \mathcal{J}$, there exists sets $\mathcal{N}, \mathcal{P} \subset \mathcal{J}$ and $j_0 \in \mathcal{J}$, such that $\mathcal{N} \cap \mathcal{P} = \emptyset$, $\mathcal{N} \cup \mathcal{P} \neq \emptyset$, $\mathcal{N} \cap \{j_0\} = \mathcal{P} \cap \{j_0\} = \emptyset$, and

$$z_{j_0} = \sum_{j\in\mathcal{P}} \zeta^{(j)} z_j - \sum_{j\in\mathcal{N}} \zeta^{(j)} z_j, \qquad \zeta^{(j)} > 0, \quad j \in \mathcal{N} \cup \mathcal{P}.$$

Moreover, set $\mathcal{M} = \mathcal{J}\backslash(\mathcal{N} \cup \mathcal{P} \cup \{j_0\})$. We distinguish two cases. First, if $\min\{\kappa^{(j)}/\zeta^{(j)} \mid j \in \mathcal{N}\} \geq \kappa^{(j_0)}$ then it holds that

$$\sum_{j\in\mathcal{J}} \kappa^{(j)} z_j = \sum_{j\in\mathcal{M}} \kappa^{(j)} z_j + \sum_{j\in\mathcal{P}} \Big(\frac{\kappa^{(j)}}{\zeta^{(j)}} + \kappa^{(j_0)}\Big)\zeta^{(j)} z_j + \sum_{j\in\mathcal{N}} \Big(\frac{\kappa^{(j)}}{\zeta^{(j)}} - \kappa^{(j_0)}\Big)\zeta^{(j)} z_j,$$

i.e. the assertion. Second, if $\min\{\kappa^{(j)}/\zeta^{(j)} \mid j \in \mathcal{N}\} < \kappa^{(j_0)}$ then one can choose $j_1 \in \mathcal{N}$, such that $\kappa^{(j_1)}/\zeta^{(j_1)} = \min\{\kappa^{(j)}/\zeta^{(j)} \mid j \in \mathcal{N}\}$. Consequently, it holds that

$$\sum_{j\in\mathcal{J}} \kappa^{(j)} z_j = \sum_{j\in\mathcal{P}} \Big(\frac{\kappa^{(j)}}{\zeta^{(j)}} + \frac{\kappa^{(j_1)}}{\zeta^{(j_1)}}\Big)\zeta^{(j)} z_j + \sum_{j\in\mathcal{N}\backslash\{j_1\}} \Big(\frac{\kappa^{(j)}}{\zeta^{(j)}} - \frac{\kappa^{(j_1)}}{\zeta^{(j_1)}}\Big)\zeta^{(j)} z_j \\ + \Big(\kappa^{(j_0)} - \frac{\kappa^{(j_1)}}{\zeta^{(j_1)}}\Big) z_{j_0},$$

i.e. the assertion.

Now, we show that the left hand side of equation (B.1) is smaller or equal to the right hand side. If $\Pi_{\mathcal{V}_0^+}(v) = 0$, the assertion is trivial. Using the previous considerations we know that there exists a set $\mathfrak{I} \subset \{1,\ldots,q\}$, $\mathfrak{I} \neq \emptyset$, such that $\Pi_{\mathcal{V}_0^+}(v) = \sum_{i\in\mathfrak{I}} \kappa^{(i)} z_i$, where $\kappa^{(i)} > 0$, $i \in \mathfrak{I}$, and z_i, $i \in \mathfrak{I}$, are linearly independent. Moreover, one easily shows that $\Pi_{\mathcal{V}_0^+}(v) = \Pi_{\mathcal{V}_{\mathfrak{I}}^+}(v)$, by checking the conditions of Proposition B.2.3.b. Using Proposition B.2.4.h yields that $\Pi_{\mathcal{V}_{\mathfrak{I}}^+}(v) = \Pi_{\mathcal{V}_{\mathfrak{I}}}(v)$. By d) one gets that

$$\big|\big|\Pi_{\mathcal{V}_0^+}(v)\big|\big|^2_{\mathscr{J}^-} = \big|\big|\Pi_{\mathcal{V}_{\mathfrak{I}}}(v)\big|\big|^2_{\mathscr{J}^-} = \pi^q_{\mathfrak{I}}(\mathscr{L}^{\mathrm{T}} v)^{\mathrm{T}} \big(\rho^q_{\mathfrak{I}}(\mathscr{L}^{\mathrm{T}} \mathscr{J} \mathscr{L})\big)^- \pi^q_{\mathfrak{I}}(\mathscr{L}^{\mathrm{T}} v),$$

where Definition 4.1.3 is also applied. As z_i, $i \in \mathfrak{I}$, are linearly independent, the matrix $\rho^q_{\mathfrak{I}}(\mathscr{L}^{\mathrm{T}} \mathscr{J} \mathscr{L})$ has full rank and it hold that

$$\Big\{\pi^{|\mathfrak{I}|}_{\{i\}}\Big(\big(\rho^q_{\mathfrak{I}}(\mathscr{L}^{\mathrm{T}} \mathscr{J} \mathscr{L})\big)^- \pi^q_{\mathfrak{I}}(\mathscr{L}^{\mathrm{T}} v)\Big) \;\Big|\; i \in \mathfrak{I}\Big\} = \{\kappa^{(i)} \mid i \in \mathfrak{I}\},$$

consequently,

$$\prod_{i\in\mathfrak{I}} \mathbb{1}\Big(\pi^{|\mathfrak{I}|}_{\{i\}}\Big(\big(\rho^q_{\mathfrak{I}}(\mathscr{L}^{\mathrm{T}} \mathscr{J} \mathscr{L})\big)^- \pi^q_{\mathfrak{I}}(\mathscr{L}^{\mathrm{T}} v)\Big) \geq 0\Big) = 1.$$

Therefore, the left hand side of equation (B.1) is smaller or equal to the right hand side.

At last, we show that the right hand side of equation (B.1) is smaller or equal to the left hand side. In the case that the right hand side is 0 the assertion is trivial. Therefore, we can assume that the right hand side of equation (B.1) is greater than 0. For any subset $\mathcal{I} \subset \{1, \ldots, q\}$, such that

$$\pi_{\mathcal{I}}^{q}(\mathscr{L}^{\mathrm{T}} v)^{\mathrm{T}} \left(\rho_{\mathcal{I}}^{q}(\mathscr{L}^{\mathrm{T}} \mathscr{J} \mathscr{L})\right)^{-} \pi_{\mathcal{I}}^{q}(\mathscr{L}^{\mathrm{T}} v) > 0$$

and

$$\prod_{i \in \mathcal{I}} \mathbb{1}\left(\pi_{\{i\}}^{|\mathcal{I}|}\left(\left(\rho_{\mathcal{I}}^{q}(\mathscr{L}^{\mathrm{T}} \mathscr{J} \mathscr{L})\right)^{-} \pi_{\mathcal{I}}^{q}(\mathscr{L}^{\mathrm{T}} v)\right) \geq 0\right) = 1,$$

it holds that

$$\pi_{\mathcal{I}}^{q}(\mathscr{L}^{\mathrm{T}} v)^{\mathrm{T}} \left(\rho_{\mathcal{I}}^{q}(\mathscr{L}^{\mathrm{T}} \mathscr{J} \mathscr{L})\right)^{-} \pi_{\mathcal{I}}^{q}(\mathscr{L}^{\mathrm{T}} v) = \left|\left|\Pi_{\mathcal{V}_{\mathcal{I}}}(v)\right|\right|_{\mathscr{J}^{-}}^{2}$$

and $\Pi_{\mathcal{V}_{\mathcal{I}}}(v) \in \mathcal{V}_{\mathcal{I}}^{+}$. Consequently, one receives that $\Pi_{\mathcal{V}_{\mathcal{I}}}(v) = \Pi_{\mathcal{V}_{\mathcal{I}}^{+}}(v)$, by checking the conditions of Proposition B.2.3.b. Using Proposition B.2.4.g gives that

$$\left|\left|\Pi_{\mathcal{V}_{\mathcal{I}}}(v)\right|\right|_{\mathscr{J}^{-}}^{2} = \left|\left|\Pi_{\mathcal{V}_{\mathcal{I}}^{+}}(v)\right|\right|_{\mathscr{J}^{-}}^{2} \leq \left|\left|\Pi_{\mathcal{V}_{0}^{+}}(v)\right|\right|_{\mathscr{J}^{-}}^{2},$$

which completes the proof. □

B.3 Results on Covariance Matrices

B.3.1 Definition (Covariance Matrix). Let $\mathscr{J}$ be a real, symmetric, positive semi-definite $(m \times m)$ matrix. $\mathscr{J}$ is called covariance matrix.

B.3.2 Proposition. For any $(m \times m)$ covariance matrix $\mathscr{J}$, it holds that

$$\mathbb{R}^{m} = \ker(\mathscr{J}) \oplus \operatorname{Im}(\mathscr{J})$$

with respect to the Euclidean inner product. In particular, we have that

a) $s \in \operatorname{Im}(\mathscr{J}) \iff s^{\mathrm{T}}\kappa = 0$ for all $\kappa \in \ker(\mathscr{J})$.

b) $s \in \ker(\mathscr{J}) \Longleftrightarrow s^{\mathrm{T}} \mathscr{J} s = 0$.

Proof. Set $l = \operatorname{rank}(\mathscr{J})$. The matrix $\mathscr{J}$ is diagonalisable, therefore there exists an orthonormal basis of eigenvectors $v_1, \ldots v_m$, such that $\mathscr{F}^{\mathrm{T}} \mathscr{J} \mathscr{F} = \mathscr{D}$, where $\mathscr{D} = \operatorname{diag}\big(\lambda_1, \ldots, \lambda_l, 0, \ldots, 0\big)$, $\lambda_i > 0$, $i = 1, \ldots, l$, and $\mathscr{F} = \big(v_1, \ldots, v_m\big)$. We have that $\operatorname{Im}(\mathscr{J}) = \operatorname{span}\big(v_i \mid i = 1, \ldots, l\big)$ and $\ker(\mathscr{J}) = \operatorname{span}\big(v_i \mid i = l+1, \ldots, m\big)$. Hence, the assertions are trivial. □

B.3.3 Proposition (Decompositions of Covariance Matrices). Let $\mathscr{J}$ be a $(m \times m)$ covariance matrix and set $l = \operatorname{rank}(\mathscr{J})$. The following assertions hold true.

a) There exists a (uniquely determined) $(m \times m)$ covariance matrix $\mathscr{J}^{1/2}$ satisfying the following properties.
   (i) $\mathscr{J}^{1/2} \mathscr{J}^{1/2} = \mathscr{J}$.
   (ii) $\operatorname{rank}(\mathscr{J}^{1/2}) = l$.
   (iii) $\operatorname{Im}(\mathscr{J}) = \operatorname{Im}(\mathscr{J}^{1/2})$.
   (iv) The linear mappings

$$\mathscr{J} : \operatorname{Im}(\mathscr{J}) \longrightarrow \operatorname{Im}(\mathscr{J}) \quad \text{and} \quad \mathscr{J}^{1/2} : \operatorname{Im}(\mathscr{J}) \longrightarrow \operatorname{Im}(\mathscr{J})$$

   are bijective.

b) There exists $(m \times l)$ matrices $\mathscr{B}$ and $\mathscr{C}$, such that
   (i) $\mathscr{J} = \mathscr{B}\mathscr{B}^{\mathrm{T}}$ and $\mathscr{J}^{-} = \mathscr{C}\mathscr{C}^{\mathrm{T}}$,
   (ii) $\operatorname{rank}(\mathscr{B}) = \operatorname{rank}(\mathscr{C}) = l$,
   (iii) $\mathscr{B}^{\mathrm{T}}\mathscr{C} = \mathscr{C}^{\mathrm{T}}\mathscr{B} = \mathscr{E}_l$, where $\mathscr{E}_l$ denotes the $(l \times l)$ unity matrix,
   (iv) $\mathscr{B}^{-} = \mathscr{C}^{\mathrm{T}}$ and $\mathscr{C}^{-} = \mathscr{B}^{\mathrm{T}}$.

Proof. We use the notation introduced in the proof of Proposition B.3.2. Set $\mathscr{J}^{1/2} = \mathscr{F}\widetilde{\mathscr{D}}\mathscr{F}^{\mathrm{T}}$, where $\widetilde{\mathscr{D}} = \operatorname{diag}\big(\sqrt{\lambda_1} \ldots, \sqrt{\lambda_l}, 0 \ldots, 0\big)$. Obviously, it holds that $\mathscr{J}^{1/2}\mathscr{J}^{1/2} = \mathscr{F}\mathscr{D}\mathscr{F}^{\mathrm{T}} = \mathscr{J}$ and that $\operatorname{rank}(\mathscr{J}^{1/2}) = \operatorname{rank}(\widetilde{\mathscr{D}}) = \operatorname{rank}(\mathscr{D}) = l$. Clearly, $\mathscr{J}^{1/2}$ is a covariance matrix. We have that $\operatorname{Im}(\mathscr{J}^{1/2}) =$

$\operatorname{span}\big(v_i \mid i = 1, \ldots, l\big) = \operatorname{Im}(\mathscr{J})$. As $v_1, \ldots, v_l$ are eigenvectors to positive eigenvalues the remaining assertions of a) are straightforward.

For the proof of b) we set

$$\mathscr{B} = \mathscr{F} \begin{pmatrix} \operatorname{diag}\big(\sqrt{\lambda_1} \ldots, \sqrt{\lambda_l}\big) \\ 0 \end{pmatrix} \quad \text{and} \quad \mathscr{C} = \mathscr{F} \begin{pmatrix} \operatorname{diag}\big(1/\sqrt{\lambda_1} \ldots, 1/\sqrt{\lambda_l}\big) \\ 0 \end{pmatrix}.$$

It holds that $\mathscr{B}\mathscr{B}^{\mathrm{T}} = \mathscr{F}\mathscr{D}\mathscr{F}^{\mathrm{T}} = \mathscr{J}$ and $\mathscr{C}\mathscr{C}^{\mathrm{T}} = \mathscr{F}\mathscr{D}^{-}\mathscr{F}^{\mathrm{T}} = \mathscr{J}^{-}$, where we use Proposition B.1.6. Obviously, it also holds that $\operatorname{rank}(\mathscr{B}) = \operatorname{rank}(\mathscr{C}) = l$. The proof of the last but one assertion is straightforward. The last assertion of b) is shown by checking the conditions of Definition B.1.1 and using Proposition B.1.2. □

B.3.4 Proposition. Let $\mathscr{J}$ be some $(m \times m)$ covariance matrix and $s \in \operatorname{Im}(\mathscr{J})$. Furthermore, assume that $\mathscr{J}$ is partitioned as follows

$$\mathscr{J} = \begin{pmatrix} \mathscr{J}_{1,1} & \mathscr{J}_{1,2} \\ \mathscr{J}_{2,1} & \mathscr{J}_{2,2} \end{pmatrix},$$

where $\mathscr{J}_{1,1}$ is a $(r \times r)$ matrix.

a) It holds the inclusion $\ker(\mathscr{J}_{2,2}) \subset \ker(\mathscr{J}_{1,2})$.

b) There exists a matrix $\mathscr{C}$, such that $\mathscr{J}_{1,2} = \mathscr{C}^{\mathrm{T}}\mathscr{J}_{2,2}$.

c) Assume that $\mathscr{V}$ is some real $(m \times q)$ matrix. It holds that

$$\mathscr{V}^{\mathrm{T}} s \in \operatorname{Im}(\mathscr{V}^{\mathrm{T}}\mathscr{J}\mathscr{V}).$$

Proof. Let $\mathscr{B}$ be the matrix defined in Proposition B.3.3.b and assume that

$$\mathscr{B} = \begin{pmatrix} \mathscr{B}_1 \\ \mathscr{B}_2 \end{pmatrix},$$

where $\mathscr{B}_1$ is a $(r \times l)$ matrix. It holds that

$$\mathscr{J} = \begin{pmatrix} \mathscr{J}_{1,1} & \mathscr{J}_{1,2} \\ \mathscr{J}_{2,1} & \mathscr{J}_{2,2} \end{pmatrix} = \begin{pmatrix} \mathscr{B}_1\mathscr{B}_1^{\mathrm{T}} & \mathscr{B}_1\mathscr{B}_2^{\mathrm{T}} \\ \mathscr{B}_2\mathscr{B}_1^{\mathrm{T}} & \mathscr{B}_2\mathscr{B}_2^{\mathrm{T}} \end{pmatrix}.$$

Proposition B.3.2.b gives that $\kappa^{\mathrm{T}}\mathscr{B}_2\mathscr{B}_2^{\mathrm{T}}\kappa = 0$ for all $\kappa \in \ker(\mathscr{J}_{2,2})$. Thus, one gets that $\mathscr{B}_2^{\mathrm{T}}\kappa = 0$ and $\mathscr{B}_1\mathscr{B}_2^{\mathrm{T}}\kappa = 0$, *i.e.* the assertion of a) holds.

Proof of b). Proposition B.3.2.a and a) give that the columns of $\mathscr{J}_{2,1}$ are elements of $\mathrm{Im}(\mathscr{J}_{2,2})$. Thus, $\mathscr{J}_{2,1} = \mathscr{J}_{2.2}\mathscr{C}$ for some matrix $\mathscr{C}$. As $\mathscr{J}_{2,1}^{\mathrm{T}} = \mathscr{J}_{1,2}$ and $\mathscr{J}_{2,2}$ is symmetric, the result is the assertion.

Proof of c). As $s \in \mathrm{Im}(\mathscr{J})$ there exists s_0, such that $s = \mathscr{J} s_0$. Assume that $\kappa \in \ker\big(\mathscr{V}^{\mathrm{T}}\mathscr{J}\mathscr{V}\big)$. Proposition B.3.2.b gives that $\mathscr{V}\kappa \in \ker(\mathscr{J})$. Consequently, $\kappa^{\mathrm{T}}\mathscr{V}^{\mathrm{T}}s = s_0^{\mathrm{T}}\mathscr{J}\mathscr{V}\kappa = 0$. Proposition B.3.2.a yields the assertion. □

B.4 Results on Stochastic Convergence

This section provides some results on stochastic convergence used in the previous chapters.

B.4.1 Definition (Stochastic Convergence). a) Assume that $(\Omega, \mathcal{A}, \mathbb{P})$ is some probability space and that $V^{(u,v)} : \Omega, \longrightarrow \mathbb{R}$, $u = 1, \dots, q$, $v = 1, \dots, r$, are measurable mappings. The mapping

$$V : \Omega \longrightarrow \mathbb{R}^{q\times r}, \qquad V = \begin{pmatrix} V^{(1,1)} & \dots & V^{(1,r)} \\ \vdots & \ddots & \vdots \\ V^{(q,1)} & \dots & V^{(q,r)} \end{pmatrix},$$

is called a real $(q \times r)$ random matrix.

b) Let V_n, $n \in \mathbb{N}$, be a sequence of real $(q \times r)$ random matrices. We say V_n converges in probability to some real $(q \times r)$ matrix $\mathscr{V}$,

$$V_n - \mathscr{V} \longrightarrow_{\mathbb{P}_n} 0, \qquad \text{as } n \to \infty,$$

if and only if, for all $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_n\big(|V_n^{(u,v)} - \mathscr{V}^{(u,v)}| \geq \varepsilon\big) = 0 \qquad u = 1, \dots, q,\ v = 1, \dots, r.$$

B.4.2 Remark. a) The vector spaces $\mathbb{R}^{q\times r}$ and $\mathbb{R}^{q\cdot r}$ are isomorph.

b) The definition of the stochastic convergence is based on the vector space isomorphism between $\mathbb{R}^{q\times r}$ and $\mathbb{R}^{q\cdot r}$ and the fact that on $\mathbb{R}^{q\cdot r}$ all norms are equivalent. Choosing the sup-norm, cf. Definition B.4.3, leads to the above definition of the stochastic convergence.

c) For sequences of real (1×1) random matrices Definition B.4.1 is the usual definition of stochastic convergence for sequences of real random variables.

B.4.3 Definition (Sup-Norm, Row-Sum-Norm). Assume that $x \in \mathbb{R}^q$ and that $\mathscr{A}$ is some real $(r \times q)$ matrix.

a) The mapping $||\cdot||_\infty : \mathbb{R}^q \longrightarrow \mathbb{R}$, $||x||_\infty = \max_{1\le j\le q}|x^{(j)}|$, is called sup-norm.

b) The mapping $||\cdot||_{\text{r.s.}} : \mathbb{R}^{r\times q} \longrightarrow \mathbb{R}$, $||\mathscr{A}||_{\text{r.s.}} = \max_{1\le u\le q}\sum_{v=1}^{r}|\mathscr{A}^{(u,v)}|$ is called row-sum-norm.

B.4.4 Proposition (Properties of Row-Sum-Norm). The following assertions hold true.

a) $||\cdot||_{\text{r.s.}}$ is a norm on the space $\mathbb{R}^{q\times r}$.

b) $||\mathscr{A}x||_\infty \le ||\mathscr{A}||_{\text{r.s.}} \cdot ||x||_\infty$, $\mathscr{A} \in \mathbb{R}^{q\times r}$, $x \in \mathbb{R}^r$.

c) $||\mathscr{A}\mathscr{B}||_{\text{r.s.}} \le ||\mathscr{A}||_{\text{r.s.}} \cdot ||\mathscr{B}||_{\text{r.s.}}$, $\mathscr{A} \in \mathbb{R}^{q\times r}$, $\mathscr{B} \in \mathbb{R}^{r\times s}$.

d) $||x||_\infty = ||x||_{\text{r.s.}}$, $x \in \mathbb{R}^q$.

Proof. Cf. Königsberger [44, pp. 26]. □

B.4.5 Proposition. Let $\mathscr{V}$ be some real $(q \times r)$ matrix. The following statements are equivalent

(i) $\widehat{V}_n - \mathscr{V} \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$.

(ii) $||\widehat{V}_n - \mathscr{V}||_{\text{r.s.}} \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$.

Proof. Assume that (i) holds. For all $\varepsilon > 0$ we have that

$$\begin{aligned}\mathbb{P}_n\big\{||\widehat{V}_n - \mathscr{V}||_{\text{r.s.}} \ge \varepsilon\big\} &\le \sum_{u=1}^{q}\mathbb{P}_n\left(\sum_{v=1}^{r}|\widehat{V}_n^{(u,v)} - \mathscr{V}^{(u,v)}| \ge \varepsilon\right)\\ &\le \sum_{u=1}^{q}\sum_{v=1}^{r}\mathbb{P}_n\Big(|\widehat{V}_n^{(u,v)} - \mathscr{V}^{(u,v)}| \ge \frac{\varepsilon}{r}\Big) \to 0,\end{aligned}$$

as $n \to \infty$. Let us assume that (ii) holds. As we have

$$\mathbb{P}_n\Big(|\widehat{V}_n^{(u,v)} - \mathscr{V}^{(u,v)}| \geq \varepsilon\Big) \leq \mathbb{P}_n\big(||\widehat{V}_n - \mathscr{V}||_{\text{r.s.}} \geq \varepsilon\big) \to 0,$$

as $n \to \infty$, for all $\varepsilon > 0$, the proof is complete. □

B.4.6 Proposition. a) Assume that

$$V_{n,i} \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} V_i \quad \text{and} \quad V_{n,i} - W_{n,i} \longrightarrow_{\mathbb{P}_n} 0, \quad \text{as } n \to \infty, \qquad i = 1, 2,$$

where $V_{n,1}$ is a real $(q \times r)$ random matrix and $V_{n,2}$ is a real $(r \times s)$ random matrix. It holds that

$$V_{n,1}V_{n,2} - W_{n,1}W_{n,2} \longrightarrow_{\mathbb{P}_n} 0, \qquad \text{as } n \to \infty.$$

b) $V_n - \mathscr{V} \longrightarrow_{\mathbb{P}_n} 0$, as $n \to \infty$, implies $V_n \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} \mathscr{V}$ as $n \to \infty$, where V_n, $n \in \mathbb{N}$, is a sequence of $(q \times r)$ random matrices and $\mathscr{V}$ is a $(q \times r)$ matrix.

Proof. We show the first assertion. Using Slutsky's Lemma, cf. Witting and Müller-Funk [72, Satz 5.45], one gets that $W_{n,i} \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} V_i$. The Continuous Mapping Theorem, cf. Witting and Müller-Funk [72, Satz 5.43], yields $||V_{n,i}||_{\text{r.s.}} \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} ||V_i||_{\text{r.s.}}$ and $||W_{n,i}||_{\text{r.s.}} \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} ||V_i||_{\text{r.s.}}$. The following estimate and a special case of Slutsky's Lemma, cf. Witting and Müller-Funk [72, Korollar 5.84], give

$$\begin{aligned} 0 &\leq ||V_{n,1}V_{n,2} - W_{n,1}W_{n,2}||_{\text{r.s.}} \\ &= ||V_{n,1}V_{n,2} - V_{n,1}W_{n,2} + V_{n,1}W_{n,2} - W_{n,1}W_{n,2}||_{\text{r.s.}} \\ &\leq ||V_{n,1}||_{\text{r.s.}}\, ||V_{n,2} - W_{n,2}||_{\text{r.s.}} + ||V_{n,1} - W_{n,1}||_{\text{r.s.}}\, ||W_{n,2}||_{\text{r.s.}} \longrightarrow_{\mathbb{P}_n} 0, \end{aligned}$$

as $n \to \infty$. Proposition B.4.5 yields the assertion.

Using Remark B.4.2 and Witting and Müller-Funk [72, Hilfssatz 5.82] one gets the second assertion. □

B.4.7 Proposition. Assume that $X_n \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} X$, where $\mathfrak{L}(X)$ is some distribution on $\mathbb{R}$, such that $\mathbb{P}\big(X = 0\big) = 0$. If $X_n - \widehat{X}_n \longrightarrow_{\mathbb{P}_n} 0$ then we have that

$$\mathbb{1}\big(X_n \geq 0\big) - \mathbb{1}\big(\widehat{X}_n \geq 0\big) \longrightarrow_{\mathbb{P}_n} 0, \qquad \text{as } n \to \infty.$$

Proof. By Slutsky's Lemma, cf. Witting and Müller-Funk [72, Satz 5.45], we also know $\widehat{X}_n \xrightarrow{\mathfrak{D}}_{\mathbb{P}_n} X$, as $n \to \infty$. Because of the inclusions

$$\begin{aligned}\big\{X_n \geq 0\big\} \cap \big\{\widehat{X}_n < 0\big\} &= \Big(\big\{X_n \geq 0\big\} \cap \big\{\widehat{X}_n < 0\big\} \cap \big\{|X_n - \widehat{X}_n| \geq \delta\big\}\Big) \\ &\qquad \cup \Big(\big\{X_n \geq 0\big\} \cap \big\{\widehat{X}_n < 0\big\} \cap \big\{|X_n - \widehat{X}_n| < \delta\big\}\Big) \\ &\subset \big\{|X_n - \widehat{X}_n| \geq \delta\big\} \cup \big\{0 \leq X_n \leq \delta\big\}\end{aligned}$$

and

$$\big\{X_n < 0\big\} \cap \big\{\widehat{X}_n \geq 0\big\} \subset \big\{|X_n - \widehat{X}_n| \geq \delta\big\} \cup \big\{0 \leq \widehat{X}_n \leq \delta\big\},$$

for all $\delta, \varepsilon > 0$, it holds that

$$\begin{aligned}&\mathbb{P}_n\Big\{\big|\mathbb{1}\big(X_n \geq 0\big) - \mathbb{1}\big(\widehat{X}_n \geq 0\big)\big| \geq \varepsilon\Big\} \\ &\qquad \leq \mathbb{P}_n\Big(\big(\{X_n \geq 0\} \cap \{\widehat{X}_n < 0\}\big) \cup \big(\{X_n < 0\} \cap \{\widehat{X}_n \geq 0\}\big)\Big) \\ &\qquad \leq 2\mathbb{P}_n\big(|X_n - \widehat{X}_n| \geq \delta\big) + \mathbb{P}_n\big(0 \leq X_n \leq \delta\big) + \mathbb{P}_n\big(0 \leq \widehat{X}_n \leq \delta\big).\end{aligned}$$

As we can choose a sequence of $\delta_k \downarrow 0$, such that $\mathbb{P}\big(X = \delta_k\big) = 0$, the Portmanteau Theorem, cf. Billingsley [9, Theorem 2.1], gives that

$$\limsup_{n\to\infty} \mathbb{P}_n\Big(\big|\mathbb{1}(X_n \geq 0) - \mathbb{1}(\widehat{X}_n \geq 0)\big| \geq \varepsilon\Big) \leq 2\,\mathbb{P}\big(0 \leq X \leq \delta_k\big) \to 0,$$

as $k \to \infty$, since $\mathbb{P}\big(0 \leq X \leq \delta_k\big) \to \mathbb{P}\big(X = 0\big)$, as $k \to \infty$. □

B.4.8 Proposition (Sub-Sub-Sequence Principle for Convergence in Probability). Let X_n, $n \in \mathbb{N}$ be a sequence of real-valued random variables that are defined on the probabilty space $(\Omega, \mathcal{F}, \mathbb{P})$, the following conditions are equivalent

i) $X_n \longrightarrow_{\mathbb{P}} 0$, as $n \to \infty$.

ii) In every sub-sequence of the natural numbers m_n, $n \in \mathbb{N}$, there exists a sub-sub-sequence k_n, $n \in \mathbb{N}$, such that $X_{k_n} \to 0$, as $n \to \infty$, $\mathbb{P}$-almost surely.

Proof. Cf. Bauer [6, Korollar 20.8]. □

B.5 Results on Measure Theory

B.5.1 Proposition. Let $(\Omega, \mathcal{F})$ be some measurable space and $X_t : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$, $t \in \mathbb{R}_+$, measurable mappings. If for all $t \in \mathbb{R}_+$ and $\omega \in \Omega$ there exists $\varepsilon > 0$, such that

$$X_t(\omega) = X_{t+s}(\omega) \qquad \text{for all } s \in [0, \varepsilon)$$

then the filtration $\{\mathcal{F}_t \mid t \in \mathbb{R}_+\}$, $\mathcal{F}_t = \sigma\big(X_s \mid s \leq t\big)$, is right continuous. For any σ-algebra $\mathcal{G}$, $\mathcal{G} \subset \mathcal{F}$, it holds that the filtration $\{\mathcal{G} \vee \mathcal{F}_t \mid t \in \mathbb{R}_+\}$ is right continuous.

Proof. The first assertion can be found in Fleming and Harrington [19, Theorem A.2.6]. Considering the mappings

$$\widetilde{X}_t : \big(\Omega, \mathcal{F}\big) \longrightarrow \big(\Omega \times \Omega', \mathcal{G} \otimes \mathcal{F}'\big), \qquad \widetilde{X}_t(\omega) = \big(\omega, X_t(\omega)\big), \qquad t \in \mathbb{R}_+,$$

one sees that $\mathcal{G} \vee \mathcal{F}_t = \sigma\big(\widetilde{X}_s \mid s \leq t\big)$. Thus, the first assertion implies the second. $\square$

B.5.2 Proposition. Set $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ and let $\bar{\mathbb{B}}$ denote the Borel σ-algebra on $\bar{\mathbb{R}}$. Assume that $\big(\Omega, \mathcal{F}, \mathbb{F} = \{F_t \mid t \in \mathbb{R}_+\}, Q\big)$ is some filtered space and that $Z_t : \Omega \longrightarrow \bar{\mathbb{R}}^p$, $t \in \mathbb{R}_+$, are measurable mappings, such that the process $\{Z_t \mid t \in \mathbb{R}_+\}$ is progressively measurable, *i.e.* for all $t \in \mathbb{R}_+$ the mapping $(\omega, s) \mapsto Z_s(\omega)$ on $\Omega \times [0, t]$ is $\mathcal{F}_t \otimes \mathbb{B}_+[0, t]$–$\bar{\mathbb{B}}^p$ measurable, where $\mathbb{B}_+[0, t] = \{B \cap [0, t] \mid B \in \mathbb{B}_+\}$. The following assertions hold true.

a) $\{Z_t \mid t \in \mathbb{R}_+\}$ is adapted to $\mathbb{F}$.

b) The mapping $s \mapsto Z_s(\omega)$ is $\mathbb{B}_+$–$\bar{\mathbb{B}}^p$ measurable for every $\omega \in \Omega$.

c) Assume that $f : \big(\bar{\mathbb{R}}^p, \bar{\mathbb{B}}^p\big) \longrightarrow \big(\bar{\mathbb{R}}, \bar{\mathbb{B}}\big)$ is non-negative and that μ is some σ-finite measure on $\mathbb{B}_+$. The processes

$$\big\{f(Z_s) \mid s \in \mathbb{R}_+\big\} \quad \text{and} \quad \left\{\int_{[0,t]} f(Z_s)\, \mathrm{d}\mu(s) \,\middle|\, t \in \mathbb{R}_+\right\}$$

are progressively measurable.

Proof. The first assertion is an easy consequence of Bauer [6, Lemma 23.5]. The same Lemma also yields that $\{s \mid Z_s(\omega) \in B,\ s \le t\} \in \mathbb{B}_+[0,t]$ for every $\omega \in \Omega$ and $B \in \bar{\mathbb{B}}^p$, $t \in \mathbb{R}_+$. Consequently, we get that

$$\begin{aligned} \{s \mid Z_s(\omega) \in B\} &= \bigcup_{n \ge 1} \{s \mid Z_s(\omega) \in B\} \cap [0,n] \\ &= \bigcup_{n \ge 1} \underbrace{\{s \mid Z_s(\omega) \in B,\ s \le n\}}_{\in \mathbb{B}_+}, \end{aligned} \tag{B.2}$$

which proves b).

Consider the mapping $(\omega, s) \mapsto f\big(Z_s(\omega)\big)$ on $\Omega \times [0,t]$. Since $f^{-1}(B) \in \bar{\mathbb{B}}^p$ for all $B \in \bar{\mathbb{B}}$, it holds that

$$\Big\{(\omega, s) \mid f\big(Z_s(\omega)\big) \in B\Big\} = \big\{(\omega, s) \mid Z_s(\omega) \in f^{-1}(B)\big\} \in \mathcal{F}_t \otimes \mathbb{B}_+[0,t],$$

where we use that the process $\{Z_s \mid s \in \mathbb{R}_+\}$ is progressively measurable.

Consider the space $\Omega \times [0,t] \times [0,t]$ equipped with the σ-algebra $\mathcal{F}_t \otimes \mathbb{B}_+[0,t] \otimes \mathbb{B}_+[0,t]$. All processes and functions are now defined on this product space. The mappings $(\omega, s, u) \mapsto f\big(Z_s(\omega)\big)$ and $(\omega, s, u) \mapsto \mathbb{1}(u \le s)$, $(\omega, s, u) \in \Omega \times [0,t] \times [0,t]$, are obviously $\mathcal{F}_t \otimes \mathbb{B}_+[0,t] \otimes \mathbb{B}_+[0,t]$–$\bar{\mathbb{B}}$ measurable. Fubini's Theorem, cf. Bauer [6, Satz 23.6], gives that

$$(\omega, s) \mapsto \int_{[0,s]} f\big(Z_s(\omega)\big)\, \mathrm{d}\mu(u), \quad (\omega, s) \in \Omega \times [0,t],$$

is $\mathcal{F}_t \otimes \mathbb{B}_+[0,t]$–$\bar{\mathbb{B}}$ measurable. □

B.5.3 Proposition. Let $\big(\Omega, \mathcal{F}, \mathbb{F} = \{F_t \mid t \in \mathbb{R}_+\}, Q\big)$ be some filtered space and assume that $Z_t : \Omega \longrightarrow \bar{\mathbb{R}}^p$, $t \in \mathbb{R}_+$, are mappings, such that the process $Z = \{Z_t \mid t \in \mathbb{R}_+\}$ is predictable. Moreover, $f : \mathbb{R}^p \longrightarrow \mathbb{R}^q$ is some $\mathbb{B}^p$–$\mathbb{B}^q$ measurable function.

a) The process $\big\{f(Z_t) \mid t \in \mathbb{R}_+\big\}$ is predictable.

b) Assume that Z is additionally locally bounded. If f is continuous or bounded then the process $\big\{f(Z_t) \mid t \in \mathbb{R}_+\big\}$ is locally bounded.

Proof. Let $\mathcal{P}$ denote the predictable σ-algebra. Clearly, for all $B \in \mathbb{B}^q$ it holds that

$$\big\{(\omega, s) \mid f(Z_s) \in B\big\} = \big\{(\omega, s) \mid Z_s \in f^{-1}(B)\big\} \in \mathcal{P}.$$

Proof of b). In the case that f is bounded the assertion is trivial. In the case that f is continuous the assertion is implied by the fact that f maps compact sets on compact sets. $\square$

B.5.4 Proposition. Assume that $(\Omega_i, \mathcal{F}_i, \mathbb{F}_i, Q_i)$, where $\mathbb{F}_i = \{\mathcal{F}_{i,t} \mid t \in \mathbb{R}_+\}$, $i = 1, \dots, n$, are filtered probability spaces and that $Z_{i,t} : \Omega_i \longrightarrow \bar{\mathbb{R}}_+$, $t \in \mathbb{R}_+$, are measurable mappings.

Define $\Omega = \bigtimes \Omega_i$, $\mathcal{F} = \bigotimes_{i=1}^n \mathcal{F}_i$, $Q = \bigotimes_{i=1}^n Q_i$ as well as

$$\mathbb{F}^0 = \Big\{\mathcal{F}_t^0 = \bigotimes_{i=1}^n \mathcal{F}_{i,t} \;\Big|\; t \in \mathbb{R}_+\Big\}, \quad \mathbb{F} = \Big\{\mathcal{F}_t = \bigcap_{s>t}^n \mathcal{F}_s^0 \;\Big|\; t \in \mathbb{R}_+\Big\}$$

and $\varpi_i : \Omega \longrightarrow \Omega_i$, $\varpi_i(\omega_1, \dots, \omega_n) = \omega_i$, $i = 1, \dots, n$.

a) $\bigvee_{t \geq 0} \mathcal{F}_{i,t} = \mathcal{F}_i$, $i = 1, \dots, n$, implies that $\bigvee_{t \geq 0} \mathcal{F}_t^0 = \bigvee_{t \geq 0} \mathcal{F}_t = \mathcal{F}$.

b) If $\{Z_{i,t} \mid t \in \mathbb{R}_+\}$ is progressively measurable then $\{Z_{i,t} \circ \varpi_i \mid t \in \mathbb{R}_+\}$ is progressively measurable with respect to $\mathbb{F}$ and $\mathbb{F}^0$.

c) If $\{Z_{i,t} \mid t \in \mathbb{R}_+\}$ is predictable then $\{Z_{i,t} \circ \varpi_i \mid t \in \mathbb{R}_+\}$ is predictable with respect to $\mathbb{F}$ and $\mathbb{F}^0$.

d) If τ_i is a $\mathbb{F}_i$ stopping time then $\tau_\circ \varpi_i$ is a $\mathbb{F}$ and $\mathbb{F}^0$ stopping time.

Proof. As $\{\mathcal{F}_t^0 \mid t \geq 0\}$ is increasing, one gets that $\bigcup_{t \geq 0} \mathcal{F}_t^0 = \bigcup_{t \geq 0} \mathcal{F}_t$. Obviously, it holds that $\bigvee_{t \geq 0} \mathcal{F}_t^0 \subset \mathcal{F}$, since

$$\mathcal{F}_t^0 = \sigma\Big(\bigtimes_{i=1}^n F_i \;\Big|\; F_i \in \mathcal{F}_{i,t}\Big) \subset \sigma\Big(\bigtimes_{i=1}^n F_i \;\Big|\; F_i \in \mathcal{F}_i\Big) = \mathcal{F}.$$

On the other hand, it holds that

$$
\begin{aligned}
\mathcal{F} &= \sigma\Big(\bigcup_{i=1}^{n}\{\varpi_i^{-1}(F) \mid F \in \mathcal{F}_i\}\Big) \\
&= \sigma\Big(\bigcup_{i=1}^{n}\Big\{\varpi_i^{-1}(F) \mid F \in \bigvee_{t\geq 0}\mathcal{F}_{i,t}\Big\}\Big) \\
&= \sigma\Big(\bigcup_{i=1}^{n}\sigma\Big(\varpi_i^{-1}(F) \mid F \in \bigcup_{t\geq 0}\mathcal{F}_{i,t}\Big)\Big) \\
&= \sigma\Big(\bigcup_{i=1}^{n}\Big\{\varpi_i^{-1}(F) \mid F \in \bigcup_{t\geq 0}\mathcal{F}_{i,t}\Big\}\Big) \\
&= \sigma\Big(\bigcup_{t\geq 0}\bigcup_{i=1}^{n}\{\varpi_i^{-1}(F) \mid F \in \mathcal{F}_{i,t}\}\Big) \\
&= \sigma\Big(\bigcup_{t\geq 0}\Big\{\bigtimes_{i=1}^{n} F_i \;\Big|\; F_i \in \mathcal{F}_{i,t}\Big\}\Big) \subset \sigma\Big(\bigcup_{t\geq 0}\mathcal{F}_t^0\Big) = \sigma\Big(\bigcup_{t\geq 0}\mathcal{F}_t\Big).
\end{aligned}
$$

Proof of b). It holds that

$$
\big\{(\omega, s) \mid Z_{i,s}\circ\varpi_i(\omega) \in B\big\} = \varpi_i^{-1}\Big(\big\{(\omega, s) \mid Z_{i,s}(\omega) \in B\big\}\Big) \in \mathcal{F}_t^0 \otimes \mathbb{B}_+[0,t]
$$

and $\mathcal{F}_t^0 \subset \mathcal{F}_t$ for all $B \in \bar{\mathbb{B}}^p$.

Proof of c). Let $\mathcal{Q}$ and $\mathcal{Q}_i$ denote the predictable σ-algebras with respect to $\mathbb{F}$ and $\mathbb{F}_i$. Define $\widetilde{\varpi}_i : \Omega\times\mathbb{R}_+ \longrightarrow \Omega_i\times\mathbb{R}_+$, $\widetilde{\varpi}_i(\omega, s) = (\omega_i, s)$. As $\mathcal{Q}_i$ is generated by the predictable rectangles

$$
F_0 \times \{0\}, \quad F_s \times (s,t], \qquad F_u \in \mathcal{F}_{i,u},\ s < t,
$$

and as $\widetilde{\varpi}_i^{-1}\big(F_0\times\{0\}\big) = \varpi_i^{-1}(F_0)\times\{0\}$ and $\widetilde{\varpi}_i^{-1}\big(F_0\times(s,t]\big) = \varpi_i^{-1}(F_0)\times(s,t]$ are also predictable rectangles, where we use that $\mathcal{F}_s^0 \subset \mathcal{F}_s$, it results that $\widetilde{\varpi}_i$ is $\mathcal{Q}$–$\mathcal{Q}_i$ measurable. As $\{Z_{i,s}\circ\varpi_i \in B\} = \widetilde{\varpi}_i^{-1}\big(\{Z_{i,s} \in B\}\big)$, $B \in \bar{\mathbb{B}}^p$, it follows the assertion.

Proof of d). It holds that

$$
\{\tau_\circ\varpi_i \leq t\} \in \varpi_i^{-1}\big(\{\tau_i \leq t\}\big) \in \mathcal{F}_t^0 \subset \mathcal{F}_t, \quad t \in \mathbb{R}_+.
$$

□

B.5.5 Proposition. Let $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$ be some probability space and assume that $\{X(t) \mid t \in \mathcal{T}\}$ is a real valued stochastic process on Ω.

a) Let $\mathcal{T}_0 \subset \mathcal{T}$ be some countable set and assume that for every $t \in \mathcal{T}$ and every $\omega \in \Omega$, there exists a sequence $t_k(\omega, t)$, $k \in \mathbb{N}$, such that

   i) $t_k(\omega, t) \in \mathcal{T}_0$, $k \in \mathbb{N}$,

   ii) $\lim_{k\to\infty} X\big(t_k(\omega, t), \omega\big) = X(t, \omega)$.

   Then it holds that $\sup_{t\in\mathcal{T}}\big|X(t)\big| = \sup_{t\in\mathcal{T}_0}\big|X(t)\big|$. Moreover, $\sup_{t\in\mathcal{T}}\big|X(t)\big|$ is $\mathbb{F}$–$\mathbb{B}$ measurable.

b) Additionally, assume that these sequences are independent of $\omega \in \Omega$, *i.e.* $t_k(t) = t_k(t, \omega)$ for all $\omega \in \Omega$, and that $\mathbb{E}\sup_{t\in\mathcal{T}}\big|X(t)\big| < \infty$, then

$$\lim_{k\to\infty} \mathbb{E}\, X\big(t_k(t)\big) = \mathbb{E}\, X(t) \quad \text{for all } t \in \mathcal{T}.$$

Proof. For every fixed $\omega \in \Omega$, we know that there exists a sequence $s_k \in \mathcal{T}$, $k \in \mathbb{N}$, such that $\lim_{k\to\infty}\big|X(s_k, \omega)\big| = \sup_{t\in\mathcal{T}}\big|X(t, \omega)\big|$ and $\big|X(s_k, \omega)\big| \le \big|X(s_{k+1}, \omega)\big|$, $k \in \mathbb{N}$. Using the assumption, we can find a sequence $t_k \in \mathcal{T}_0$, $k \in \mathbb{N}$, such that

$$\lim_{k\to\infty}\Big|\big|X(s_k, \omega)\big| - \big|X(t_k, \omega)\big|\Big| = 0.$$

This establishes the equality of the suprema. For the measurability, see Bauer [6, Satz 9.5].

Proof of b). As $\big|X\big(t_k(t)\big)\big| \le \sup_{t\in\mathcal{T}}\big|X(t)\big|$, the result is an immediate consequence of the Dominated Convergence Theorem, cf. Bauer[6, Satz 15.1 and Satz 15.6]. $\square$

Index

Bibliography

[1] Aalen, O. Nonparametric inference for a family of counting processes. *The Annals of Statistic 6*, 4 (1978), 701–726.

[2] Andersen, P. K., and Borgan, Ø. Counting process models for life history data: A review (with discussion). *Scandinavian Journal of Statistics 12* (1985), 97–158.

[3] Andersen, P. K., Borgan, Ø., Gill, R. D., and Keiding, N. Linear nonparametric tests for the comparison of counting processes, with application to censored survival data. *International Statistical Reviews 50* (1982), 219–258.

[4] Andersen, P. K., Borgan, Ø., Gill, R. D., and Keiding, N. *Statistical models based on counting processes.* Springer, 1993.

[5] Andersen, P. K., and Gill, R. D. Cox's regression model for counting processes: A large sample study. *The Annals of Statistics 10*, 4 (1982), 1100–1120.

[6] Bauer, H. *Maß- und Integrationstheorie*, 2. ed. Walter de Gruyter, 1992.

[7] Behnen, K., and Neuhaus, G. *Rank tests with estimated scores and their applications.* B.G. Teubner, 1989.

[8] Behnen, K., and Neuhaus, G. Likelihood ratio rank tests for the two-sample problem with randomly censored data. *Kybernetika 27*, 2 (1991), 81–99.

[9] Billingsley, P. *Convergence of Probability Measures.* John Wiley & Sons, 1968.

[10] Borwein, J. M., and Lewis, A. S. *Convex Analysis and Nonlinear Optimization*. Springer, 2000.

[11] Brendel, M. Testing hypotheses under a generalized Koziol-Green model with partially informative censoring. *Statistics 39*, 4 (2005), 329–345.

[12] Brunner, E., and Munzel, U. *Nichtparametrische Datenanalyse*. Springer-Verlag, 2002.

[13] Cox, D. R. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society B 34* (1972), 187–220.

[14] Cox, D. R. Partial likelihood. *Biometrika 62* (1975), 269–276.

[15] Dabrowska, D. M. Smoothed Cox regression. *The Annals of Statistic 25*, 4 (1997), 1510–1540.

[16] Dellacherie, C., and Meyer, P.-A. *Probabilities and Patential*. Hermann and North-Holland Publishing Company, 1978.

[17] Dümbgen, L. Empirische Prozesse. Skripten zur Mathematischen Statistik Nr. 35, 2000.

[18] Eaton, M. L. *Multivariate Statistics*. John Wiley & Sons, 1983.

[19] Fleming, T. R., and Harrington, D. P. *Counting Processes and Survival Analysis*. John Wiley & Sons, 1991.

[20] Gänssler, P., and Stute, W. *Wahrscheinlichkeitstheorie*. Springer-Verlag, 1977.

[21] Gentlemen, R., and Crowley, J. Local full likelihood estimation for the proportional hazard model. *Biometrics 47* (1991), 1283–1296.

[22] Gill, R. D. *Censoring and Stochastic Integral*, 3rd ed. Mathematisch Centrum Amsterdam, 1986.

[23] Grambsch, P. M., and Therneau, T. M. Proportional hazards tests and dianostics based on weighted residuals. *Biometrika 81*, 3 (1994), 515–526.

[24] Graybill, F. A. *Matrices with application in statistics*, 2nd ed. Wadsworth, Inc., 1983.

[25] Greenwood, P. E., and Wefelmeyer, W. Cox's factoring of regression model likelihoods for continuous-time processes. *Bernoulli 4*, 1 (1998), 65–80.

[26] Hájek, J., and Šidák, Z. *Theory of Rank Tests.* Academic Press, Inc., 1967.

[27] Heller, G. The Cox proportional hazard model with a partly linear relative risk function. *Lifetime Data Analysis 7* (2001), 255–277.

[28] Heller, G., and Venkatraman, E. S. A nonparametrix test to compare survival distributions with covariate adjustment. *Journal of the Royal Statistical Society B 66* (2004), 719–733.

[29] Huang, J. Efficient estimation of the partly linear additive Cox model. *The Annals of Statistics 27*, 5 (1999), 1536–1563.

[30] Jacod, J. Multivariate point processes: predictable projection, Randon-Nikodym derivatives, representation of martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 31* (1975), 235–253.

[31] Jacod, J. Partial likelihood process and asymptotic normality. *Stochastic Processes and their Application 26* (1987), 47–71.

[32] Jacod, J., and Shiryaev, A. N. *Limit Theorems for Stochastic Processes.* Springer-Verlag, 1987.

[33] Janssen, A. Local asymptotic normality for randomly censored data with applications to rank tests. *Statistica Neerlandica 43*, 2 (1989), 109–125.

[34] Janssen, A. Optimal k-sample tests for randomly censored data. *Scandinavian Journal of Statistics 18* (1991), 135–152.

[35] Janssen, A. On local odds and hazard rate models in survival analysis. *Statistics & Probability Letters 20* (1994), 355–365.

[36] Janssen, A. Global power functions of goodness of fit tests. *The Annals of Statistic 28*, 1 (2000), 239–253.

[37] Janssen, A., and Mayer, C.-D. Conditional studentized survival tests for randomly censored models. *Scandinavian Journal of Statistics 28* (2001), 283–293.

[38] Janssen, A., and Werft, W. A survey about the efficiency of two-sample survival tests for randomly censored data. *Mitteilungen aus dem Mathematischen Seminar Giessen 254* (2004), 1–47.

[39] Jones, M. P., and Crowley, J. A general class of nonparametric tests for survival analysis. *Biometrics 45* (1989), 157–170.

[40] Jones, M. P., and Crowley, J. Asymptotic properties of a gerneral class of nonparametric tests for survival analysis. *The Annals of Statistics 18*, 3 (1990), 1203–1220.

[41] Kalbfleisch, J. D., and Prentice, R. L. *The Statistical Analysis of Failure Time Data.* John Wiley & Sons, 1980.

[42] Kauermann, G., and Berger, U. A smooth test in proportional hazard survival models using local partial likelihood fitting. *Lifetime Data Analysis 9* (2003), 373–393.

[43] Klein, J. P., and Moeschberger, M. L. *Survival Analysis – Techniques for Censored and Truncated Data*, 2nd ed. Springer, 2003.

[44] Königsberger, K. *Analysis 2*, 2nd ed. Springer, 1997.

[45] Kraus, D. Goodness-of-fit for the Cox-Aalen additive-multiplicative regression model. *Statistics & Probability Letters 70* (2004), 285–298.

[46] LeBlanc, M., and Crowley, J. Adaptive regression splines in the Cox model. *Biometrics 55*, 1 (1999), 204–213.

[47] Lehmann, E. L. *Testing Statistical Hypotheses*, 2nd ed. John Wiley & Sons, 1986.

[48] Lin, C.-Y., and Kosorok, M. R. A general class of function-indexed nonparametric tests for survival analysis. *The Annals of Statistic 27*, 5 (1999), 1722–1744.

[49] Lin, D., and Ying, Z. Semiparametric analysis of general additive-multiplicative models for counting processes. *The Annals of Statistic 23*, 5 (1995), 1712–1734.

[50] Lin, D. Y. Goodness-of-fit analysis for the Cox regression model based on a class of parameter estmators. *Journal of the American Statistical Association 86*, 415 (1991), 725–728.

[51] Martinussen, T., and Scheike, T. H. A flexible additive multiplicative hazard model. *Biometrika 89*, 2 (2002), 283–298.

[52] Martinussen, T., Scheike, T. H., and Skovgaard, I. M. Efficient estimation of fixed and time-varying covariate effects in multiplicative intensity models. *Scandinavian Journal of Statistics 29*, 1 (2002), 57–74.

[53] Mayer, C.-D. *Projektionstests für das Zweistichprobenproblem mit zensierten Daten.* Inauguraldissertation, Heinrich-Heine-Universität Düsseldorf, 1996.

[54] McKeague, I. W., and Utikal, K. J. Identifying nonlinear covariate effects in semimartingale regression models. *Probability Theory and Related Fields 87*, 1 (1990), 1–25.

[55] Murphy, S. A. Testing for time dependent coefficients in Cox's regresion model. *Scandinavian Journal of Statistics 20*, 1 (1993), 35–50.

[56] Murphy, S. A., and Sen, P. K. Time-dependent coefficients in a Cox-type regression model. *Stochastic Processes and their Applications 39*, 1 (1991), 153–180.

[57] Neuhaus, G. Asymptotically optimal rank tests. *Communications in Statistics: Theory and Methods 17*, 6 (1988), 2037–2058.

[58] Neuhaus, G. Einige Kapitel der finiten und asymptotischen Entscheidungstheorie von Le Cam. Skripten zur Mathematischen Statistik Nr. 17, Gesellschaft zur Förderung der Mathematischen Statistik, Münster, 1989.

[59] Neuhaus, G. Conditional rank tests for the two-sample problem under random censorship. *The Annals of Statistics 21*, 4 (1993), 1760–1769.

[60] Neuhaus, G. A method of constructing rank tests in survival analysis. *Journal of Statistical Planning and Inference 91* (2000), 481–497.

[61] Rebolledo, R. Central limit theorems for local martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 51* (1980), 269–286.

[62] Sargent, D. J. A flexible approach to time-varying coefficients in the Cox regression setting. *Lifetime Data Analysis 3* (1997), 13–25.

[63] Scheike, T. H., and Martinussen, T. On estimation and tests of time-varying effects in the proportional hazards model. *Scandinavian Journal of Statistics 31* (2004), 51–62.

[64] Scheike, T. H., and Zhang, M.-J. An additive-multiplicative Cox-Aalen regression model. *Scandinavian Journal of Statistics 29*, 1 (2002), 75–88.

[65] Shen, Y., and Fleming, T. R. Weighted mean survival test statistics: A class of distance tests for survival data. *Journal of the Royal Statistical Society B 59*, 1 (1997), 269–280.

[66] Slud, E. V. Partial likelihoods for continuous-time processes. *Scandinavian Journal of Statistics 19* (1992), 97–109.

[67] Stoer, J. *Numerische Mathematik 1*, 5., verbesserte ed. Springer-Verlag, 1989.

[68] Strasser, H. *Mathematical Theory of Statistics – Statistical Eperiments and Asymptotic Decision Theory*. Walter de Gruyter, 1985.

[69] Tsiatis, A. A. A large sample study of Cox's regression model. *The Annals of Statistic 9*, 1 (1981), 93–108.

[70] Verweij, P. J. M., and van Houwelingen, H. C. Time-dependent effects of fixed covariates in Cox regression. *Biometrics 51*, 4 (1995), 1550–1556.

[71] Witting, H. *Mathematische Statistik I – Parametrische Verfahren bei festem Stichprobenumfang*. B.G. Teubner, 1985.

[72] Witting, H., and Müller-Funk, U. *Mathematische Statistik II – Asymptotische Statistik: Parametrische Modelle und nicht-parametrische Funktionale*. B.G. Teubner, 1995.

[73] Wong, W. H. Theory of partial likelihood. *The Annals of Statistic 14*, 1 (1986), 88–123.

Summary

The determination of the influence of covariates on survival times is a common issue in biomedical research. The interaction between covariates and survival times can be specified by the Modified Cox Regression Model (MCRM). This model incorporates crucial aspects of the popular and frequently applied Cox Regression Model and the basic concept of the rank tests with estimated scores provided by Behnen and Neuhaus. On the basis of localized, parametric sub-models of the MCRM, tests for various hypotheses are rigorously developed.

The considered models are stated as counting process models; therefore a general result on asymptotic normality for such models is discussed and applied to localized, parametric sub-models of the MCRM. Using the likelihood ratio test statistic of the limit experiment, asymptotically unbiased and asymptotically admissible tests are derived.

In order to receive test statistics that are independent of the special choice of the underlying localized, parametric sub-model of the MCRM, sequences of hardest parametric sub-models are considered. In particular, statistical considerations are made to shape and provide a comprehensible and coherent definition of sequences of hardest parametric sub-models.

Examples addressing the applicability of the MCRM are given and the connection to known results is shown. Moreover, the underlying general assumptions are investigated in detail for important special cases. Additionally, a descriptive illustration of the tests is provided by presenting them as projective-type tests.

Finally, a permutation method to determine critical values is introduced. The resulting conditional permutation tests are asymptotically equivalent to the above constructed tests, but keep the level even for finite sample-sizes in certain situations.

Zusammenfassung

Die Bestimmung des Effekts von Kovariablen auf Überlebenszeiten ist eine in der biomedizinischen Forschung häufig auftretende Fragestellung. Das Zusammenspiel zwischen Kovariablen und Überlebenszeiten kann mit dem modifizierten Cox'schen Regressionsmodell (MCRM) beschrieben werden. Dieses Modell verbindet die wesentlichen Aspekte des populären und häufig angewandten Cox'schen Regressionsmodells mit dem Konzept der Rangtests mit geschätzten Gewichten von Behnen und Neuhaus. Auf der Grundlage von lokalisierten, parametrischen Teilmodellen des MCRM werden Tests für verschiedene Hypothesen entwickelt.

Die betrachteten Modelle werden als Zählprozessmodelle formuliert, deshalb wird ein allgemeines Resultat über asymptotische Normalität für solche Modelle erörtert und auf lokalisierte, parametrische Teilmodelle des MCRM angewandt. Unter Verwendung der Likelihood-Quotienten-Teststatistik des Limesexperiments werden asymptotisch unverfälschte und asymptotisch zulässige Tests hergeleitet.

Um Tests zu erhalten, die von einer speziellen Wahl des lokalisierten, parametrischen Teilmodells unabhängig sind, werden Folgen von härtesten parametrischen Teilmodellen betrachtet. Insbesondere wird aufgrund von statistischen Überlegungen eine anschauliche und verständliche Definition der härtesten parametrischen Teilmodelle entwickelt.

Weiterhin werden Beispiele, die die Anwendungsmöglichkeiten des MCRM demonstrieren, diskutiert und die Verbindung zu bekannten Resultaten aufgezeigt. Auch werden die allgemeinen Voraussetzungen für wichtige Spezialfälle näher untersucht. Durch den Nachweis, dass es sich bei den vorgestellten Verfahren um Projektionstests handelt, wird zusätzlich eine anschauliche Deutung der Ergebnisse gegeben.

Abschließend wird eine Permutationsmethode vorgestellt, um kritische Werte für die Tests zu bestimmen. Die so konstruierten bedingten Permutationstests sind asymptotisch äquivalent mit den oben behandelten Tests, aber halten das Niveau bereits bei endlichen Stichprobenumfängen in bestimmten Situationen ein.

Curriculum Vitae

Michael Brendel

geboren am 7. September 1975 in Garmisch-Partenkirchen

verheiratet

Berufliche Tätigkeiten

06/2006 –	Bernhard-Nocht-Institut für Tropenmedizin in Hamburg, wissenschaftlicher Mitarbeiter in der tropenmedizinischen Grundlagenforschung
10/2003	Technische Universität Hamburg-Harburg, Lehrauftrag für den *Mathe-Vorkurs*
04/2003 – 03/2006	Universität Hamburg, Department Mathematik, wissenschaftlicher Mitarbeiter in Forschung und Lehre
10/1999 – 02/2003	Universität Hamburg, Fachbereich Mathematik, studentische Hilfskraft

Studium und Auslandsaufenthalte

05/2005	Gastaufenthalt an der Karls Universität zu Prag
02/2003 – 09/2006	Promotionsstudium am Department Mathematik der Universität Hamburg
10/2001 – 03/2002	Studium an der University of Dundee (United Kingdom) im Rahmen des ERASMUS-Programms
10/1997 – 01/2003	Studium der Wirtschaftsmathematik an der Universität Hamburg, Diplomarbeit: *Testen in einem Modell mit informativen und nicht-informativen Zensierungen* betreut von Prof. Dr. Georg Neuhaus

Zivildienst

07/1996 – 07/1997	Tagespflegestätte der Arbeiterwohlfahrt in Wedel (Holst.)

Schulbildung

08/1992 – 06/1996	Gymnasium Bahrenfeld, Hamburg
08/1986 – 07/1992	Ernst-Barlach-Schule, Realschule, Wedel (Holst.)
08/1982 – 07/1986	Grund- und Hauptschule Holm (Kreis Pinneberg)

Interessen	Naturwissenschaften, Reiseberichte, Kochen, Kunst