Functional weak limit theorem for a local empirical process of non-stationary time series and its application to von Mises-statistics

Dissertation zur Erlangung des Grades des
Doktors der Naturwissenschaften
der Fakultät für Mathematik und Informatik
der Universität des Saarlandes

von

Ulrike Mayer

Saarbrücken

2019

Tag des Kolloquiums:	14.06.2019

Mitglieder des Prüfungsausschusses:

Vorsitzender:	Professor Dr. Vladimir Lazić
1. Berichterstatter:	Professor Dr. Henryk Zähle
2. Berichterstatter:	Professor Dr. Anne Leucht (während des Kolloquiums vertreten durch Professor Dr. Christian Bender)
Protokollführer:	Dr. Dominik Schillo
Dekan:	Professor Dr. Sebastian Hack

Abstract

This thesis is concerned with the asymptotics of a local empirical process of piece-wise locally stationary (PLS) time series. In this context we prove a weak limit theorem that can be seen as analogue of a result for the classical empirical process of stationary time series provided by Wu (2008). The class of PLS time series, based on the locally stationary time series model of Zhou and Wu (2009), is illustrated by means of the PLS linear process and PLS ARCH process.

Moreover, we extend the continuous mapping approach for deriving the asymptotics of V-statistics of Beutner and Zähle (2014) to multi-sample V-statistics of degree d. In combination with the weak limit theorem for the local empirical process, this enables to determine the asymptotic distribution of V-statistics of degree d for non-stationary time series. We further use our extended continuous mapping approach to investigate the asymptotic distribution of the skewness of probability distributions.

In addition, we develop a multivariate integration by parts formula and a Jordan decomposition for functions on $\mathbb{R}^d$ of locally bounded variation, which is required for the extension of the approach of Beutner and Zähle.

Zusammenfassung

Die Arbeit beschäftigt sich mit der Asymptotik eines lokalen empirischen Prozesses stückweise lokal stationärer (PLS) Zeitreihen. In diesem Zusammenhang beweisen wir ein schwaches Grenzwerttheorem, ein Analogon zu einem Resultat für den klassischen empirischen Prozess stationärer Zeitreihen von Wu (2008). Die Klasse der stückweise lokal stationären Zeitreihen, die auf dem lokal stationären Zeitreihenmodell von Zhou and Wu (2009) basiert, wird mittels des PLS linearen Prozesses und des PLS ARCH Prozesses veranschaulicht.

Darüber hinaus erweitern wir den Continuous Mapping-Ansatz von Beutner und Zähle (2014) zur Herleitung der Asymptotik von V-Statistiken auf Mehrfachstichproben-V-Statistiken von Grad d. Kombiniert mit dem schwachen Grenzwerttheorem für den lokalen empirischen Prozess ermöglicht dies, die asymptotische Verteilung der V-Statistiken von Grad d nicht-stationärer Zeitreihen zu bestimmen. Des Weiteren wenden wir unseren erweiterten Continuous Mapping-Ansatz an, um die Asymptotik der Schiefe von Wahrscheinlichkeitsverteilungen zu untersuchen.

Überdies wird eine multivariate partielle Integrationsformel und eine Jordanzerlegung für Funktionen auf $\mathbb{R}^d$ von lokal beschränkter Variation hergeleitet, die zur Erweiterung des Ansatzes von Beutner und Zähle erforderlich sind.

Acknowledgements

First and foremost, I want to express my gratitude towards my supervisor Professor Henryk Zähle for giving me the opportunity to write this thesis (temporarily with financial support by the Federal Ministry of Education and Research (BMBF) raised by him). I deeply value his continuous support and mathematical guidance, his patience and encouragement, and that he made my visiting stay in Toronto possible.

My special thanks also go to Professor Zhou Zhou for the collaboration, for taking so much time to introduce me to non-stationary time series and for his inspiring enthusiasm. Furthermore, I would like to thank Professor Anne Leucht for reviewing this thesis. Many thanks also go to Professor Eric Beutner for the collaboration and valuable discussions.

I am grateful to the German Academic Exchange Service (DAAD) for financial support during my stay at the University of Toronto.

I would like to thank all my colleagues in Saarbrücken and in Toronto, in particular Alexandra Lauer, Patrick Kern, Christian Gärtner, Robert Knobloch, Steffen Meyer, Philip Oberacker, Nikolaus Schweitzer, Matthias Thiel and Wei Lin. Thanks for exchanging views, for mathematical discussions and the necessary amount of distraction.

Many thanks to all my friends for encouraging me and having fun together. Special thanks go to Robert Knobloch, Helge Rhodin and Katharina Rojan for sharing their research experience and for helping to improve this thesis.

Last but not least, I would like to express my deep gratitude towards my family Horst, Birgit, Isabel, Florian and Helge for the endless support in any possible way, for encouraging me and for going with me through all ups and downs.

Contents

Introduction

Locally stationary time series analysis has attracted much attention in the statistics community over the last two decades. In contrast to a stationary time series $\{X_t\}_{t=0,1,2...}$ whose joint probability distributions do not change over time or at least whose second moments are finite for all t and both mean function $\mathbb{E}[X_t]$ and covariance function $\mathbb{C}\mathrm{ov}(X_{t+h}, X_t)$ are independent of t for each h, locally stationary time series merely show a stationary behavior over a short period of time (locally at each point). However, their parameters and covariances are successively changing in an unspecific way.

The study of these non-stationary time series goes back to Priestley [61] who introduced spectral representations of processes that are time-varying (see also [62]). While Priestley's approach describes physically how the process moves on with increasing time, Dahlhaus [19, 20] managed to establish a reasonable asymptotic theory for non-stationary time series. Instead of letting the time parameter tend to infinity, Dahlhaus rescaled the time to the interval $[0, 1]$ by observing the process at points i/n for $i = 1, \dots, n$. Hence, with increasing n more and more data of each local structure is available, which enables the study of asymptotic behavior. From that moment on locally stationary processes have been investigated from different points of view adopting this rescaling technique. While Dahlhaus [20] proposed a class of locally stationary time series based on time-varying spectral representations, Neumann and von Sachs [57] and Nason et al. [56] studied locally stationary time series via the time varying wavelet spectrum. In [84], Zhou and Wu formulated locally stationary time series from the perspective of a time-varying physical system, and Dahlhaus et al. [22] recently combined this approach with stationary approximations to present a general theory for locally stationary time series. We refer to Dahlhaus [21] for a comprehensive survey and additional references.

In this thesis we will investigate locally stationary time series in the sense of Zhou [82] who extended the framework of Zhou and Wu [84] to a class of piece-wise locally stationary time series models allowing both smooth and abrupt changes in the physical system. The latter class of time series includes some common examples. For instance the time-varying linear process and the time-varying ARCH-process can be extended in such a way that they are piece-wise locally stationary, see Subsection 1.2.3. The time-varying linear process was originally introduced in Dahlhaus [20], whereas the time-varying ARCH-process is known from Dahlhaus and Subba Rao [23] and investigated further by Fryzclewicz et al. [35], Fryzclewicz and Subba Rao [36] and others.

Chapter 1 of the thesis is devoted to the study of the asymptotics for the local empirical process of these piece-wise locally stationary time series with respect to a nonuniform sup-norm. Empirical processes play a powerful role in mathematical statistics. As many statistical estimators and test statistics are functionals of an empirical distribution function, weak convergence results for the empirical process serve as fundamental tools for deriving the asymptotics of these functionals by means of methods such as the (extended) continuous mapping theorem or the functional delta method.

To formulate the weak convergence theorem explicitly, let $(X_{n,i})_{i=1}^n$ be a piece-wise locally stationary time series on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that will precisely be defined in Section 1.1. Suppose that we are interested in (a characteristic derived from) the distribution of $X_{n,i_{p,n}}$ for $i_{p,n} := \lfloor pn \rfloor$ for some fixed $p \in (0,1)$. Under some assumptions the distribution function of $X_{n,i_{p,n}}$, denoted by $F_{p,n}$, stabilizes as $n \to \infty$. In Subsection 1.3.1 we will see that indeed $F_{p,n} \to F_p$ for some distribution function F_p in some (nonuniform) sup-norm, provided the assumptions are fulfilled. Thus, under suitable conditions it can be reasonable to use

$$\widehat{F}_{p,n} := c_n \sum_{i=1}^{n} \kappa\Big(\frac{i/n - i_{p,n}/n}{b_n}\Big) \mathbb{1}_{[X_{n,i},\infty)}$$

as an estimator for $F_{p,n}$, where $\kappa : \mathbb{R} \to \mathbb{R}_+$ is a kernel function, $b_n \in \mathbb{R}_+ \setminus \{0\}$ is a bandwidth, and $c_n := 1/\sum_{i=1}^n \kappa((i/n - i_{p,n}/n)/b_n)$ is a normalizing constant. In Chapter 1, we will show that under suitable assumptions

$$\sqrt{nb_n}\big(\widehat{F}_{p,n}(\cdot) - F_{p,n}(\cdot)\big) \rightsquigarrow^* B_p \tag{1}$$

(with respect to a nonuniform sup-norm) for a non-degenerate Gaussian process B_p, where $\rightsquigarrow^*$ means convergence in distribution in the Hoffmann-Jørgensen sense [45]. In fact we will show that under suitable assumptions

$$\sqrt{nb_n}\big(\widehat{F}_{p,n}(\cdot) - \mathbb{E}[\widehat{F}_{p,n}(\cdot)]\big) \rightsquigarrow^* B_p, \tag{2}$$

and we will discuss additional assumptions under which $\sqrt{nb_n}(F_{p,n}(\cdot) - \mathbb{E}[\widehat{F}_{p,n}(\cdot)]) \to 0$ (with respect to a nonuniform sup-norm). The convergence in (2) can be seen as the analogue of Theorem 1 in [78] where a similar result was proven for stationary time series (and with $\widehat{F}_{p,n}$ replaced by the classical empirical distribution function).

On the one hand, (1) yields consistency and the rate of convergence of the function-valued estimator $\widehat{F}_{p,n}(\cdot)$ for the distribution function $F_{p,n}(\cdot)$. On the other hand, in view of tools as the (extended) continuous mapping theorem and the functional delta-method, (1) can also be seen as a building stone for deriving the asymptotic distribution of the empirical plug-in estimator $\mathcal{T}(\widehat{F}_{p,n})$ for some characteristic $\mathcal{T}(F_{p,n})$ derived from $F_{p,n}$. In two specific examples in Section 1.3 the asymptotics of weighted empirical quantiles and weighted von Mises-statistics (or V-statistics for short) of degree 2 will be discussed. The latter example makes use of the continuous mapping approach to V-statistics of degree 2 with kernel functions h in Beutner and Zähle [11] and provides an analogous result to Zhou [83] where V-statistics of degree 2 are studied under similar assumptions from the perspective of Fourier analysis.

In **Chapter 3** we will extend this continuous mapping approach to multi-sample V-statistics of degree $d \geq 2$ with kernel functions h_n depending on n. This enables to study even the asymptotics of weighted V-statistics of degree $d \geq 2$ for non-stationary time series, as we will see in Section 3.4. Apart from that, the asymptotic distribution of V-statistics is a matter of particular interest.

The theory of V-statistics goes back to the 1940s with pioneering publications of Halmos [42], Hoeffding [44] and von Mises [74]. Since that time many weak central limit theorems have been established to determine the asymptotics of V-statistics, where most efforts have been put on stationary sequences of random variables. We refer for instance to Beutner and Zähle [10, 11], Beutner et al. [8], Dehling and Taqqu [25], Dehling and Wendler [26], Dewan and Prakasa Rao [29, 30, 31], Denker and Keller [28], Garg and Dewan [37, 38], Leucht [52], Sen [65], Yoshihara [80] and Zhou [83] for several approaches under various (dependence-) conditions.

To describe multi-sample V-statistics of degree d with kernel function h_n, let us use $\mathcal{V}_{h_n}$ to denote the functional playing the role of $\mathcal{T}$ above. For some Borel measurable kernel function $h_n : \mathbb{R}^d \to \mathbb{R}$ the functional $\mathcal{V}_{h_n}$ is defined by

$$\mathcal{V}_{h_n}(F^{(1)}, \ldots, F^{(d)}) := \int_{\mathbb{R}^d} h_n(x_1, \ldots, x_d)\, (\mu_{F^{(1)}} \otimes \cdots \otimes \mu_{F^{(d)}})(d(x_1, \ldots, x_d))$$

on the set of d-tuples of distribution functions $F^{(1)}, \ldots, F^{(d)}$ on the real line for which the latter integral exists. If $\widehat{F}_n^{(j)}$ is the empirical distribution function of random variables $X_1^{(j)}, \ldots, X_n^{(j)}$, $j = 1, \ldots, d$, then $\mathcal{V}_{h_n}(\widehat{F}_n^{(1)}, \ldots, \widehat{F}_n^{(d)})$ is referred to as d-sample V-statistics of degree d. Chapter 3 is concerned with the question of the asymptotic distribution of $\mathcal{V}_{h_n}(\widehat{F}_n^{(1)}, \ldots, \widehat{F}_n^{(d)})$ or rather of the weak convergence of the empirical error

$$a_n\big(\mathcal{V}_{h_n}(\widehat{F}_n^{(1)}, \ldots, \widehat{F}_n^{(d)}) - \mathcal{V}_{h_n}(F^{(1)}, \ldots, F^{(d)})\big) \tag{3}$$

for some $a_n \to \infty$, if $\widehat{F}_n^{(j)}$ is not necessarily the empirical distribution function but any (suitable) estimator of $F^{(j)}$ for every $j = 1, \ldots, d$ and $n \in \mathbb{N}$. More precisely, in Section 3.3 we will provide a weak central limit theorem for a vector-valued random variable with components being of the form (3) for different kernel functions $h_{n,1} : \mathbb{R}^{d_1} \to \mathbb{R}, \ldots, h_{n,k} : \mathbb{R}^{d_k} \to \mathbb{R}$, which allows to study also the asymptotic distribution of suitable compositions of different V-statistics such as the skewness or kurtosis of probability distributions. The example of the skewness will be discussed in Subsection 3.3.3.

To determine the limit distribution of (3) via our extended continuous mapping approach, we primarily need the limit distribution of $a_n(\widehat{F}_n - F)$. If $\widehat{F}_n$ is the empirical distribution function, several weak limit results for the empirical process $a_n(\widehat{F}_n - F)$ with respect to a nonuniform sup-norm, analogous to (1), can be found in the literature under various (dependence-) conditions, see for instance Arcones and Yu [4], Beutner et al. [8], Shao and Yu [69], Shorack and Wellner [70] and Wu [78]. With our approach, we may thus regain many asymptotic results that exist in the literature concerning one-sample V-statistics of degree $d \geq 2$ for stationary sequences under various dependence conditions. Moreover, weighted V-statistics of degree d for non-stationary time series and also multi-sample V-statistics can be dealt with.

We emphasize that the extended continuous mapping approach is only applicable for kernel functions h_n that are locally of bounded variation. To prove the weak convergence of the empirical error in (3), we will apply the (extended) continuous mapping theorem to a special representation of (3) that we obtain by means of a multivariate integration by parts formula.

In **Chapter 2** we will thus develop an integration by parts formula for multivariable functions of locally bounded variation. For that purpose, we will recall the notions of d-fold monotonically increasing functions and of functions that are locally of bounded d-fold variation and their connections to positive and signed Borel measures on $\mathbb{R}^d$. Moreover, we will prove several auxiliary results including a Jordan decomposition for functions on $\mathbb{R}^d$ that are locally of bounded variation.

The results of the first chapter can also be found in the submitted paper [55], jointly with Professor Henryk Zähle and Professor Zhou Zhou:

> Mayer, U., Zähle, H. and Zhou, Z. (2019). Functional weak limit theorem for a local empirical process of non-stationary time series and its application, *submitted*.

The results of the third chapter are based on joint work with Professor Eric Beutner and Professor Henryk Zähle:

> Beutner, E., Mayer, U. and Zähle, H., project on the "Extended continuous mapping approach to the asymptotics of V-statistics", *work in progress*.

Chapter 1

Functional weak limit theorem for a local empirical process of non-stationary time series

1.1 Introduction

This chapter is devoted to the study of the asymptotics for the local empirical process of piece-wise locally stationary time series. The latter class of time series is based on the approach of Zhou and Wu [84] who formulated locally stationary time series from the perspective of a time-varying physical system. In Zhou [82], the framework in Zhou and Wu [84] was extended to the class of the piece-wise locally stationary (PLS) models of the form (1.2) below by allowing both smooth and abrupt changes in the physical system.

To define our time series model explicitly, we fix a finite partition $0 = p_0 < p_1 < \cdots < p_\ell < p_{\ell+1} = 1$ of the unit interval $[0, 1]$. For every $j = 0, \ldots, \ell$, let $G_j : (p_j, p_{j+1}] \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ be any $(\mathcal{B}((p_j, p_{j+1}]) \otimes \mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}}, \mathcal{B}(\mathbb{R}))$-measurable map. For every $n \in \mathbb{N} := \{1, 2, \ldots\}$, define by

$$\mathfrak{G}_n\big(i, (x_k)_{k\in\mathbb{N}}\big) := \sum_{j=0}^{\ell} G_j\big(i/n, (x_k)_{k\in\mathbb{N}}\big)\,\mathbb{1}_{(p_j,p_{j+1}]}(i/n) \tag{1.1}$$

a time dependent filter $\mathfrak{G}_n : \{1, \ldots, n\} \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$. Then, given a two-sided sequence $\boldsymbol{\epsilon} = (\varepsilon_k)_{k\in\mathbb{Z}}$ of i.i.d. real-valued random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we can define a non-stationary time series $(X_{n,i})_{i=1}^{n}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ by

$$X_{n,i} \;:=\; \mathfrak{G}_n(i, \boldsymbol{\epsilon}_i) \;=\; \sum_{j=0}^{\ell} G_j(i/n, \boldsymbol{\epsilon}_i)\,\mathbb{1}_{(p_j,p_{j+1}]}(i/n), \qquad i = 1, \ldots, n, \tag{1.2}$$

where $\boldsymbol{\epsilon}_i := (\varepsilon_i, \varepsilon_{i-1}, \varepsilon_{i-2}, \ldots)$. For every $j = 1, \ldots, \ell$, this times series is subject to a structural break at the smallest time point i with $i > np_j$. Note that the number of observations between any two adjacent structural break points increases linearly in n.

Under suitable assumptions on $G_0, \dots, G_\ell$ and $\mathbb{P}_{\varepsilon_0}$ such times series are approximatively stationary in every small (relative to n) time range in between adjacent structural break points. Meanwhile the series can experience abrupt changes in its data generating mechanism at break points $p_1, \dots, p_\ell$. Hence the above PLS framework allows for a very flexible modeling of complexly time-varying temporal dynamics with both smooth and abrupt changes. We refer to [82] and [79] for more discussions and examples of the PLS time series models.

Suppose that we are interested in (a characteristic derived from) the distribution of $X_{n,i_{p,n}}$ for $i_{p,n} := \lfloor pn \rfloor$ for some fixed $p \in (0,1)$. For our mathematical results we will assume that $p \notin \{p_1, \dots, p_\ell\}$. Let us use $F_{p,n}$ to denote the distribution function of $X_{n,i_{p,n}}$. Under some assumptions $F_{p,n}$ stabilizes as $n \to \infty$. In Lemma 1.3.1 below we will see that under some assumptions indeed $F_{p,n} \to F_p$ in some (nonuniform) sup-norm, where F_p denotes the distribution function of $\xi_p := \sum_{j=0}^{\ell} G_j(p, \boldsymbol{\epsilon}_0)\mathbb{1}_{(p_j,p_{j+1}]}(p)$. Thus, under suitable conditions it can be reasonable to use

$$\widehat{F}_{p,n} := c_n \sum_{i=1}^{n} \kappa\Big(\frac{i/n - i_{p,n}/n}{b_n}\Big)\mathbb{1}_{[X_{n,i},\infty)} = c_n \sum_{i=1}^{n} \kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big)\mathbb{1}_{[X_{n,i},\infty)} \tag{1.3}$$

as an estimator for $F_{p,n}$, where $\kappa : \mathbb{R} \to \mathbb{R}_+$ is a suitable (kernel) function, $b_n \in \mathbb{R}_+ \setminus \{0\}$ is a bandwidth, and $c_n := 1/\sum_{i=1}^{n} \kappa((i/n - i_{p,n}/n)/b_n)$ is a normalizing constant. In the main result of this chapter, Theorem 1.2.4 in conjunction with Remark 1.2.5, we will show that under suitable assumptions

$$\mathcal{E}_{p,n}(\cdot) := \sqrt{nb_n}\big(\widehat{F}_{p,n}(\cdot) - F_{p,n}(\cdot)\big) \rightsquigarrow^* B_p \tag{1.4}$$

(with respect to a nonuniform sup-norm) for a non-degenerate Gaussian process B_p, where $\rightsquigarrow^*$ means convergence in distribution in the Hoffmann-Jørgensen sense [45]. In fact we will show that under suitable assumptions

$$\widetilde{\mathcal{E}}_{p,n}(\cdot) := \sqrt{nb_n}\big(\widehat{F}_{p,n}(\cdot) - \mathbb{E}[\widehat{F}_{p,n}(\cdot)]\big) \rightsquigarrow^* B_p, \tag{1.5}$$

and we will discuss additional assumptions under which $\sqrt{nb_n}(F_{p,n}(\cdot) - \mathbb{E}[\widehat{F}_{p,n}(\cdot)]) \to 0$ (with respect to a nonuniform sup-norm). Assertion (1.4) yields consistency and the rate of convergence of the function-valued estimator $\widehat{F}_{p,n}(\cdot)$ for the distribution function $F_{p,n}(\cdot)$. Since many statistical estimators and test statistics are functionals of an empirical distribution function, the weak limit result in (1.4) can also be seen as building stone for deriving the asymptotic distribution of the empirical plug-in estimator $\mathcal{T}(\widehat{F}_{p,n})$ for some characteristic $\mathcal{T}(F_{p,n})$ derived from $F_{p,n}$ in view of tools as the (extended) continuous mapping theorem and the functional delta-method. Two specific examples will be discussed in Section 1.3.

The rest of the chapter is organized as follows. In Section 1.2, we present our main result, Theorem 1.2.4. The latter result can be seen as the analogue of Theorem 1 in [78] where a similar statement was proven for stationary time series (and with $\widehat{F}_{p,n}$ replaced by the classical empirical distribution function). The imposed assumptions, that might look somewhat cumbersome at first glance, are in line with the assumptions imposed by Wu [78] in the stationary case. We will demonstrate that they are satisfied by two relevant PLS time series models, namely PLS linear processes and PLS ARCH processes in Subsection 1.2.3. In Section 1.3, the functional weak limit theorem of Theorem 1.2.4 is applied to derive the asymptotic distribution of point estimators for quantiles and von Mises-characteristics of $F_{p,n}$. The proof of Theorem 1.2.4 is carried out in Section 1.4. All the others results will be proven in Section 1.5.

1.2 Main result

1.2.1 Physical dependence measure revisited

Before presenting our main result, we recall the definition of the physical dependence measure introduced by Wu [76] and extended by Zhou and Wu [84]. The dependence measure (more precisely the objects introduced in (1.6) and (1.7) below) will appear in assumptions (A5) and (A8) in Subsection 1.2.2. Let ε^* be a real-valued random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{P}_{\varepsilon_0} = \mathbb{P}_{\varepsilon^*}$ and being independent of $\boldsymbol{\epsilon} = (\varepsilon_k)_{k\in\mathbb{Z}}$. If necessary, consider an enlargement of $(\Omega, \mathcal{F}, \mathbb{P})$. For every $i \in \mathbb{Z}$ and $r \in \mathbb{N}$, let

$$\boldsymbol{\epsilon}^*_{i,i-r} := (\varepsilon_i, \varepsilon_{i-1}, \ldots, \varepsilon_{i-r+1}, \varepsilon^*, \varepsilon_{i-r-1}, \ldots).$$

Note that $\boldsymbol{\epsilon}^*_{i,i-r}$ is a coupled version of $\boldsymbol{\epsilon}_i$ with ε_{i-r} replaced by the i.i.d. copy ε^*. Let $I \subseteq \mathbb{R}$ be an interval, and $H : I \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ be any $(\mathcal{B}(I) \otimes \mathcal{B}(\mathbb{R})^{\otimes\mathbb{N}}, \mathcal{B}(\mathbb{R}))$-measurable map. For any $r \in \mathbb{N}$, $q > 0$, and $t \in I$, the *physical dependence measure* (associated with $H(t, \cdot\,)$ and $\boldsymbol{\epsilon}$) is defined by

$$\delta_{\boldsymbol{\epsilon},r;q}(H;t) := \big\|H(t, \boldsymbol{\epsilon}_0) - H(t, \boldsymbol{\epsilon}^*_{0,-r})\big\|_q, \tag{1.6}$$

where $\|\cdot\|_q := \mathbb{E}[|\cdot|^q]^{1/q}$. Moreover, for any $r \in \mathbb{N}$ and $q > 0$, the *physical dependence measure* (associated with H and $\boldsymbol{\epsilon}$) is defined by

$$\delta_{\boldsymbol{\epsilon},r;q}(H) := \sup_{t\in I} \delta_{\boldsymbol{\epsilon},r;q}(H;t). \tag{1.7}$$

Note that $\delta_{\boldsymbol{\epsilon},r;q}(H;t)$ and $\delta_{\boldsymbol{\epsilon},r;q}(H)$ will not change if in (1.6) $\boldsymbol{\epsilon}_0$ and $\boldsymbol{\epsilon}^*_{0,-r}$ are replaced by $\boldsymbol{\epsilon}_k$ and $\boldsymbol{\epsilon}^*_{k,k-r}$, respectively, for any $k \in \mathbb{Z} \setminus \{0\}$. According to [76], the time series model (1.2) can be seen as a time-varying physical system with $\boldsymbol{\epsilon}_i$ being the input and $G_j(i/n, \boldsymbol{\epsilon}_i)$ being the output (if $i/n \in (p_j, p_{j+1}]$), where G_j serves as filter or as transform.

From this perspective, $\delta_{\boldsymbol{\epsilon},r;q}(G_j)$ quantifies the dependence of $G_j(i/n, \boldsymbol{\epsilon}_i)$ on ε_{i-r} for any $i = 1, \dots, n$ by measuring uniformly in t the distance between $G_j(t, \boldsymbol{\epsilon}_i)$ and the coupled version $G_j(t, \boldsymbol{\epsilon}^*_{i,i-r})$. The following Example 1.2.1 was already discussed on page 6 in [82].

Example 1.2.1 In the setting of Section 1.1, assume that specifically $G_j(\pi, (x_k)_{k\in\mathbb{N}}) := \sum_{s=0}^{\infty} a_{j,s}(\pi) x_{i+s}$ for some arbitrary functions $a_{j,s} : (p_j, p_{j+1}] \to \mathbb{R}$, $s \in \mathbb{N}_0$. Then

$$\delta_{\boldsymbol{\epsilon},r;q}(G_j) \;=\; \sup_{\pi\in(p_j,p_{j+1}]} \|a_{j,r}(\pi)(\varepsilon^*_{-r} - \varepsilon_{-r})\|_q \;\leq\; 2\|\varepsilon_0\|_q \sup_{\pi\in(p_j,p_{j+1}]} |a_{j,r}(\pi)|$$

for every $r \in \mathbb{N}$ and $q > 0$. ◇

1.2.2 Assumptions and main result

As already mentioned in Section 1.1, our main result (Theorem 1.2.4 below) is a variant of Theorem 1 in [78]. In the latter theorem, Wu studied the case of stationary time series (i.e. $\ell = 0$ and G_0 independent of the first argument), where the role of $\widehat{F}_{p,n}$ was played by the classical empirical distribution function. For our result we will impose nine assumptions, (A1)–(A9). Assumptions (A7) and (A8) are the analogues of Wu's assumptions (6) and (7), respectively. Assumption (A3) is the analogue of a moment condition on the marginal distribution of the time series in [78], and the analogue of (A6) was tacitly assumed in [78]. The additional assumptions (A1), (A2), (A4), and (A9) are due to the non-stationarity of our underlying time series model, and the additional assumption (A5) is a short range dependence condition.

In Theorem 1.2.4 below we will assume that the following conditions (A3), (A7), and (A8) hold for a common $\lambda \geq 0$. Thus let $\lambda \geq 0$ be arbitrary but fixed. We will frequently use the function $\phi_s : \mathbb{R} \to [1, \infty)$ defined by $\phi_s(x) := (1 + |x|)^s$ for different $s \in \mathbb{R}$. We will also use the corresponding nonuniform sup-norm $\|\cdot\|_{(s)}$ defined by $\|v\|_{(s)} := \|v\phi_s\|_\infty$ with $\|v\|_\infty := \sup_{x\in\mathbb{R}} |v(x)|$. Please do not confuse the nonuniform sup-norm $\|\cdot\|_{(s)}$ for real-valued functions on $\mathbb{R}$ with the L^q-norm $\|\cdot\|_q := \mathbb{E}[|\cdot|^q]^{1/q}$ for random variables on $(\Omega, \mathcal{F}, \mathbb{P})$.

Regarding the kernel and the bandwidth we make the following assumptions.

(A1) The kernel function κ is twice continuously differentiable on $\mathbb{R}$ with support $[-1, 1]$ and (without loss of generality) $\int_{\mathbb{R}} \kappa(u)\, du = 1$.

(A2) $\lim_{n\to\infty} nb_n = \infty$ and $\lim_{n\to\infty} b_n = 0$.

Let j_p be the unique index j with $p \in (p_j, p_{j+1})$. Then we have for n sufficiently large (depending only on p_{j_p} and p_{j_p+1}) that $i_{p,n}/n \in (p_{j_p}, p_{j_p+1})$. For every $n \in \mathbb{N}$ we use $I_{n;p}$ to denote the set of all $i \in \{1, \dots, n\}$ with $i/n \in (p_{j_p}, p_{j_p+1})$. We make the following assumptions.

(A3) The distribution of $X_{n,i}$ has a Lebesgue density $f_{n,i}$ for any $i = 1, \dots, n$ and $n \in \mathbb{N}$, and $\sup_{n\in\mathbb{N}} \max_{1\le i\le n} \|f_{n,i}\|_{(\gamma)} < \infty$ for some $\gamma \in (2\lambda+1, \infty)$.

(A4) $\|G_{j_p}(\pi, \boldsymbol{\epsilon}_0) - G_{j_p}(\pi', \boldsymbol{\epsilon}_0)\|_1 \le C_p\, |\pi - \pi'|$ for all $\pi, \pi' \in (p_{j_p}, p_{j_p+1}]$, and some $C_p > 0$.

(A5) $\delta_{\boldsymbol{\epsilon},r;q}(G_{j_p}) = \mathcal{O}(a^r)$ in $r \in \mathbb{N}$, for some constants $a \in [0,1)$ and $q \in (2, \infty)$.

Here $\delta_{\boldsymbol{\epsilon},r;q}$ refers to the physical dependence measure as defined in (1.7). Thus assertion (A5) means that $\delta_{\boldsymbol{\epsilon},r;q}(G_{j_p};\pi)$ decays exponentially in r uniformly in $\pi \in (p_{j_p}, p_{j_p+1}]$.

Now, denote by $\mathbb{P}_{X_{n,i}\|\boldsymbol{\epsilon}_{i-1}}$ a factorized regular version of the conditional distribution of $X_{n,i}$ (w.r.t. $\mathbb{P}$) given $\boldsymbol{\epsilon}_{i-1}$, i.e. a probability kernel satisfying $\mathbb{P}_{X_{n,i}\|\boldsymbol{\epsilon}_{i-1}}(\boldsymbol{x}, B) = \mathbb{P}[X_{n,i} \in B \| \boldsymbol{\epsilon}_{i-1} = \boldsymbol{x}]$ for $\mathbb{P}_{\boldsymbol{\epsilon}_{i-1}}$-a.e. $\boldsymbol{x} \in \mathbb{R}^{\mathbb{N}}$, for all $B \in \mathcal{B}(\mathbb{R})$. Define a map $\mathfrak{F}_{n,i} : \mathbb{R} \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ by

$$\mathfrak{F}_{n,i}(x, \boldsymbol{x}) := \mathbb{P}_{X_{n,i}\|\boldsymbol{\epsilon}_{i-1}}(\boldsymbol{x}, (-\infty, x]) \quad \big(= \mathbb{E}[\mathbb{1}_{(-\infty,x]}(X_{n,i}) \| \boldsymbol{\epsilon}_{i-1} = \boldsymbol{x}] \big),$$

which we refer to as *factorized conditional distribution function* of $X_{n,i}$ given $\boldsymbol{\epsilon}_{i-1}$. If $x \mapsto \mathfrak{F}_{n,i}(x, \boldsymbol{x})$ is twice differentiable for $\mathbb{P}_{\boldsymbol{\epsilon}_{i-1}}$-a.e. $\boldsymbol{x} \in \mathbb{R}^{\mathbb{N}}$, then we may define maps $\mathfrak{f}_{n,i} : \mathbb{R} \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ and $\mathfrak{f}'_{n,i} : \mathbb{R} \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ by

$$\mathfrak{f}_{n,i}(x, \boldsymbol{x}) := \begin{cases} \frac{\partial}{\partial x}\mathfrak{F}_{n,i}(x, \boldsymbol{x}) & , \quad \boldsymbol{x} \notin N_{i-1} \\ 0 & , \quad \boldsymbol{x} \in N_{i-1} \end{cases} \qquad \text{and} \qquad \mathfrak{f}'_{n,i}(x, \boldsymbol{x}) := \frac{\partial}{\partial x}\mathfrak{f}_{n,i}(x, \boldsymbol{x}),$$

respectively, where $N_{i-1} \in \mathcal{B}(\mathbb{R})^{\otimes\mathbb{N}}$ is the respective $\mathbb{P}_{\boldsymbol{\epsilon}_{i-1}}$-null set. In this case, we refer to $\mathfrak{f}_{n,i}$ as *factorized conditional density* of $X_{n,i}$ given $\boldsymbol{\epsilon}_{i-1}$, and to $\mathfrak{f}'_{n,i}$ as its derivative. We make the following assumptions, where $\delta_{\boldsymbol{\epsilon},r;2}$ is defined as in (1.6) and the constant q in (A7) might differ from the constant q in (A5).

(A6) For any $n \in \mathbb{N}$ and $i \in I_{n;p}$, the factorized conditional distribution function $x \mapsto \mathfrak{F}_{n,i}(x, \boldsymbol{x})$ is twice continuously differentiable for $\mathbb{P}_{\boldsymbol{\epsilon}_{i-1}}$-a.e. $\boldsymbol{x} \in \mathbb{R}^{\mathbb{N}}$.

(A7) For some $q \in (2, \infty)$ we have $\lim_{w\to\infty} M_q(\mathbb{R}\setminus(-w, w)) = 0$ and $M_q(\mathbb{R}) < \infty$, where

$$M_q(J) \;:=\; \sup_{n\in\mathbb{N}} \int_J \max_{i\in I_{n;p}} \|\mathfrak{f}_{n,i}(x, \boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\, \phi_{q\lambda-1+q/2}(x)\, dx.$$

(A8) For some $\alpha \in [0,1]$ and $\beta \in (0, \infty)$ we have $\lim_{w\to\infty} M_{i,\alpha}(\mathbb{R} \setminus (-w, w)) = 0$ and $M_{i,\alpha}(\mathbb{R}) < \infty$ for $i = 1, 2$ as well as $M_\beta(\mathbb{R}) < \infty$, where

$$M_{1,\alpha}(J) \;:=\; \sup_{n\in\mathbb{N}} \sum_{r=1}^{\infty} \Big\{ \max_{i\in I_{n;p}} \int_J \delta^2_{\boldsymbol{\epsilon},r-1;2}(\mathfrak{F}_{n,i}; x)\, \phi_{2\lambda-\alpha}(x)\, dx \Big\}^{1/2},$$

$$M_{2,\alpha}(J) \;:=\; \sup_{n\in\mathbb{N}} \sum_{r=1}^{\infty} \Big\{ \max_{i\in I_{n;p}} \int_J \delta^2_{\boldsymbol{\epsilon},r-1;2}(\mathfrak{f}_{n,i}; x)\, \phi_{2\lambda+\alpha}(x)\, dx \Big\}^{1/2},$$

$$M_\beta(\mathbb{R}) \;:=\; \sup_{n\in\mathbb{N}} \sum_{r=1}^{\infty} \Big\{ \max_{i\in I_{n;p}} \int_{\mathbb{R}} \delta^2_{\boldsymbol{\epsilon},r-1;2}(\mathfrak{f}'_{n,i}; x)\, \phi_{-\beta}(x)\, dx \Big\}^{1/2}.$$

(A9) The distribution of $\xi_p := G_{j_p}(p, \boldsymbol{\epsilon}_0)$ has a bounded Lebesgue density f_p.

Before stating our main result (Theorem 1.2.4), we present two lemmas which are needed for (the statement of) the main result.

Lemma 1.2.2 *Let $\kappa_2 := \int \kappa(x)^2\, dx$ and assume that (A1)–(A5) and (A9) hold. Then*

$$\gamma_p(x, y) := \kappa_2 \sum_{k=-\infty}^{\infty} \mathbb{C}\mathrm{ov}\Big(\mathbb{1}_{(-\infty,x]}\big(G_{j_p}(p, \boldsymbol{\epsilon}_k)\big), \mathbb{1}_{(-\infty,y]}\big(G_{j_p}(p, \boldsymbol{\epsilon}_0)\big)\Big) \tag{1.8}$$

is well-defined for any $x, y \in \mathbb{R}$, and the mapping $(x, y) \mapsto \gamma_p(x, y)$ is symmetric and positive semi-definite. Moreover $\lim_{n\to\infty} \mathbb{E}[\widetilde{\mathcal{E}}_{p,n}(x)\widetilde{\mathcal{E}}_{p,n}(y)] = \gamma_p(x, y)$ for any $x, y \in \mathbb{R}$.

As a consequence of Lemma 1.2.2 there exists a centered Gaussian process with covariance function γ_p. This Gaussian process (respectively a suitable modification of it) will play the role of the limiting process in Theorem 1.2.4 below. Convergence in distribution will take place in a suitable càdlàg space. As càdlàg spaces are nonseparable w.r.t. sup-norms, we regard convergence in distribution as convergence in distribution "w.r.t. the open-ball σ-algebra" (in symbols $\rightsquigarrow^\circ$) as used in [60, 70]; see also [15, Section 1.6] and the Appendices of [13, 14] for further details on this sort of convergence. Let $\boldsymbol{D}_{(\lambda)}$ be the set of all bounded càdlàg functions $v : \mathbb{R} \to \mathbb{R}$ with $\lim_{x\to\pm\infty} v(x) = 0$ and $\|v\|_{(\lambda)}\, (= \sup_{x\in\mathbb{R}} |v(x)|\phi_\lambda(x)) < \infty$. We equip $\boldsymbol{D}_{(\lambda)}$ with the nonuniform sup-norm $\|\cdot\|_{(\lambda)}$ and the corresponding open-Ball σ-algebra $\mathcal{B}^\circ_{(\lambda)}$. The latter is known to coincide with the σ-algebra generated by the one-dimensional coordinate projections; see e.g. Lemma 4.1 in [13].

Lemma 1.2.3 *Assume that assumptions (A1)–(A5) and (A9) hold and let γ_p be defined as in (1.8). Then any centered Gaussian process with covariance function γ_p possesses a modification whose paths all lie in the set $\boldsymbol{C}_{(\lambda)}$ of all continuous elements of $\boldsymbol{D}_{(\lambda)}$.*

Lemma 1.2.3 ensures that we may and do assume that the Gaussian limiting process in the following theorem takes values only in a separable and measurable subset of $\boldsymbol{D}_{(\lambda)}$. This is crucial for the claim of the theorem. The processes $\mathcal{E}_{p,n}$ and $\widetilde{\mathcal{E}}_{p,n}$ were defined in (1.4) and (1.5), respectively.

Theorem 1.2.4 *If conditions (A1)–(A9) hold true for some common $\lambda \geq 0$, then*

$$\widetilde{\mathcal{E}}_{p,n}(\cdot) \rightsquigarrow^\circ B_p \qquad \text{in } (\boldsymbol{D}_{(\lambda)}, \mathcal{B}^\circ_{(\lambda)}, \|\cdot\|_{(\lambda)}) \tag{1.9}$$

for a continuous centered Gaussian process B_p with covariance function γ_p as defined in (1.8). In particular, if we assume in addition $\sqrt{nb_n}\|F_{p,n}(\cdot) - \mathbb{E}[\widehat{F}_{p,n}(\cdot)]\|_{(\lambda)} \to 0$,

$$\mathcal{E}_{p,n}(\cdot) \rightsquigarrow^\circ B_p \qquad \text{in } (\boldsymbol{D}_{(\lambda)}, \mathcal{B}^\circ_{(\lambda)}, \|\cdot\|_{(\lambda)}). \tag{1.10}$$

Remark 1.2.5 As the limiting process B_p in (1.9) and (1.10) is continuous, we may replace in either case $\rightsquigarrow^\circ$ by convergence in distribution in the Hoffmann-Jørgensen sense [45] (usually denoted by $\rightsquigarrow^*$). This is ensured by part (i) of Theorem 1.7.2 in [73]. $\diamond$

The following Lemma 1.2.6 provides sufficient conditions for the additional condition in the second part of Theorem 1.2.4 to hold. It involves the following two conditions.

(B2) $\lim_{n\to\infty} nb_n^{(3q+1)/(q+1)} = 0$.

(B4) $\|G_{j_p}(\pi, \boldsymbol{\epsilon}_0) - G_{j_p}(\pi', \boldsymbol{\epsilon}_0)\|_q \le C_{p,q}\,|\pi - \pi'|$ for all $\pi, \pi' \in (p_{j_p}, p_{j_p+1}]$.

Note that conditions (A2) and (B2) on the bandwidth b_n are simultaneously fulfilled if, for instance, $b_n = n^{-\beta}$ for some $\beta \in (\frac{q+1}{3q+1}, 1)$.

Lemma 1.2.6 *If (B2), (A3), (B4) hold true for some $\lambda \in [0,\infty)$, $q \in [\lambda,\infty) \cap (0,\infty)$, $C_{p,q} \in [0,\infty)$, then $\lim_{n\to\infty} \sqrt{nb_n}\,\|\mathbb{E}[\widehat{F}_{p,n}] - F_{p,n}\|_{(\lambda)} = 0$.*

The proofs of Theorem 1.2.4 and Lemmas 1.2.2, 1.2.3, 1.2.6 will be carried out in Sections 1.4 and 1.5. There we will avail the projection operator $P_k(\cdot) : L^1(\Omega, \mathcal{F}, \mathbb{P}) \to L^1(\Omega, \sigma(\boldsymbol{\epsilon}_k), \mathbb{P})$ defined by

$$P_k(Z) := \mathbb{E}[Z|\boldsymbol{\epsilon}_k] - \mathbb{E}[Z|\boldsymbol{\epsilon}_{k-1}] \tag{1.11}$$

for any fixed $k \in \mathbb{Z}$. In the proofs we will also frequently use that under (A1) and (A2)

$$c_n = \mathcal{O}\big((nb_n)^{-1}\big) \qquad \big(\text{in particular } c_n\sqrt{nb_n} = \mathcal{O}\big((nb_n)^{-1/2}\big)\big), \tag{1.12}$$

which follows from $\sum_{i=1}^n \kappa(\frac{i-i_{p,n}}{nb_n}) = nb_n \int_{-1}^{+1} \kappa(u)\,du + \mathcal{O}(1)$ under (A1) and (A2).

1.2.3 Illustrating examples

PLS linear processes

Let for any $j = 0, \ldots, \ell$ specifically $G_j(\pi, (x_k)_{k\in\mathbb{N}}) := \sum_{s=0}^\infty a_{j,s}(\pi)x_{i+s}$ for some functions $a_{j,s} : (p_j, p_{j+1}] \to \mathbb{R}$, $s \in \mathbb{N}_0$ as in Example 1.2.1. In this case the corresponding process $(X_{n,i})_{i=1}^n$ can be seen as a piecewise locally stationary linear process. Without loss of generality we assume $a_{j,0} \equiv 1$.

Corollary 1.2.7 *Let assumptions (A1) and (A2) be fulfilled. Assume that $a_{j_p,k}$ is continuously differentiable on $(p_{j_p}, p_{j_p+1}]$ for any $k \in \mathbb{N}$, and that the distribution of ε_0 has a Lebesgue density f_ε that is twice continuously differentiable. Moreover assume that for some given $\lambda \in [0,\infty)$ the following assertions hold.*

(a) $\sum_{k=1}^{\infty} \sup_{\pi\in(p_j,p_{j+1}]} |a_{j,k}(\pi)| < \infty$, $j = 0,\ldots,\ell$, *and* $\sup_{\pi\in(p_{j_p},p_{j_p+1}]} |a_{j_p,k}(\pi)| = \mathcal{O}(a^k)$ *for some* $a \in [0,1)$.

(b) $\sum_{k=1}^{\infty} \sup_{\pi\in(p_{j_p},p_{j_p+1}]} |a'_{j_p,k}(\pi)| < \infty$.

(c) $\|f_\varepsilon\|_{(\gamma)} < \infty$ *for some* $\gamma \in (2\lambda+5,\infty)$.

(d) $\|f'_\varepsilon\|_{(\lambda+1)} < \infty$ *and* $\|f''_\varepsilon\|_{(1-\lambda)} < \infty$.

Then (1.9) holds true. Moreover, if in addition condition (B2) is satisfied for $q := 2\lambda+4$, *then also (1.10) holds true.*

In the proof of Corollary 1.2.7 in Subsection 1.5.5, we will show that the assumptions of the corollary imply (A3)–(A9) and (B4).

PLS ARCH processes

Recall that the filters $\mathfrak{G}_n$, $n \in \mathbb{N}$, introduced in (1.1) are generated by $G_0,\ldots,G_\ell$, and that $\boldsymbol{\epsilon} = (\varepsilon_k)_{k\in\mathbb{Z}}$ is a two-sided sequence of i.i.d. real-valued random variables on some probability space $(\Omega,\mathcal{F},\mathbb{P})$. Assume that ε_k, $k \in \mathbb{Z}$, are nonnegative and that

$$G_j(\pi,\boldsymbol{x}_i) = \Big(a_{j,0}(\pi)+\sum_{s=1}^{\mathcal{P}} a_{j,s}(\pi)G_j(\pi,\boldsymbol{x}_{i-s})\Big)x_i \quad \text{for any } \pi \in (p_j,p_{j+1}], \quad \mathbb{P}_{\boldsymbol{\epsilon}}\text{-a.e. } \boldsymbol{x} \in \mathbb{R}^{\mathbb{Z}} \tag{1.13}$$

for any $j = 0,\ldots,\ell$ and $i \in \mathbb{N}$. Here, $\mathcal{P} \in \mathbb{N}$ is fixed, $a_{j,s} : [p_j,p_{j+1}] \to \mathbb{R}_+$, $s = 0,\ldots,\mathcal{P}$, are any functions, and $\boldsymbol{x} := (x_k)_{k\in\mathbb{Z}}$ as well as $\boldsymbol{x}_i := (x_i,x_{i-1},x_{i-2},\ldots)$. The existence of such functions $G_0,\ldots,G_\ell$ under certain restrictions on $a_{j,s}$ and ε_0 will be provided in Lemma 1.2.8 below. In this case, we have in particular

$$G_j(i/n,\boldsymbol{\epsilon}_i) = \rho_{n,i,j}\,\varepsilon_i \quad \mathbb{P}\text{-a.s.,} \quad \text{where} \quad \rho_{n,i,j} := a_{j,0}(i/n) + \sum_{s=1}^{\mathcal{P}} a_{j,s}(i/n)G_j(i/n,\boldsymbol{\epsilon}_{i-s}) \tag{1.14}$$

for any $j = 0,\ldots,\ell$, $n \in \mathbb{N}$, and $i = 1,\ldots,n$ with $i/n \in (p_j,p_{j+1}]$. If no structural break is possible (i.e. $\ell = 0$), then (1.14) can be seen as a variant of the time-varying ARCH (tvARCH) model introduced by Dahlhaus and Subba Rao [23] (and developed further by Fryzlewicz et al. [35], Fryzlewicz and Subba Rao[36], and others). In the latter references the roles of $\rho_{n,i,0}$ and $G_0(i/n,\boldsymbol{\epsilon}_{i-s})$ are played by σ_i^2 and X_{i-s}^2 respectively (similarly as in [41, p. 4] in the stationary case). However we do not only allow for smooth but also for abrupt changes of the coefficients (i.e. $\ell \geq 1$).

As before let $X_{n,i}$ be defined by (1.2) (with $G_0,\ldots,G_\ell$ defined by (1.15) below). In view of (1.14) and the preceding comments, we refer to the process $(X_{n,i})_{i=1}^n$ as PLS ARCH($\mathcal{P}$) process. With regard to applications one might think of $X_{n,i}$ for instance as the absolute value or squared value of an asset return.

Let us give a criterion for (1.13) to be valid (see Lemma 1.2.8 below). To this end let $v_{(1)}$ refer to the first entry of a vector $v \in \mathbb{R}^{\mathcal{P}}$ and set

$$\bar{b}_j(\pi, x) := \begin{bmatrix} a_{j,0}(\pi)x \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \; A_j(\pi, x) := \begin{bmatrix} a_{j,1}(\pi)x & a_{j,2}(\pi)x & \dots & a_{j,\mathcal{P}-1}(\pi)x & a_{j,\mathcal{P}}(\pi)x \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

for any $\pi \in [0, 1]$ and $x \in \mathbb{R}$. Under the validity of assertion (i) of Lemma 1.2.8 below we may define a function $G_j : [p_j, p_{j+1}] \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ by

$$G_j\big(\pi, (x_k)_{k\in\mathbb{N}}\big) \tag{1.15}$$
$$:= \begin{cases} \bar{b}_j(\pi, x_1)_{(1)} + \Big(\sum_{r=0}^{\infty} \Big\{ \prod_{t=0}^{r} A_j(\pi, x_{t+1}) \Big\} \bar{b}_j(\pi, x_{r+2})\Big)_{(1)} & , \quad (x_k)_{k\in\mathbb{N}} \notin N \\ 0 & , \quad (x_k)_{k\in\mathbb{N}} \in N \end{cases}$$

for some suitable $\mathbb{P}_{\boldsymbol{\epsilon}}$-null set N. In this case we have

$$G_j(\pi, \boldsymbol{\epsilon}_i) \tag{1.16}$$
$$= \bar{b}_j(\pi, \varepsilon_i)_{(1)} + \sum_{r=0}^{\infty} \Big\{ \Big[\prod_{s=0}^{r} A_j(\pi, \varepsilon_{i-s}) \Big] \bar{b}_j(\pi, \varepsilon_{i-r-1}) \Big\}_{(1)} \quad \text{for any } \pi \in (p_j, p_{j+1}], \; \mathbb{P}\text{-a.s.}$$

for any $j = 0, \dots, \ell$ and $i \in \mathbb{N}$. Note that (1.16) is in line with the vector representation of ARCH and GARCH processes considered in [5, 16, 36, 71] and others.

In the following lemma we mean by solution to (1.13) a measurable map $G_j : [p_j, p_{j+1}] \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ for which (1.13) holds for any $i \in \mathbb{N}$. We say that two solutions G_j and H_j generate the same samples almost surely if $G_j(i/n, \boldsymbol{\epsilon}_i) = H_j(i/n, \boldsymbol{\epsilon}_i)$ for all $j = 0, \dots, \ell, n \in \mathbb{N}$ and $i = 1, \dots, n$ with $i/n \in [p_j, p_{j+1}]$ $\mathbb{P}$-a.s. The proof of the lemma can be found in Subsection 1.5.6.

Lemma 1.2.8 *Assume that* $\|\varepsilon_0\|_q \max_{j=0,\dots,\ell} \sum_{s=0}^{\mathcal{P}} \sup_{\pi\in[p_j,p_{j+1}]} a_{j,s}(\pi) < 1$ *for some* $q \in [1, \infty)$*. Then for any* $j = 0, \dots, \ell$ *the following assertions hold true.*

(i) *For any fixed* $t \in \mathbb{N}$, $\| \sup_{\pi\in[p_j,p_{j+1}]} \sum_{r=0}^{\infty} \{[\prod_{s=0}^{r} A_j(\pi, \varepsilon_{t-s})] \bar{b}_j(\pi, \varepsilon_{t-r-1})\}_{(1)} \|_q < \infty$ *and, in particular,* $\mathbb{P}$*-a.s. the series* $\sum_{r=0}^{\infty} \{[\prod_{s=0}^{r} A_j(\pi, \varepsilon_{t-s})] \bar{b}_j(\pi, \varepsilon_{t-r-1})\}_{(1)}$ *converges for any* $\pi \in [p_j, p_{j+1}]$.

(ii) *The function* G_j *defined by (1.15) is a solution of (1.13).*

(iii) *If another solution* H_j *of (1.13) satisfies* $\|H_j(i/n, \boldsymbol{\epsilon}_0)\|_q < \infty$ *for all* $n \in \mathbb{N}$ *and* $i = 1, \dots, n$ *with* $i/n \in [p_j, p_{j+1}]$*, then* H_j *and* G_j *generate the same samples almost surely.*

Corollary 1.2.9 *Let assumptions (A1) and (A2) be fulfilled. Assume that $a_{j_p,s}$ is continuously differentiable on $[p_{j_p}, p_{j_p+1}]$ for any $s = 0, \ldots, \mathcal{P}$, and that the distribution of ε_0 has a Lebesgue density f_ε that is twice continuously differentiable. Moreover assume that for some given $\lambda \in [0, \infty)$ the following assertions hold.*

(a) $\min_{j=0,\ldots,\ell} \inf_{\pi \in (p_{j_p}, p_{j_p+1}]} a_{j,0}(\pi) > 0$, *and* $\|\varepsilon_0\|_q \max_{j=0,\ldots,\ell} \sum_{s=0}^{\mathcal{P}} \sup_{\pi \in (p_j, p_{j+1}]} a_{j,s}(\pi) < 1$ *for some* $q \in (4\lambda + 2, \infty)$.

(b) $\|f_\varepsilon\|_{(\gamma)} + \|f'_\varepsilon\|_{(\gamma)} < \infty$ *for some* $\gamma \in (2\lambda + 1, \infty)$.

(c) $\|f''_\varepsilon\|_{(0)} < \infty$.

Then (1.9) holds true. Moreover, if in addition condition (B2) is satisfied for q from assumption (a), then also (1.10) holds true.

In the proof of Corollary 1.2.9 in Subsection 1.5.7, we will show that the assumptions of the corollary imply (A3)–(A9) and (B4).

1.3 Applications

1.3.1 A preliminary result

Theorem 1.2.4 and Lemma 1.2.6 show that the convergence in (1.10) holds true if conditions (A1)–(A9) as well as (B2) and (B4) are satisfied. By the following Lemma 1.3.1 (and Slutsky's theorem in the form of Corollary A.2 in [14]) we can immediately conclude that under the same assumptions

$$\sqrt{nb_n}\big(\widehat{F}_{p,n}(\cdot) - F_p(\cdot)\big) \rightsquigarrow^\circ B_p \qquad \text{in } (\boldsymbol{D}_{(\lambda)}, \mathcal{B}^\circ_{(\lambda)}, \|\cdot\|_{(\lambda)}), \tag{1.17}$$

because Lemma 1.3.1 ensures

$$\sqrt{nb_n}\|F_{p,n} - F_p\|_{(\lambda)} \to 0. \tag{1.18}$$

Here F_p refers to the distribution function of ξ_p introduced a few lines before (1.3). Lemma 1.3.1 involves the following condition.

(C2) $\lim_{n\to\infty} n^{(1-q)/(1+q)} b_n = 0$.

Note that (B2) implies (C2), and that (A2) implies (C2) if $q \geq 1$.

Lemma 1.3.1 *If (C2), (A3), (B4) hold true for some $\lambda \in [0, \infty)$, $q \in [\lambda, \infty) \cap (0, \infty)$, $C_{p,q} \in [0, \infty)$, then (1.18) holds.*

In the proof of Lemma 1.3.1 (see Subsection 1.5.8) we will show that (A3) and (B4) imply $\|F_{p,n} - F_p\|_{(\lambda)} = \mathcal{O}(n^{-q/(q+1)})$; together with (C2) this ensures the claim of the lemma. Let us summarize our findings.

Corollary 1.3.2 *Assume that (A1)–(A9) hold for some common $\lambda \in [0, \infty)$. Moreover assume that (B2) and (B4) hold for some $q \in [\lambda, \infty) \cap (0, \infty)$ with the same λ. Then (1.17) and (1.18) hold.*

1.3.2 Weighted empirical quantiles

The (lower) α-quantile functional associated with some given level $\alpha \in (0, 1)$ is defined by

$$\mathcal{Q}_\alpha(F) := \inf\big\{x \in \mathbb{R} : F(x) \geq \alpha\big\}$$

on the set of all distribution functions F on the real line. Given the time series $X_{n,1}, \ldots, X_{n,n}$, it can be reasonable to use $\mathcal{Q}_\alpha(\widehat{F}_{p,n})$ as an estimator for $\mathcal{Q}_\alpha(F_{p,n})$. Note that $\mathcal{Q}_\alpha(\widehat{F}_{p,n})$ can be seen as a weighted α-quantile. The estimator $\widehat{F}_{p,n}$ is indeed supported by the finite set $\{X_{n,1}, \ldots, X_{n,n}\}$, but the mass assigned to the individual points of this set is not uniform. More precisely, denoting by $X_{n,1(n)}, \ldots, X_{n,n(n)}$ the order statistics of $X_{n,1}, \ldots, X_{n,n}$, we have

$$\mathcal{Q}_\alpha(\widehat{F}_{p,n}) = X_{n,k(n)} \text{ for the smallest } k \in \{1, \ldots, n\} \text{ with } \sum_{i=1}^{k} w_n(i(n)) \geq \alpha,$$

where $w_n(i(n)) := c_n \kappa(\frac{i(n)-i_{p,n}}{nb_n})$ refers to the mass assigned to $X_{n,i(n)}$.

Given (1.17) and (1.18), we can use the functional delta-method to obtain

$$\sqrt{nb_n}\big(\mathcal{Q}_\alpha(\widehat{F}_{p,n}) - \mathcal{Q}_\alpha(F_{p,n})\big) \rightsquigarrow Z \tag{1.19}$$

for some centered normally distributed random variable Z with variance

$$\mathbb{V}\mathrm{ar}[Z] = \frac{\gamma_p(F_p^{-1}(\alpha), F_p^{-1}(\alpha))}{F_p'(F_p^{-1}(\alpha))^2} \tag{1.20}$$

under some assumption on F_p, where γ_p is the covariance function defined by (1.8).

Theorem 1.3.3 *Assume that (1.17) and (1.18) hold for $\lambda = 0$ and that F_p is continuously differentiable in a neighborhood of $F_p^{-1}(\alpha)$ with strictly positive derivative at $F_p^{-1}(\alpha)$. Then (1.19) holds.*

Proof In view of (1.17) and Remark 1.2.5, we obtain by Lemma 21.4 and Theorem 20.8 in [72] that $\sqrt{nb_n}(\mathcal{Q}_\alpha(\widehat{F}_{p,n}) - \mathcal{Q}_\alpha(F_p)) \rightsquigarrow Z$, noting that $Z := -B_p(F_p^{-1}(\alpha))/F_p'(F_p^{-1}(\alpha))$ is normally distributed with variance as in (1.20). Moreover, in view of (1.18), we obtain by another application of Lemma 21.4 and Theorem 20.8 in [72] (to purely deterministic variables) that $\sqrt{nb_n}(\mathcal{Q}_\alpha(F_{p,n}) - \mathcal{Q}_\alpha(F_p)) \to 0$. Along with Slutsky's theorem this gives (1.19). □

1.3.3 Weighted V-statistics

The V-functional (von Mises functional) of degree two associated with some given measurable function $h : \mathbb{R}^2 \to \mathbb{R}$ (often referred to as kernel) is defined by

$$\mathcal{V}_h(F) := \int_{\mathbb{R}}\int_{\mathbb{R}} h(x_1, x_2)\, \mu_F(dx_1)\mu_F(dx_2) \tag{1.21}$$

on the set $\mathbf{F}_h$ of all distribution functions F on the real line for which the double integral (with respect to the measure μ_F generated by F) in (1.21) exists. Given the time series $X_{n,1}, \ldots, X_{n,n}$, it can be reasonable to use $\mathcal{V}_h(\widehat{F}_{p,n})$ as an estimator for $\mathcal{V}_h(F_{p,n})$. Note that $\mathcal{V}_h(\widehat{F}_{p,n})$ can be seen as a weighted V-statistic. It indeed admits the representation

$$\mathcal{V}_h(\widehat{F}_{p,n}) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_n(i,j) h(X_{n,i}, X_{n,j})$$

with $w_n(i,j) := c_n^2 \kappa(\frac{i-i_{p,n}}{nb_n})\kappa(\frac{j-i_{p,n}}{nb_n})$.

Given (1.17) and (1.18), we can follow the continuous mapping approach of Beutner and Zähle [11] to show that under some assumptions (see Theorem 1.3.4 below)

$$\sqrt{nb_n}\big(\mathcal{V}_h(\widehat{F}_{p,n}) - \mathcal{V}_h(F_{p,n})\big) \rightsquigarrow Z \tag{1.22}$$

for some centered normally distributed random variable Z with variance

$$\mathbb{V}\mathrm{ar}[Z] = \int_{\mathbb{R}}\int_{\mathbb{R}} \gamma_p(x_1, x_2)\, \mu_{h_{F_p}}(dx_1)\mu_{h_{F_p}}(dx_2), \tag{1.23}$$

where γ_p is the covariance function defined by (1.8), and $h_{F_p} := h_{1,F_p} + h_{2,F_p}$ with $h_{1,F_p}(\cdot) := \int_{\mathbb{R}} h(\,\cdot\,, x_2)\, \mu_{F_p}(dx_2)$ and $h_{2,F_p}(\cdot) := \int_{\mathbb{R}} h(x_1, \,\cdot\,)\, \mu_{F_p}(dx_1)$.

Let us collect the assumptions we need for (1.22). Assume $F_p \in \mathbf{F}_h$, $F_{p,n} \in \mathbf{F}_h$ and $\int_{\mathbb{R}}\int_{\mathbb{R}} |h(x_1,x_2)|\mu_{F_{p,n}}(dx_1)\mu_{F_p}(dx_2) < \infty$ and $\int_{\mathbb{R}}\int_{\mathbb{R}} |h(x_1,x_2)|\mu_{F_p}(dx_1)\mu_{F_{p,n}}(dx_2) < \infty$ for any $n \in \mathbb{N}$. Assume that h_{1,F_p} and h_{2,F_p} are right-continuous and locally of bounded variation, that h is upper right-continuous and locally of bounded bivariation, and that $h_{x_1}(\cdot) := h(x_1, \cdot)$ and $h_{x_2}(\cdot) := h(\cdot, x_2)$ are locally of bounded variation for every fixed real x_1 and x_2, respectively. Under some weak additional assumptions (see e.g. Remark 1.3.5) making the tail behavior of h_{F_p} and F_p and of h and F_p compatible, one can derive from (1.21) the decomposition

$$\begin{aligned}\mathcal{V}_h(\widehat{F}_{p,n}) - \mathcal{V}_h(F_p) \;&=\; -\int_{\mathbb{R}} (\widehat{F}_{p,n} - F_p)(x-)\, \mu_{h_{F_p}}(dx) \\ &\quad + \int_{\mathbb{R}^2} (\widehat{F}_{p,n} - F_p)(x_1-)(\widehat{F}_{p,n} - F_p)(x_2-)\, \mu_h(d(x_1,x_2))\end{aligned} \tag{1.24}$$

and its analogue with $\widehat{F}_{p,n}$ replaced by $F_{p,n}$. Then, under (1.17) and (1.18), the continuous mapping theorem (in the form of Theorem 6.4 of [15]) and Slutsky's theorem (in the form of Corollary A.2 in [14]) imply the following theorem, where one should note that $Z := -\int_{\mathbb{R}} B_p(x)\mu_{h_{F_p}}(x)$ is normally distributed with variance as in (1.23).

Theorem 1.3.4 *Assume that (1.17) and (1.18) hold for some* $\lambda \in [0, \infty)$*. Moreover assume that (1.24) and its analogue with* $\widehat{F}_{p,n}$ *replaced by* $F_{p,n}$ *hold for any* $n \in \mathbb{N}$*, and that* $\int_{\mathbb{R}} \phi_{-\lambda}(x)\,|\mu_{h_{F_p}}|(dx) < \infty$ *and* $\int_{\mathbb{R}^2} \phi_{-\lambda}(x_1)\phi_{-\lambda}(x_2)\,|\mu_h|(d(x_1, x_2)) < \infty$*. Then (1.22) holds.*

Remark 1.3.5 The conditions in Lemmas 3.4 and 3.6 in [11] (with $\widehat{F}_n, F$ replaced by $\widehat{F}_{p,n}, F_p$) provide simple (but lengthy) conditions for (1.24). The analogous assumptions with $\widehat{F}_{p,n}$ replaced by $F_{p,n}$ ensure (1.24) with $\widehat{F}_{p,n}$ replaced by $F_{p,n}$. ◇

As elaborated in Section 3.2 of [11], the set of kernels h that satisfy the conditions mentioned in Remark 1.3.5 (and thus admit the representation (1.24)) include the kernels corresponding to the variance, to Gini's mean difference, to the Cramér–von Mises goodness-of-fit test statistic, and to the Arcones–Giné test statistic for symmetry.

In Corollary 1 and Example 2 in [83], Zhou presents the analogue of (1.22) with $\mathcal{V}_h(F_{p,n})$ replaced by $\mathcal{V}_h(\mathbb{E}[\widehat{F}_{p,n}])$. More precisely, he proves that the standardized V-statistic $(\mathcal{V}_h(\widehat{F}_{p,n}) - \mathbb{E}[\mathcal{V}_h(\widehat{F}_{p,n})])/\mathbb{V}\mathrm{ar}[\mathcal{V}_h(\widehat{F}_{p,n})]^{1/2}$ is asymptotically standard normal under similar assumptions.

1.4 Proof of Theorem 1.2.4

In the following we will only show that (1.9) holds true, because (1.10) is a trivial consequence of (1.9) and Slutsky's theorem (in the form of Corollary A.2 in [14]). For (1.9) it suffices to show

$$\phi_\lambda \widetilde{\mathcal{E}}_{p,n} \rightsquigarrow^\circ \phi_\lambda B_p \qquad \text{in } (\boldsymbol{D}_{(0)}, \mathcal{B}^\circ_{(0)}, \|\cdot\|_{(0)}) \tag{1.25}$$

(note that $\|\cdot\|_{(0)} = \|\cdot\|_\infty$). Indeed, the continuous mapping theorem (in the form of Theorem 6.4 of [15]) and the continuity of the mapping $v \mapsto v/\phi_\lambda$ from $(\boldsymbol{D}_{(0)}, \|\cdot\|_{(0)})$ to $(\boldsymbol{D}_{(\lambda)}, \|\cdot\|_{(\lambda)})$ together ensure that (1.25) implies (1.9).

To show (1.25), we derive in Subsection 1.4.1 a Donsker-type theorem (see Theorem 1.4.1 below). After a brief introduction to Burkholder's inequality in Subsection 1.4.2, we verify in Subsections 1.4.3 and 1.4.4 that conditions (a) and (b) of Theorem 1.4.1 below are satisfied in our setting, so that (1.25) is a direct consequence of Theorem 1.4.1.

1.4.1 Auxiliary result: Donsker-type theorem

The following Donsker-type theorem is a generalization of Theorem V.1.3 in [60].

Theorem 1.4.1 *Let ξ_n be a $(\boldsymbol{D}_{(0)}, \mathcal{B}^\circ_{(0)})$-valued random variable on some probability space $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ for every $n \in \mathbb{N}$. Let $\boldsymbol{C}_{(0)} \in \mathcal{B}^\circ_{(0)}$ be separable, and ξ be a $(\boldsymbol{D}_{(0)}, \mathcal{B}^\circ_{(0)})$-valued random variable on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\xi \in \boldsymbol{C}_{(0)}$ $\mathbb{P}$-a.s. Assume that the following two conditions hold.*

(a) *The finite-dimensional distributions of ξ_n converge in distribution to those of ξ.*

(b) *For every $\epsilon > 0$ and $\delta > 0$ there exist $k \in \mathbb{N}$ and a partition $-\infty = x_0 < x_1 < \cdots < x_k < x_{k+1} = \infty$ such that*

$$\limsup_{n\to\infty} \mathbb{P}_n\Big[\max_{i=0,\dots,k} \sup_{x\in[x_i,x_{i+1})} |\xi_n(x) - \xi_n(x_i)| \geq \delta\Big] \leq \epsilon.$$

Then $\xi_n \rightsquigarrow^\circ \xi$ in $(\boldsymbol{D}_{(0)}, \mathcal{B}^\circ_{(0)}, \|\cdot\|_{(0)})$.

In Subsections 1.4.3 and 1.4.4 we will verify conditions (a) and (b) of Theorem 1.4.1, if $\phi_\lambda \widetilde{\mathcal{E}}_{p,n}$ and $\phi_\lambda B_p$ play the roles of ξ_n and ξ respectively. We note that $\phi_\lambda \widetilde{\mathcal{E}}_{p,n}$ and $\phi_\lambda B_p$ take values in $\boldsymbol{D}_{(0)}$ and $\boldsymbol{C}_{(0)}$, respectively. This is ensured by Lemma 1.4.6 ahead and Lemma 1.2.3, respectively.

For the proof of Theorem 1.4.1, we first need two auxiliary results.

Lemma 1.4.2 *For every $v \in \boldsymbol{D}_{(0)}$ and $\varepsilon > 0$ there exist $m \in \mathbb{N}$ and a partition $-\infty = y_0 < y_1 < \cdots < y_m < y_{m+1} = \infty$ such that*

$$\max_{i=0,\dots,m} \sup_{x,x'\in[y_i,y_{i+1})} |v(x) - v(x')| \leq \varepsilon. \tag{1.26}$$

Proof Pick $\varepsilon > 0$. Let $\overline{y}$ be the supremum of those $y \in \mathbb{R}$ for which one can find $m \in \mathbb{N}$ and a partition $-\infty = y_0 < y_1 < \cdots < y_m < y_{m+1} = y$ such that (1.26) holds. Here we use the usual convention $\sup \mathbb{R} := \infty$. Since v as an element of $\boldsymbol{D}_{(0)}$ satisfies $\lim_{x\to-\infty} v(x) = 0$, we can find some $\underline{x} \in \mathbb{R}$ such that $|v(-\infty) - v(x)| \leq \varepsilon/2$ for all $x \leq \underline{x}$. Hence $\sup_{x,x'\leq\underline{x}} |v(x) - v(x')| \leq \varepsilon$. Thus $\overline{y} \geq \underline{x}$.

Next observe that one can find $m \in \mathbb{N}$ and a partition $-\infty = y_0 < y_1 < \cdots < y_m < y_{m+1} = \overline{y}$ such that (1.26) holds, i.e. one can find such a partition for $\overline{y}$ itself. Indeed: Since $c := \lim_{x\nearrow\overline{y}} v(x)$ exists in $\mathbb{R}$ (note that $c = 0$ if $\overline{y} = \infty$), we can find some $y^* \in (-\infty, \overline{y})$ such that $|c - v(x)| \leq \varepsilon/2$ for all $x \in [y^*, \overline{y})$, and thus $\sup_{x,x'\in[y^*,\overline{y})} |v(x) - v(x')| \leq \varepsilon$. By definition of $\overline{y}$ we can find $m \in \mathbb{N}$ and a partition $-\infty = y_0 < y_1 < \cdots < y_m = y^*$ such that $\sup_{x,x'\in[y_i,y_{i+1})} |v(x) - v(x')| \leq \varepsilon$ holds for $i = 0, \dots, m-1$. Hence (1.26) holds for $-\infty = y_0 < y_1 < \cdots < y_k < y_m = y^* < y_{m+1}$ with $y_{m+1} := \overline{y}$.

Finally suppose that $\overline{y} < \infty$. Then, since $\lim_{x\searrow\overline{y}} v(x) = v(\overline{y})$, one could find some $\delta > 0$ such that $|v(\overline{y}) - v(x)| \leq \varepsilon/2$ for all $x \in [\overline{y}, \overline{y} + \delta)$, and thus $\sup_{x,x'\in[\overline{y},\overline{y}+\delta)} |v(x) - v(x')| \leq \varepsilon$. This would lead to a contradiction to the definition of $\overline{y}$. Hence $\overline{y} = \infty$. $\square$

For any points $z_1, \ldots, z_\ell \in \mathbb{R}$, let the map $A_{\{z_1,\ldots,z_\ell\}} : \boldsymbol{D}_{(0)} \to \boldsymbol{D}_{(0)}$ be defined by

$$A_{\{z_1,\ldots,z_\ell\}}(v)(\cdot) := \sum_{i=1}^{\ell-1} v(z_{i:\ell}) \mathbb{1}_{[z_{i:\ell}, z_{i+1:\ell})}(\cdot),$$

where $z_{1:\ell}, \ldots, z_{\ell:\ell}$ is the order statistics of $z_1, \ldots, z_\ell$.

Lemma 1.4.3 *There exists a sequence $(z_p)_{p\in\mathbb{N}}$ of real numbers such that*

$$\lim_{p\to\infty} \|A_{\{z_1,\ldots,z_p\}}(v) - v\|_\infty = 0 \qquad \text{for all } v \in \boldsymbol{C}_{(0)}. \tag{1.27}$$

Proof *Step 1.* By the separability of $\boldsymbol{C}_{(0)}$ we can find a countable dense subset $\widetilde{\boldsymbol{C}}_{(0)} \subseteq \boldsymbol{C}_{(0)}$. Let $(\widetilde{v}_j)_{j\in\mathbb{N}}$ be an enumeration of $\widetilde{\boldsymbol{C}}_{(0)}$. By Lemma 1.4.2 we can find for every $j, \ell \in \mathbb{N}$ an $m_{j,\ell} \in \mathbb{N}$ and a partition $-\infty = y_0^{j,\ell} < y_1^{j,\ell} < \cdots < y_{m_{j,\ell}}^{j,\ell} < y_{m_{j,\ell}+1}^{j,\ell} = \infty$ such that

$$\max_{i=0,\ldots,m_{j,\ell}} \sup_{x,x' \in [y_i^{j,\ell}, y_{i+1}^{j,\ell})} |\widetilde{v}_j(x) - \widetilde{v}_j(x')| \le 1/\ell. \tag{1.28}$$

Set $U_\ell(\widetilde{v}_j) := \{y_1^{j,\ell}, \ldots, y_{m_{j,\ell}}^{j,\ell}\}$, $j, \ell \in \mathbb{N}$. Note that the left-hand side of (1.28) does not increase as the partition is getting finer. So we may assume without loss of generality that

$$U_1(\widetilde{v}_1) \subseteq U_1(\widetilde{v}_2) \subseteq U_2(\widetilde{v}_1) \subseteq U_1(\widetilde{v}_3) \subseteq U_2(\widetilde{v}_2) \subseteq U_3(\widetilde{v}_1) \subseteq \cdots \tag{1.29}$$

(here the order of the double indices are determined by Cantor's diagonal method), and for any $j, \ell \in \mathbb{N}$ and any $U_{\ell_*}(\widetilde{v}_{j_*})$ which occurs in between $U_\ell(\widetilde{v}_j)$ and $U_{\ell+1}(\widetilde{v}_j)$ in (1.29) we have

$$\max_{i=0,\ldots,m_{j_*,\ell_*}} \sup_{x,x' \in [y_i^{j_*,\ell_*}, y_{i+1}^{j_*,\ell_*})} |\widetilde{v}_j(x) - \widetilde{v}_j(x')| \le 1/\ell$$

and thus

$$\|A_{U_{\ell_*}(\widetilde{v}_{j_*})}(\widetilde{v}_j) - \widetilde{v}_j\|_\infty \le \max_{i=0,\ldots,m_{j_*,\ell_*}} \sup_{x \in [y_i^{j_*,\ell_*}, y_{i+1}^{j_*,\ell_*})} |\widetilde{v}_j(x) - \widetilde{v}_j(y_i^{j_*,\ell_*})| \le 1/\ell. \tag{1.30}$$

Now choose the sequence $(z_p)_{p\in\mathbb{N}}$ as follows. The first $\#U_1(\widetilde{v}_1)$ terms are the elements of $U_1(\widetilde{v}_1)$, the next $\#(U_1(\widetilde{v}_2) \setminus U_1(\widetilde{v}_1))$ terms are the elements of $U_1(\widetilde{v}_2) \setminus U_1(\widetilde{v}_1)$, the next $\#(U_2(\widetilde{v}_1) \setminus U_1(\widetilde{v}_2))$ terms are the elements of $U_2(\widetilde{v}_1) \setminus U_1(\widetilde{v}_2)$, the next $\#(U_1(\widetilde{v}_3) \setminus U_2(\widetilde{v}_1))$ terms are the elements of $U_1(\widetilde{v}_3) \setminus U_2(\widetilde{v}_1)$, and so on. Then (1.30) implies

$$\lim_{p\to\infty} \|A_{\{z_1,\ldots,z_p\}}(\widetilde{v}) - \widetilde{v}\|_\infty = 0 \qquad \text{for all } \widetilde{v} \in \widetilde{\boldsymbol{C}}_{(0)}. \tag{1.31}$$

Step 2. It remains to show that (1.31) extends to (1.27). Let $v \in \boldsymbol{C}_{(0)}$ and $\varepsilon > 0$ arbitrary but fixed. Choose $\widetilde{v} \in \widetilde{\boldsymbol{C}}_{(0)}$ such that $\|v - \widetilde{v}\|_\infty \le \varepsilon/3$; recall that $\widetilde{\boldsymbol{C}}_{(0)}$ is a dense subset of $\boldsymbol{C}_{(0)}$. Moreover, by (1.31) we can choose $p_{\widetilde{v}} \in \mathbb{N}$ such that $\|A_{\{z_1,\dots,z_p\}}(\widetilde{v}) - \widetilde{v}\|_\infty \le \varepsilon/3$ for all $p \ge p_{\widetilde{v}}$. It follows that

$$\begin{aligned} &\|A_{\{z_1,\dots,z_p\}}(v) - v\|_\infty \\ &\quad\le \|A_{\{z_1,\dots,z_p\}}(v) - A_{\{z_1,\dots,z_p\}}(\widetilde{v})\|_\infty + \|A_{\{z_1,\dots,z_p\}}(\widetilde{v}) - \widetilde{v}\|_\infty + \|\widetilde{v} - v\|_\infty \\ &\quad\le \|v - \widetilde{v}\|_\infty + \|A_{\{z_1,\dots,z_p\}}(\widetilde{v}) - \widetilde{v}\|_\infty + \|\widetilde{v} - v\|_\infty \\ &\quad\le \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \\ &\quad= \varepsilon \end{aligned}$$

for all $p \ge p_{\widetilde{v}}$. This gives (1.27). □

Proof of Theorem 1.4.1 According to the Portmanteau theorem in the form of Theorem 6.3 in [15] it suffices to show

$$\lim_{n\to\infty} \int f\, d\mathbb{P}^n_{\xi_n} = \int f\, d\mathbb{P}_\xi \tag{1.32}$$

for any bounded, uniformly continuous and $(\mathcal{B}^\circ_{(0)}, \mathcal{B}(\mathbb{R}))$-measurable function $f : \boldsymbol{D}_{(0)} \to \mathbb{R}$. Let f be any such function. Pick $\varepsilon > 0$, and choose $\delta > 0$ such that

$$|f(v) - f(w)| \le \varepsilon/4 \qquad \text{for any } v, w \in \boldsymbol{D}_{(0)} \text{ with } \|v - w\|_\infty \le \delta. \tag{1.33}$$

Step 1. By assumption (b) we can find a grid partition $-\infty = x_0 < x_1 < \cdots < x_k < x_{k+1} = \infty$ (depending on ε and δ) such that

$$\begin{aligned} &\limsup_{n\to\infty} \mathbb{P}_n\big[\|A_{\{x_1,\dots,x_k\}}(\xi_n) - \xi_n\|_\infty \ge \delta/2\big] \\ &\quad\Big(= \limsup_{n\to\infty} \mathbb{P}_n\Big[\max_{i=0,\dots,k} \sup_{x\in[x_i,x_{i+1})} |\xi_n(x) - \xi_n(x_i)| \ge \delta/2\Big]\Big) \\ &\quad\le \varepsilon/(4\|f\|_\infty). \end{aligned} \tag{1.34}$$

Moreover, by Lemma 1.4.3 we can choose a sequence $(z_p)_{p\in\mathbb{N}}$ of real numbers such that (1.27) holds. Since we assumed $\xi \in \boldsymbol{C}_{(0)}$ $\mathbb{P}$-a.s., we can conclude that

$$\lim_{p\to\infty} \|A_{\{z_1,\dots,z_p\}}(\xi) - \xi\|_\infty = 0 \qquad \mathbb{P}\text{-a.s.} \tag{1.35}$$

The map $A_{\{z_1,\dots,z_p\}} : \boldsymbol{D}_{(0)} \to \boldsymbol{D}_{(0)}$ is $(\mathcal{B}^\circ_{(0)}, \mathcal{B}^\circ_{(0)})$-measurable for any $p \in \mathbb{N}$, because

$$\begin{aligned} &A^{-1}_{\{z_1,\dots,z_p\}}(B_r(v)) \\ &\quad= \big\{w \in \boldsymbol{D}_{(0)} : \|A_{\{z_1,\dots,z_p\}}(w) - v\|_\infty < r\big\} \\ &\quad= \Big\{w \in \boldsymbol{D}_{(0)} : \max_{i=1,\dots,p-1} \sup_{x\in[z_i,z_{i+1})} |w(z_i) - v(x)| < r, \sup_{x\in(-\infty,z_1)\cup[z_p,\infty)} |v(x)| < r\Big\} \\ &\quad= \Big\{w \in \boldsymbol{D}_{(0)} : \max_{i=1,\dots,p-1} \sup_{x\in[z_i,z_{i+1})\cap\mathbb{Q}} |w(z_i) - v(x)| < r, \sup_{x\in((-\infty,z_1)\cup[z_p,\infty))\cap\mathbb{Q}} |v(x)| < r\Big\} \end{aligned}$$

$$= \begin{cases} \bigcap_{i=1}^{p-1} \bigcap_{x\in[z_i,z_{i+1})\cap\mathbb{Q}} \pi_{z_i}^{-1}(B_r(v(x))) & , \ \sup_{x\in(-\infty,z_1)\cup[z_p,\infty)} |v(x)| < r \\ \emptyset & , \ \text{otherwise} \end{cases}$$

lies in $\mathcal{B}_{(0)}^{\circ}$ for any $p \in \mathbb{N}$, $v \in \boldsymbol{D}_{(0)}$ and $r > 0$; take into account that the projection map $\pi_z : \boldsymbol{D}_{(0)} \to \mathbb{R}$ is $(\mathcal{B}_{(0)}^{\circ}, \mathcal{B}(\mathbb{R}))$-measurable for any $z \in \mathbb{R}$. It follows that $A_{\{z_1,\dots,z_p\}}(\xi)$ is $(\mathcal{F}, \mathcal{D}_0)$-measurable for any $p \in \mathbb{N}$, and thus $\|A_{\{z_1,\dots,z_p\}}(\xi) - \xi\|_\infty = \sup_{x\in\mathbb{Q}} |A_{\{z_1,\dots,z_p\}}(\xi)(x) - \xi(x)|$ is $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$-measurable for any $p \in \mathbb{N}$. Therefore (1.35) implies

$$\lim_{p\to\infty} \mathbb{P}\big[\|A_{\{z_1,\dots,z_p\}}(\xi) - \xi\|_\infty \geq \eta\big] = 0 \qquad \text{for all } \eta > 0. \tag{1.36}$$

In view of (1.36), we can choose $p_* \in \mathbb{N}$ (depending on ε and δ) such that

$$\mathbb{P}\big[\|A_{\{z_1,\dots,z_{p_*}\}}(\xi) - \xi\|_\infty \geq \delta/2\big] \leq \varepsilon/(4\|f\|_\infty). \tag{1.37}$$

Now if we set $U := \{x_1, \dots, x_k, z_1, \dots, z_{p_*}\}$ we obtain the following analogues of (1.34) and (1.37), with $\delta/2$ replaced by δ:

$$\limsup_{n\to\infty} \mathbb{P}_n\big[\|A_U(\xi_n) - \xi_n\|_\infty \geq \delta\big] \ \leq \ \varepsilon/(4\|f\|_\infty), \tag{1.38}$$

$$\mathbb{P}\big[\|A_U(\xi) - \xi\|_\infty \geq \delta\big] \ \leq \ \varepsilon/(4\|f\|_\infty). \tag{1.39}$$

Step 2. To verify (1.32) we apply the triangle inequality to obtain

$$\begin{aligned} \Big|\int f\, d\mathbb{P}_{\xi_n}^n - \int f\, d\mathbb{P}_\xi\Big| \ &\leq \ \int |f(\xi_n) - f(A_U(\xi_n))|\, d\mathbb{P}^n \\ &\quad + \Big|\int f(A_U(\xi_n))\, d\mathbb{P}^n - \int f(A_U(\xi))\, d\mathbb{P}\Big| \\ &\quad + \int |f(A_U(\xi)) - f(\xi)|\, d\mathbb{P} \\ &=: \ S_1(n) + S_2(n) + S_3. \end{aligned} \tag{1.40}$$

For the first summand we obtain by (1.33) and (1.38)

$$\begin{aligned} \limsup_{n\to\infty} S_1(n) \ &\leq \ \limsup_{n\to\infty} \int |f(\xi_n) - f(A_U(\xi_n))| \mathbb{1}_{\{\|\xi_n - A_U(\xi_n)\|_\infty \leq \delta\}}\, d\mathbb{P}^n \\ &\quad + \limsup_{n\to\infty} \int |f(\xi_n) - f(A_U(\xi_n))| \mathbb{1}_{\{\|\xi_n - A_U(\xi_n)\|_\infty \geq \delta\}}\, d\mathbb{P}^n \\ &\leq \ \varepsilon/4 + \|f\|_\infty \limsup_{n\to\infty} \mathbb{P}^n\big[\|\xi_n - A_U(\xi_n)\|_\infty \geq \delta\big] \\ &\leq \ \varepsilon/2. \end{aligned} \tag{1.41}$$

For the third summand we analogously obtain by (1.33) and (1.39)

$$S_3 \ \leq \ \int |f(\xi) - f(A_U(\xi))| \mathbb{1}_{\{\|\xi - A_U(\xi)\|_\infty \leq \delta\}}\, d\mathbb{P}$$

$$
\begin{aligned}
&\quad + \int |f(\xi) - f(A_U(\xi))| \mathbb{1}_{\{\|\xi - A_U(\xi)\|_\infty \geq \delta\}} \, d\mathbb{P} \\
&\leq \varepsilon/4 + \|f\|_\infty \mathbb{P}\big[\|\xi - A_U(\xi)\|_\infty \geq \delta\big] \\
&\leq \varepsilon/2. \qquad (1.42)
\end{aligned}
$$

Let $m := k + p_*$, and $y_1, \dots, y_m$ be an enumeration of $U := \{x_1, \dots, x_k, z_1, \dots, z_{p_*}\}$ such that $y_1 \leq \cdots \leq y_m$. For any $(a_1, \dots, a_m) \in \mathbb{R}^m$, let $\varphi_{a_1,\dots,a_m} : \mathbb{R} \to \mathbb{R}$ be defined by $\varphi_{a_1,\dots,a_m}(x) := \sum_{i=0}^{m-1} a_i \mathbb{1}_{[y_i, y_{i+1})}(x)$, with the conventions $a_0 := 0$, $y_0 := -\infty$ and $y_{m+1} := \infty$. Then the function $g : \mathbb{R}^m \to \mathbb{R}$ defined by $g(a_1, \dots, a_m) := f(\varphi_{a_1,\dots,a_m})$ is bounded and continuous, where the continuity follows from the continuity of f and the continuity of the mapping $(a_1, \dots, a_m) \mapsto \varphi_{a_1,\dots,a_m}(\cdot)$ from $\mathbb{R}^m$ to $\boldsymbol{D}_{(0)}$. Since $f(A_U) = g(\pi_{y_1,\dots,y_m})$, assumption (a) ensures

$$
\begin{aligned}
\limsup_{n\to\infty} S_2(n) &= \limsup_{n\to\infty} \Big| \int f(A_U(\xi_n)) \, d\mathbb{P}^n - \int f(A_U(\xi)) \, d\mathbb{P} \Big| \\
&= \limsup_{n\to\infty} \Big| \int g(\pi_{y_1,\dots,y_m}(\xi_n)) \, d\mathbb{P}^n - \int g(\pi_{y_1,\dots,y_m}(\xi)) \, d\mathbb{P} \Big| \\
&= \limsup_{n\to\infty} \Big| \int g(\xi_n(y_1), \dots, \xi_n(y_m)) \, d\mathbb{P}^n - \int g(\xi(y_1), \dots, \xi(y_m)) \, d\mathbb{P} \Big| \\
&= 0. \qquad (1.43)
\end{aligned}
$$

By (1.40)–(1.43) we have $\limsup_{n\to\infty} | \int f \, d\mathbb{P}^n_{\xi_n} - \int f \, d\mathbb{P}_\xi | \leq \varepsilon$. Since $\varepsilon > 0$ was chosen arbitrarily, we arrive at (1.32). □

1.4.2 Auxiliary result: Burkholder's inequality

Many approaches in dealing with asymptotic issues of (piecewise locally) stationary processes are based on martingale techniques. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$. Recall that a real-valued L^1-process $(M_n)_{n\in\mathbb{N}}$ with M_n being $\mathcal{F}_n$-measurable for all $n \in \mathbb{N}$ is called a *martingale*, if $\mathbb{E}[M_{n+1}|\mathcal{F}_n] = M_n$ $\mathbb{P}$-a.s. for all $n \in \mathbb{N}$. To verify assumption (a) and (b) of Theorem 1.4.1, we will frequently use the following Burkholder inequality (cf. Theorem 11.2.1 in [18]).

Lemma 1.4.4 (Burkholder inequality) *Let $q \in (1, \infty)$. There exist constants $c_q := (q-1)/(18q^{3/2})$ and $C_q := 18q^{3/2}/(q-1)^{1/2}$ such that for any martingale $(M_n)_{n\in\mathbb{N}}$*

$$
c_q \Big\| \Big(\sum_{i=1}^n \big(M_i - M_{i-1}\big)^2 \Big)^{1/2} \Big\|_q \leq \|M_n\|_q \leq C_q \Big\| \Big(\sum_{i=1}^n \big(M_i - M_{i-1}\big)^2 \Big)^{1/2} \Big\|_q .
$$

In our applications, we only need the upper bound of $\|M_n\|_q$. We note that the differences $M_n - M_{n-1}$ are called *martingale differences*. By definition a *martingale difference sequence* is a real-valued L^1-process $(D_n)_{n\in\mathbb{N}}$ with D_n being $\mathcal{F}_n$-measurable for all $n \in \mathbb{N}$ that fulfills $\mathbb{E}[D_{n+1}|\mathcal{F}_n] = 0$ $\mathbb{P}$-a.s. for all $n \in \mathbb{N}$. And indeed, for any martingale $(M_n)_{n\in\mathbb{N}}$ the difference $M_n - M_{n-1}$ is $\mathcal{F}_n$-measurable and $\mathbb{E}[|M_n - M_{n-1}|] \leq \mathbb{E}[|M_n|] + \mathbb{E}[|M_{n-1}|] < \infty$ for all $n \in \mathbb{N}$. Moreover, $\mathbb{E}[M_n - M_{n-1}|\mathcal{F}_{n-1}] = \mathbb{E}[M_n|\mathcal{F}_{n-1}] - M_{n-1} = M_{n-1} - M_{n-1} = 0$ $\mathbb{P}$-a.s. for all $n \in \mathbb{N}$ so that $(M_n - M_{n-1})_{n\in\mathbb{N}}$ fulfills all the properties of a martingale difference sequence.

Conversely, given a martingale difference sequence $(D_n)_{n\in\mathbb{N}}$, then for all $n \in \mathbb{N}$ we obviously have that $\sum_{i=1}^n D_i$ is $\mathcal{F}_n$-measurable with $\mathbb{E}[|\sum_{i=1}^n D_i|] < \infty$. Since additionally $\mathbb{E}[\sum_{i=1}^n D_i|\mathcal{F}_{n-1}] = \sum_{i=1}^{n-1} D_i + \mathbb{E}[D_n|\mathcal{F}_{n-1}] = \sum_{i=1}^{n-1} D_i$ $\mathbb{P}$-a.s. for all $n \in \mathbb{N}$, the process $(\sum_{i=1}^n D_i)_{n\in\mathbb{N}}$ is a martingale for any martingale difference sequence $(D_n)_{n\in\mathbb{N}}$.

We thus arrive at the following corollary of Lemma 1.4.4. The third assertion can also be found as Lemma 3 in [78].

Corollary 1.4.5 *Let $q \in (1,\infty)$. There exists a constant $C_q := 18q^{3/2}/(q-1)^{1/2}$ such that for every martingale difference sequence $(D_n)_{n\in\mathbb{N}}$*

(i) $\big\| \sum_{i=1}^n D_i \big\|_q \leq C_q \big\| \big(\sum_{i=1}^n D_i^2 \big)^{1/2} \big\|_q$,

(ii) $\big\| \sum_{i=1}^n D_i \big\|_q^2 \leq C_q^2 \, \big\| \sum_{i=1}^n D_i^2 \big\|_{q/2}$,

(iii) $\big\| \sum_{i=1}^n D_i \big\|_q^{\min\{q,2\}} \leq C_q^{\min\{q,2\}} \, \sum_{i=1}^n \big\| D_i \big\|_q^{\min\{q,2\}}$.

Proof (i): As previously mentioned, the process $(\sum_{i=1}^n D_i)_{n\in\mathbb{N}}$ is a martingale. Assertion (i) is thus a direct consequence of Lemma 1.4.4 applied to $(\sum_{i=1}^n D_i)_{n\in\mathbb{N}}$.

(ii): Since

$$\Big\| \Big(\sum_{i=1}^n D_i^2 \Big)^{1/2} \Big\|_q = \mathbb{E}\Big[\Big| \sum_{i=1}^n D_i^2 \Big|^{q/2} \Big]^{1/q} = \Big(\Big\| \sum_{i=1}^n D_i^2 \Big\|_{q/2} \Big)^{1/2},$$

assertion (ii) follows immediately from (i).

(iii): If $q > 2$, then we obtain by (ii) and Minkowski's inequality

$$\Big\| \sum_{i=1}^n D_i \Big\|_q^2 \leq C_q^2 \Big\| \sum_{i=1}^n D_i^2 \Big\|_{q/2} \leq C_q^2 \sum_{i=1}^n \| D_i^2 \|_{q/2} = C_q^2 \sum_{i=1}^n \big\| D_i \big\|_q^2. \tag{1.44}$$

If $q \leq 2$, then (i) yields

$$\begin{aligned} \Big\| \sum_{i=1}^n D_i \Big\|_q^q &\leq C_q^q \Big\| \Big(\sum_{i=1}^n D_i^2 \Big)^{1/2} \Big\|_q^q = C_q^q \, \mathbb{E}\Big[\Big(\sum_{i=1}^n D_i^2 \Big)^{q/2} \Big] \\ &\leq C_q^q \, \mathbb{E}\Big[\sum_{i=1}^n D_i^q \Big] = C_q^q \sum_{i=1}^n \| D_i \|_q^q. \end{aligned}$$

Along with (1.44) this implies (ii). □

1.4.3 Verification of condition (a) of Theorem 1.4.1

Let $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ be arbitrary but fixed, and assume that $x_1 < \cdots < x_d$. Here we show that (under assumptions (A1)–(A5) and (A9)) we have

$$\big(\phi_\lambda(x_1)\widetilde{\mathcal{E}}_{p,n}(x_1), \ldots, \phi_\lambda(x_d)\widetilde{\mathcal{E}}_{p,n}(x_d)\big)' \rightsquigarrow \big(\phi_\lambda(x_1)B_p(x_1), \ldots, \phi_\lambda(x_d)B_p(x_d)\big)',$$

where for any vector $\boldsymbol{v} \in \mathbb{R}^d$ we denote by $\boldsymbol{v}'$ the transpose of $\boldsymbol{v}$. By the continuous mapping theorem (in the form of Theorem 6.4 of [15]) it suffices to show that

$$\big(\widetilde{\mathcal{E}}_{p,n}(x_1), \ldots, \widetilde{\mathcal{E}}_{p,n}(x_d)\big)' \rightsquigarrow \big(B_p(x_1), \ldots, B_p(x_d)\big)'.$$

Due to the Cramér–Wold theorem it even suffices to show that

$$\sum_{k=1}^{d} \lambda_k \widetilde{\mathcal{E}}_{p,n}(x_k) \rightsquigarrow \sum_{k=1}^{d} \lambda_k B_p(x_k) \tag{1.45}$$

for every $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_d) \in \mathbb{R}^d$. For the proof of (1.45), we borrow arguments from the proof of Theorem 1 in [83]. Setting

$$\begin{aligned} Y_{n,i}(\boldsymbol{x}, \boldsymbol{\lambda}) &:= \sum_{k=1}^{d} \lambda_k\, \kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big) \Big(\mathbb{1}_{[X_{n,i},\infty)}(x_k) - \mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x_k)\big]\Big), \\ Y_{n,i}^{\{m_n\}}(\boldsymbol{x}, \boldsymbol{\lambda}) &:= \mathbb{E}[Y_{n,i}(\boldsymbol{x}, \boldsymbol{\lambda}) | \boldsymbol{\epsilon}_{i:i-m_n}] \end{aligned}$$

with $\boldsymbol{\epsilon}_{i:i-m_n} := (\varepsilon_i, \varepsilon_{i-1}, \ldots, \varepsilon_{i-m_n+1})$ and $m_n := \lceil \log(n) \rceil$, the left-hand side of (1.45) can be written as

$$\begin{aligned} &\sum_{k=1}^{d} \lambda_k \widetilde{\mathcal{E}}_{p,n}(x_k) \\ &= c_n \sqrt{nb_n} \sum_{i=1}^{n} Y_{n,i}(\boldsymbol{x}, \boldsymbol{\lambda}) \\ &= c_n \sqrt{nb_n} \Big(\sum_{i=1}^{n} Y_{n,i}(\boldsymbol{x}, \boldsymbol{\lambda}) - \sum_{i=1}^{n} Y_{n,i}^{\{m_n\}}(\boldsymbol{x}, \boldsymbol{\lambda}) \Big) + c_n \sqrt{nb_n} \sum_{i=1}^{n} Y_{n,i}^{\{m_n\}}(\boldsymbol{x}, \boldsymbol{\lambda}) \\ &=: S_{n,1}(\boldsymbol{x}, \boldsymbol{\lambda}) + S_{n,2}(\boldsymbol{x}, \boldsymbol{\lambda}). \end{aligned} \tag{1.46}$$

The summand $S_{n,1}(\boldsymbol{x}, \boldsymbol{\lambda})$ converges in probability to 0 by Lemma 1.4.9 ahead and (1.12). We will now prove that the summand $S_{n,2}(\boldsymbol{x}, \boldsymbol{\lambda})$ converges in distribution to the right-hand side in (1.45), which is a centered normally distributed random variable with variance

$$\mathbb{V}\mathrm{ar}\Big[\sum_{k=1}^{d} \lambda_k B_p(x_k)\Big] = \sum_{k=1}^{d} \sum_{l=1}^{d} \lambda_k \lambda_l\, \mathbb{E}\big[B_p(x_k) B_p(x_l)\big] = \sum_{k=1}^{d} \sum_{l=1}^{d} \lambda_k \lambda_l\, \gamma_p(x_k, x_l). \tag{1.47}$$

Along with Slutsky's theorem, this gives (1.45).

If the expression in (1.47) vanishes, then $\sum_{k=1}^d \lambda_k B_p(x_k) = 0$ $\mathbb{P}$-almost surely and $\lim_{n\to\infty} \mathbb{V}\mathrm{ar}[\sum_{k=1}^d \lambda_k \widetilde{\mathcal{E}}_{p,n}(x_k)] = 0$ by Lemma 1.2.2. The latter convergence implies $\lim_{n\to\infty} \| \sum_{k=1}^d \lambda_k \widetilde{\mathcal{E}}_{p,n}(x_k)\|_2 = 0$, i.e. $\lim_{n\to\infty} \| \sum_{k=1}^d \lambda_k \widetilde{\mathcal{E}}_{p,n}(x_k) - \sum_{k=1}^d \lambda_k B_p(x_k)\|_2 = 0$. Thus $\sum_{k=1}^d \lambda_k \widetilde{\mathcal{E}}_{p,n}(x_k)$ converges in distribution to $\sum_{k=1}^d \lambda_k B_p(x_k)$, i.e. (1.45) holds.

Now assume that the expression in (1.47) is strictly greater than 0. Then it suffices to show that

$$\frac{S_{n,2}(\boldsymbol{x},\boldsymbol{\lambda})}{\sqrt{\mathbb{V}\mathrm{ar}[\sum_{k=1}^d \lambda_k B_p(x_k)]}} \rightsquigarrow Z$$

for a standard normally distributed random variable Z. By Slutsky's theorem and Lemma 1.4.12(iv) ahead this is equivalent to

$$\frac{S_{n,2}(\boldsymbol{x},\boldsymbol{\lambda})}{\sqrt{\mathbb{V}\mathrm{ar}[S_{n,2}(\boldsymbol{x},\boldsymbol{\lambda})]}} \rightsquigarrow Z. \tag{1.48}$$

To verify (1.48), we split $S_{n,2}(\boldsymbol{x},\boldsymbol{\lambda})$ into sums

$$S_{n,2}(\boldsymbol{x},\boldsymbol{\lambda}) = c_n\sqrt{nb_n}\sum_{j=1}^{\lceil n/s_n\rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) + c_n\sqrt{nb_n}\sum_{j=1}^{\lceil n/s_n\rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \tag{1.49}$$

of $\lceil n/s_n\rceil$ many big blocks

$$R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) := \sum_{i=1}^{l_n} Y_{n,(j-1)s_n+i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda}), \qquad j = 1,2,\ldots,\lceil n/s_n\rceil, \tag{1.50}$$

and $\lceil n/s_n\rceil$ many small blocks

$$r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) := \sum_{i=l_n+1}^{s_n} Y_{n,(j-1)s_n+i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda}), \qquad j = 1,2,\ldots,\lceil n/s_n\rceil, \tag{1.51}$$

where $l_n := \lceil\sqrt{nb_n}\rceil$ and $s_n := l_n + \lceil(\log n)^2\rceil$. Recall $m_n = \lceil\log(n)\rceil$, and note that the big blocks are independent since $s_n - l_n > m_n - 2$, and that the small blocks are independent since $m_n < l_n + 2$.

Now, $c_n\sqrt{nb_n}\sum_{j=1}^{\lceil n/s_n\rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})$ converges in probability to 0 by (1.76) and (1.12). Moreover, $\lim_{n\to\infty} \mathbb{V}\mathrm{ar}[S_{n,2}(\boldsymbol{x},\boldsymbol{\lambda})]/\mathbb{V}\mathrm{ar}[c_n\sqrt{nb_n}\sum_{j=1}^{\lceil n/s_n\rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})] = 1$ by part (ii) of Lemma 1.4.12 and (1.12). Thus, in view of (1.49) and Slutsky's theorem, for (1.48) it suffices to show

$$\frac{c_n\sqrt{nb_n}\sum_{j=1}^{\lceil n/s_n\rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})}{\big(\mathbb{V}\mathrm{ar}[c_n\sqrt{nb_n}\sum_{j=1}^{\lceil n/s_n\rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})]\big)^{1/2}} = \frac{\sum_{j=1}^{\lceil n/s_n\rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})}{\big(\mathbb{V}\mathrm{ar}[\sum_{j=1}^{\lceil n/s_n\rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})]\big)^{1/2}} \rightsquigarrow Z. \tag{1.52}$$

The big blocks, i.e. the random variables in (1.50), are independent and centered. Thus, in view of Lyapunov's central limit theorem, for (1.52) it suffices to verify that the Lyapunov condition holds for $\sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})$. For $q \in (2,\infty)$ as in (A5) and sufficiently large n we have

$$\begin{aligned}\frac{\sum_{j=1}^{\lceil n/s_n \rceil} \|R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\|_q^q}{\big(\mathbb{V}\mathrm{ar}[\sum_j R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})]\big)^{q/2}} &\le \frac{1}{c_{\boldsymbol{\lambda}}^{q/2}} \frac{\sum_{j=1}^{\lceil n/s_n \rceil} \|R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\|_q^q}{(nb_n)^{q/2}} \\ &\le \frac{1}{c_{\boldsymbol{\lambda}}^{q/2}} C_{\boldsymbol{\lambda},q}^{q/2} \Big(\frac{2nb_n - \sqrt{nb_n}}{\sqrt{nb_n} + \log^2(n)}\Big) \frac{\big(\|\kappa\|_\infty^2 \sqrt{nb_n} + \mathcal{O}(1)\big)^{q/2}}{(nb_n)^{q/2}},\end{aligned}$$

where we used Lemma 1.4.13 for the first step and Lemma 1.4.10(ii) for the second step. The latter bound converges to 0 by (A2) (and $q > 2$). This shows that the Lyapunov condition indeed holds. □

1.4.4 Verification of condition (b) of Theorem 1.4.1

In this subsection we will show that (under assumptions (A1)–(A3) and (A6)–(A8)) there exist for every $\epsilon > 0$ and $\delta > 0$ some $k \in \mathbb{N}$ and a partition $-\infty = x_0 < x_1 < \cdots < x_k < x_{k+1} = \infty$ such that

$$\limsup_{n\to\infty} \mathbb{P}\Big[\max_{i=0,\dots,k} \sup_{x\in[x_i,x_{i+1})} \big|\widetilde{\mathcal{E}}_{p,n}(x)\phi_\lambda(x) - \widetilde{\mathcal{E}}_{p,n}(x_i)\phi_\lambda(x_i)\big| \ge 2\delta\Big] \le 2\epsilon.$$

For the proof, we use the same idea as in [78]. Since we can write

$$\widetilde{\mathcal{E}}_{p,n}(x) = c_n\sqrt{nb_n}\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\big(\mathbb{1}_{[X_{n,i},\infty)}(x) - \mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x)\big]\big)$$

as $\widetilde{\mathcal{E}}_{p,n}(x) = H_{p,n}(x) + Q_{p,n}(x)$ with

$$\begin{aligned} H_{p,n}(x) &:= c_n\sqrt{nb_n}\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\big(\mathbb{1}_{[X_{n,i},\infty)}(x) - \mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}]\big) \\ &= c_n\sqrt{nb_n}\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\big(\mathbb{1}_{[X_{n,i},\infty)}(x) - \mathfrak{F}_{n,i}(x,\boldsymbol{\epsilon}_{i-1})\big), \\ Q_{p,n}(x) &:= c_n\sqrt{nb_n}\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\big(\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}] - \mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x)\big]\big) \\ &= c_n\sqrt{nb_n}\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\big(\mathfrak{F}_{n,i}(x,\boldsymbol{\epsilon}_{i-1}) - \mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x)\big]\big), \end{aligned}$$

it suffices to prove that for every $\epsilon > 0$ and $\delta > 0$ there exist $k_1, k_2 \in \mathbb{N}$ and partitions $-\infty = x_0 < x_1 < \cdots < x_{k_1} < x_{k_1+1} = \infty$ and $-\infty = x_0 < x_1 < \cdots < x_{k_2} < x_{k_2+1} = \infty$

with

$$\limsup_{n\to\infty} \mathbb{P}\Big[\max_{i=0,\dots,k_1} \sup_{x\in[x_i,x_{i+1})} \big|H_{p,n}(x)\phi_\lambda(x) - H_{p,n}(x_i)\phi_\lambda(x_i)\big| \ge \delta\Big] \le \epsilon \tag{1.53}$$

and

$$\limsup_{n\to\infty} \mathbb{P}\Big[\max_{i=0,\dots,k_2} \sup_{x\in[x_i,x_{i+1})} \big|Q_{p,n}(x)\phi_\lambda(x) - Q_{p,n}(x_i)\phi_\lambda(x_i)\big| \ge \delta\Big] \le \epsilon. \tag{1.54}$$

Let $\epsilon > 0$ and $\delta > 0$ be arbitrary but fixed. By Lemma 1.4.17(ii) and Lemma 1.4.19(ii), we can find $w_{1,\epsilon} \ge 0$ and $w_{2,\epsilon} \ge 0$ such that

$$\sup_{n\in\mathbb{N}} \mathbb{P}\Big[\sup_{|x|\ge w_{1,\epsilon}} |H_{p,n}(x)|\phi_\lambda(x) \ge \delta\Big] \le \epsilon, \quad \sup_{n\in\mathbb{N}} \mathbb{P}\Big[\sup_{|x|\ge w_{2,\epsilon}} |Q_{p,n}(x)|\phi_\lambda(x) \ge \delta\Big] \le \epsilon.$$

Then (1.53) and (1.54) follow directly from Lemma 1.4.17(iii) and Lemma 1.4.19(iii). □

1.4.5 Technical details

We first show that $\mathcal{E}_{p,n}$ and $\widetilde{\mathcal{E}}_{p,n}$ can be seen as random variables in $(\boldsymbol{D}_{(\lambda)}, \mathcal{B}^\circ_{(\lambda)})$.

Lemma 1.4.6 *If condition (A3) holds true, then we have* $\lim_{x\to\pm\infty} \mathcal{E}_{p,n}(x)\phi_\lambda(x) = 0$ *and* $\lim_{x\to\pm\infty} \widetilde{\mathcal{E}}_{p,n}(x)\phi_\lambda(x) = 0$ *for every* $n \in \mathbb{N}$. *Moreover, the mappings* $\omega \mapsto \mathcal{E}_{p,n}(x,\omega)\phi_\lambda(x)$ *and* $\omega \mapsto \widetilde{\mathcal{E}}_{p,n}(x,\omega)\phi_\lambda(x)$ *are* $(\mathcal{F}, \mathcal{B}^\circ_{(\lambda)})$*-measurable.*

Proof The measurability easily follows from the fact that $\mathcal{B}^\circ_{(\lambda)}$ coincides with the σ-algebra generated by the one-dimensional coordinate projections. Concerning the first part, we will prove only the latter convergence. The proof of the former convergence follows the same line of arguments and is even easier. Let $n \in \mathbb{N}$. If $x \ge 1$ is sufficiently large such that $x \ge X_{n,i}$ for all $i = 1, \dots, n$, then

$$\sqrt{nb_n}\big(\widehat{F}_{p,n}(x) - \mathbb{E}[\widehat{F}_{p,n}(x)]\big)\,\phi_\lambda(x) \;=\; \sqrt{nb_n}\big(1 - \mathbb{E}[\widehat{F}_{p,n}(x)]\big)\,\phi_\lambda(x). \tag{1.55}$$

We have

$$\begin{aligned}
0 \;&\le\; \limsup_{x\to\infty} \big(1 - \mathbb{E}[\widehat{F}_{p,n}(x)]\big)\,\phi_\lambda(x) \\
&=\; \limsup_{x\to\infty} c_n \sum_{i=1}^n \kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big)\big(1 - \mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x)]\big)\,\phi_\lambda(x) \\
&\le\; \limsup_{x\to\infty} c_n \sum_{i=1}^n \kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big)\,\frac{1}{\phi_\lambda(x)} \int_x^\infty f_{n,i}(y)\,\phi_\lambda(y)\,dy\,\phi_\lambda(x) \\
&\le\; \limsup_{x\to\infty} c_n \sum_{i=1}^n \kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big)\,\|f_{n,i}\|_{(\gamma)} \int_x^\infty \phi_{\lambda-\gamma}(y)\,dy
\end{aligned}$$

$$\leq \limsup_{x\to\infty} \max_{1\leq i\leq n} \|f_{n,i}\|_{(\gamma)} \int_x^\infty \phi_{\lambda-\gamma}(y)\,dy\; c_n \sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)$$
$$= \limsup_{x\to\infty} \max_{1\leq i\leq n} \|f_{n,i}\|_{(\gamma)} \int_x^\infty \phi_{\lambda-\gamma}(y)\,dy.$$

By assumption (A3) we have $\max_{1\leq i\leq n}\|f_{n,i}\|_{(\gamma)} < \infty$ and $\int_0^\infty \phi_{\lambda-\gamma}(y)\,dy < \infty$ (recall $\gamma > 2\lambda+1$, so that $\lambda-\gamma < -\lambda-1$). Thus the latter expression vanishes, which implies that the left-hand side of (1.55) converges to 0 as $x \to \infty$.

If $x \leq -1$ is sufficiently small such that $x \leq X_{n,i}$ for all $i = 1,\dots,n$, then

$$\sqrt{nb_n}\big(\widehat{F}_{p,n}(x) - \mathbb{E}[\widehat{F}_{p,n}(x)]\big)\,\phi_\lambda(x) \;=\; -\sqrt{nb_n}\,\mathbb{E}[\widehat{F}_{p,n}(x)]\,\phi_\lambda(x). \tag{1.56}$$

Proceeding as above we obtain

$$0 \;\geq\; \liminf_{x\to-\infty}\Big(-\mathbb{E}[\widehat{F}_{p,n}(x)]\,\phi_\lambda(x)\Big) \;\geq\; \liminf_{x\to-\infty}\Big(-\max_{1\leq i\leq n}\|f_{n,i}\|_{(\gamma)}\int_{-\infty}^x \phi_{\lambda-\gamma}(y)\,dy\Big)$$

and we can again conclude that the latter expression vanishes, which implies that the left-hand side of (1.56) converges to 0 as $x \to -\infty$. □

Auxiliary results for the proof of assumption (a) in Theorem 1.4.1

Let $X^*_{n,i;i-r} := G_{j_p}(i/n, \boldsymbol{\epsilon}^*_{i,i-r})$ and

$$Y^*_{n,i;i-r}(\boldsymbol{x},\boldsymbol{\lambda}) \;:=\; \sum_{k=1}^d \lambda_k\,\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\,\big(\mathbb{1}_{[X^*_{n,i;i-r},\infty)}(x_k) - \mathbb{E}[\mathbb{1}_{[X^*_{n,i;i-r},\infty)}(x_k)]\big) \tag{1.57}$$

for any $\boldsymbol{x} = (x_1,\dots,x_d) \in \mathbb{R}^d$, $\boldsymbol{\lambda} = (\lambda_1,\dots,\lambda_d) \in \mathbb{R}^d$, $n \in \mathbb{N}$, $i = 1,\dots,n$, and $r \in \mathbb{N}$.

Lemma 1.4.7 *Let assumptions (A1), (A3), and (A5) be fulfilled. Let $a \in [0,1)$ and $q \in (2,\infty)$ be as in (A5), and let $\boldsymbol{\lambda} = (\lambda_1,\dots,\lambda_d) \in \mathbb{R}^d$ be arbitrary. Then there exist constants $C_{\boldsymbol{\lambda},q} > 0$ (depending on $\boldsymbol{\lambda}$ and q) and $n_* \in \mathbb{N}$ such that for any $\boldsymbol{x} = (x_1,\dots,x_d) \in \mathbb{R}^d$, $n \geq n_*$, $i = 1,\dots,n$, and $r \in \mathbb{N}$*

$$\|Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^*_{n,i;i-r}(\boldsymbol{x},\boldsymbol{\lambda})\|_q \;\leq\; C_{\boldsymbol{\lambda},q}\,\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\,a^{r/(2q)}. \tag{1.58}$$

Proof By Minkowski's inequality and Jensen's inequality, we have

$$\|Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^*_{n,i;i-r}(\boldsymbol{x},\boldsymbol{\lambda})\|_q$$
$$= \Big\|\sum_{k=1}^d \lambda_k\,\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)$$
$$\cdot\Big(\mathbb{1}_{[X_{n,i},\infty)}(x_k) - \mathbb{1}_{[X^*_{n,i;i-r},\infty)}(x_k) - \mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x_k) - \mathbb{1}_{[X^*_{n,i;i-r},\infty)}(x_k)\big]\Big)\Big\|_q$$

$$\begin{aligned}
&\le \sum_{k=1}^{d}|\lambda_k|\,\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\\
&\qquad\cdot\Big(\big\|\mathbb{1}_{[X_{n,i},\infty)}(x_k)-\mathbb{1}_{[X^*_{n,i;i-r},\infty)}(x_k)\big\|_q+\big\|\mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x_k)-\mathbb{1}_{[X^*_{n,i;i-r},\infty)}(x_k)\big]\big\|_q\Big)\\
&\le 2\sum_{k=1}^{d}|\lambda_k|\,\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\,\big\|\mathbb{1}_{[X_{n,i},\infty)}(x_k)-\mathbb{1}_{[X^*_{n,i;i-r},\infty)}(x_k)\big\|_q \qquad (1.59)
\end{aligned}$$

for any $\boldsymbol{x}\in\mathbb{R}^d$, $n\in\mathbb{N}$, $i=1,\dots,n$, and $r\in\mathbb{N}$.

As before use j_p to denote the unique index j with $p\in(p_j,p_{j+1})$, and recall that $I_{n;p}$ was defined to be the set of all $i\in\{1,\dots,n\}$ with $i/n\in(p_{j_p},p_{j_p+1})$. Moreover let I_n^{++} be the set of all $i\in\{1,\dots,n\}$ with $\kappa((i-i_{p,n})/(nb_n))\neq 0$. Note that under assumption (A1) we have $i\in I_n^{++}$ only if $|i/n-i_{p,n}/n|\le b_n$. Since $|i_{p,n}/n-p|\le 1/n$, we can conclude that $i\in I_n^{++}$ only if $|i/n-p|\le b_n+n^{-1}$. Now let $n_*\in\mathbb{N}$ so large so that the open ball around p with radius b_n+n^{-1} is contained in (p_{j_p},p_{j_p+1}). Then, for any $n\ge n_*$, we have $i\in I_n^{++}$ only if $i\in I_{n;p}$. That is, $I_n^{++}\subseteq I_{n;p}$ for any $n\ge n_*$.

In view of (1.59), for (1.58) it remains to show that there exists a constant $C_q>0$ such that

$$\|\mathbb{1}_{[X_{n,i},\infty)}(x_k)-\mathbb{1}_{[X^*_{n,i;i-r},\infty)}(x_k)\|_q\le C_q a^{r/(2q)}\quad\text{for all } n\ge n_*,\ i=1,\dots,n,\ r\in\mathbb{N}. \qquad (1.60)$$

To prove (1.60), we set $\delta_r:=\delta_{\boldsymbol{\epsilon},r;q}(G_{j_p})^{1/2}$, where $\delta_{\boldsymbol{\epsilon},r;q}(G_{j_p})$ is the dependence measure of $X_{n,i}$ defined in (1.7), and split the left-hand side into two parts:

$$\begin{aligned}
&\big\|\mathbb{1}_{[X_{n,i},\infty)}(x_k)-\mathbb{1}_{[X^*_{n,i;i-r},\infty)}(x_k)\big\|_q\\
&\quad\le \big\|\big(\mathbb{1}_{[X_{n,i},\infty)}(x_k)-\mathbb{1}_{[X^*_{n,i;i-r},\infty)}(x_k)\big)\,\mathbb{1}_{\{|X_{n,i}-X^*_{n,i;i-r}|\le\delta_r\}}\big\|_q\\
&\qquad+\big\|\big(\mathbb{1}_{[X_{n,i},\infty)}(x_k)-\mathbb{1}_{[X^*_{n,i;i-r},\infty)}(x_k)\big)\,\mathbb{1}_{\{|X_{n,i}-X^*_{n,i;i-r}|>\delta_r\}}\big\|_q\\
&\quad=: S_1(n,i,x_k,r)+S_2(n,i,x_k,r). \qquad (1.61)
\end{aligned}$$

For the first summand we have for any $n\ge n_*$, $i=1,\dots,n$, and $r\in\mathbb{N}$

$$\begin{aligned}
S_1(n,i,x_k,r)&\le \|\mathbb{1}_{\{x_k-\delta_r\le X_{n,i}\le x_k+\delta_r\}}\mathbb{1}_{\{|X_{n,i}-X^*_{n,i;i-r}|\le\delta_r\}}\|_q \le \mathbb{E}\big[|\mathbb{1}_{\{x_k-\delta_r\le X_{n,i}\le x_k+\delta_r\}}|\big]^{1/q}\\
&=\Big(\int_{x_k-\delta_r}^{x_k+\delta_r} f_{n,i}(u)\,du\Big)^{1/q}\le (2C\delta_r)^{1/q} = (2C)^{1/q}\,\delta_{\boldsymbol{\epsilon},r;q}(G_{j_p})^{1/(2q)} \qquad (1.62)
\end{aligned}$$

with $C:=\sup_{n\in\mathbb{N}}\max_{1\le i\le n}\|f_{n,i}\|_\infty$ (recall assumption (A3)). Concerning the second summand, we can apply Markov's inequality to obtain

$$\begin{aligned}
S_2(n,i,x_k,r)&\le \|\mathbb{1}_{\{|X_{n,i}-X^*_{n,i;i-r}|>\delta_r\}}\|_q = \mathbb{P}\big[|X_{n,i}-X^*_{n,i;i-r}|>\delta_r\big]^{1/q} \qquad (1.63)\\
&\le \Big(\frac{1}{\delta_r^q}\,\mathbb{E}\big[|X_{n,i}-X^*_{n,i;i-r}|^q\big]\Big)^{1/q} = \frac{1}{\delta_r}\,\|X_{n,i}-X^*_{n,i;i-r}\|_q \le \delta_{\boldsymbol{\epsilon},r;q}(G_{j_p})^{1/2}.
\end{aligned}$$

By (A5) we may conclude from (1.61)–(1.63) that for any $n \geq n_*$, $i = 1, \ldots, n$, and $r \in \mathbb{N}$

$$\big\|\mathbb{1}_{[X_{n,i},\infty)}(x_k) - \mathbb{1}_{[X^*_{n,i;i-r},\infty)}(x_k)\big\|_q \;\leq\; (2C)^{1/q}\,\widetilde{C}^{1/q}\,a^{r/(2q)} + \widetilde{C}\,a^{r/2} \;\leq\; ((2C)^{1/q}\widetilde{C}^{1/q} + \widetilde{C})\,a^{r/(2q)}.$$

This gives (1.60) with $C_{\boldsymbol{\lambda},q} := 2\sum_{k=1}^d |\lambda_k|((2C)^{1/q}\widetilde{C}^{1/q} + \widetilde{C})$. □

The following lemma involves the projection operator P_k defined in (1.11).

Lemma 1.4.8 *Let assumptions (A1), (A3), and (A5) be fulfilled. Let $a \in [0,1)$ and $q \in (2,\infty)$ be as in (A5), and let $\boldsymbol{\lambda} \in \mathbb{R}^d$ be arbitrary. Then there exist constants $C_{\boldsymbol{\lambda},q} > 0$ (depending on $\boldsymbol{\lambda}$ and q) and $n_* \in \mathbb{N}$ such that for any $\boldsymbol{x} \in \mathbb{R}^d$, $n \geq n_*$, $i = 1, \ldots, n$, and $r \in \mathbb{N}$*

$$\big\|P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big)\big\|_q \leq C_{\boldsymbol{\lambda},q}\,\kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big)\,a^{\max\{m_n,r\}/(2q)}. \tag{1.64}$$

Proof Below we will show in two steps that the following two inequalities hold true:

$$\big\|P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big)\big\|_q \;\leq\; 2\widetilde{C}_{\boldsymbol{\lambda},q}\,\kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big)\,a^{r/(2q)}, \tag{1.65}$$

$$\big\|P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big)\big\|_q \;\leq\; 2\widetilde{C}_{\boldsymbol{\lambda},q}\,\kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big)\,\frac{a^{m_n/(2q)}}{1 - a^{1/(2q)}} \tag{1.66}$$

for some constant $\widetilde{C}_{\boldsymbol{\lambda},q} > 0$. Then (1.65)–(1.66) imply (1.64) with $C_{\boldsymbol{\lambda},q} := 2\,\widetilde{C}_{\boldsymbol{\lambda},q}\,(1 - a^{1/(2q)})^{-1}$.

Step 1. We first show (1.65). We have

$$\begin{aligned}
&\big\|P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big)\big\|_q \\
&\leq\; \|P_{i-r}(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}))\|_q + \|P_{i-r}(Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}))\|_q \\
&=\; \|P_{i-r}(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}))\|_q + \|P_{i-r}(\mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-m_n}])\|_q \\
&=\; \|P_{i-r}(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}))\|_q + \|\mathbb{E}[P_{i-r}(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}))|\boldsymbol{\epsilon}_{i:i-m_n}]\|_q \\
&\leq\; 2\|P_{i-r}(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}))\|_q,
\end{aligned} \tag{1.67}$$

where we used the conditional Jensen inequality for the last step. For the third step we used that

$$\begin{aligned}
\mathbb{E}\big[\mathbb{E}[Y_{n,i}|\boldsymbol{\epsilon}_{i:i-m_n}]|\boldsymbol{\epsilon}_{i-j}\big] &= \mathbb{E}\big[\mathbb{E}[Y_{n,i}|\boldsymbol{\epsilon}_{i:i-m_n}]|\boldsymbol{\epsilon}_{i-j:i-m_n}\big] \;=\; \mathbb{E}[Y_{n,i}|\boldsymbol{\epsilon}_{i-j:i-m_n}] \\
&= \mathbb{E}\big[\mathbb{E}[Y_{n,i}|\boldsymbol{\epsilon}_{i-j}]|\boldsymbol{\epsilon}_{i-j:i-m_n}\big] \\
&= \mathbb{E}\big[\mathbb{E}[Y_{n,i}|\boldsymbol{\epsilon}_{i-j}]|\boldsymbol{\epsilon}_{i:i-m_n}\big]
\end{aligned} \tag{1.68}$$

for all $j \in \mathbb{N}_0$ (with the convention $\sigma(\boldsymbol{\epsilon}_{i-t:i-m_n}) := \{\emptyset,\Omega\}$ if $i - j < i - m_n + 1$); for (1.68) we used that the random variables ε_k, $k \in \mathbb{Z}$, are independent.

Further, note that $\mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i-r-1}] = \mathbb{E}[Y^*_{n,i;i-r}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i-r-1}]$, because ε_{i-r} (which is used for the definition of $Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})$) is independent of $\boldsymbol{\epsilon}_{i-r-1}$ and may thus be replaced by an independent identically distributed copy ε^*. By means of the conditional Jensen inequality, we obtain analogously to the proof of Theorem 1(i) and (ii) in [76]

$$\begin{aligned}
\|P_{i-r}(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}))\|_q &= \big\|\mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i-r}] - \mathbb{E}[Y^*_{n,i;i-r}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i-r-1}]\big\|_q \\
&= \big\|\mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i-r}] - \mathbb{E}[Y^*_{n,i;i-r}(\boldsymbol{x},\boldsymbol{\lambda})|(\boldsymbol{\epsilon}_{i-r-1},\varepsilon_{i-r})]\big\|_q \\
&= \big\|\mathbb{E}\big[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^*_{n,i;i-r}(\boldsymbol{x},\boldsymbol{\lambda})\big|\boldsymbol{\epsilon}_{i-r}\big]\big\|_q \\
&\leq \|Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^*_{n,i;i-r}(\boldsymbol{x},\boldsymbol{\lambda})\|_q. \qquad (1.69)
\end{aligned}$$

Now (1.67), (1.69), and Lemma 1.4.7 together imply (1.65).

Step 2. We now show (1.66). By the conditional Jensen inequality

$$\begin{aligned}
&\big\|P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big)\big\|_q \\
&\leq \big\|\mathbb{E}\big[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big|\boldsymbol{\epsilon}_{i-r}\big]\big\|_q + \big\|\mathbb{E}\big[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big|\boldsymbol{\epsilon}_{i-r-1}\big]\big\|_q \\
&\leq 2\|Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\|_q. \qquad (1.70)
\end{aligned}$$

Since for every $k \in \mathbb{N}$

$$\begin{aligned}
Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) &= Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-m_n}] \\
&= \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_i] - \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-m_n}] \\
&= \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_i] \\
&\quad -\mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-m_n}] + \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-m_n-1}] \\
&\quad -\mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-m_n-1}] + \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-m_n-2}] \\
&\quad \vdots \\
&\quad -\mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-m_n-k}] + \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-m_n-k-1}] \\
&\quad -\mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-m_n-k-1}],
\end{aligned}$$

and $\lim_{k\to\infty} \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-m_n-k-1}] = \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_i]$ $\mathbb{P}$-a.s. (by Corollary 11.1.4 in [18]; see also Theorem 7.4.3 in [32]) we can write

$$Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) = \sum_{j=m_n}^{\infty} \big(\mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-j}] - \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-j-1}]\big).$$

Plugging this in (1.70) gives

$$\begin{aligned}
\|Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\|_q &= \Big\|\sum_{j=m_n}^{\infty} \big(\mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-j}] - \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-j-1}]\big)\Big\|_q \\
&\leq \sum_{j=m_n}^{\infty} \big\|\mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-j}] - \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i:i-j-1}]\big\|_q
\end{aligned}$$

$$\leq \sum_{j=m_n}^{\infty} \|Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^*_{n,i;i-j}(\boldsymbol{x},\boldsymbol{\lambda})\|_q, \tag{1.71}$$

where the last step is valid by the same line of arguments as in (1.69). By Lemma 1.4.7 combined with (1.70) and (1.71), we obtain

$$\begin{aligned}
\|P_{i-r}(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}))\|_q &\leq 2\,\widetilde{C}_{\boldsymbol{\lambda},q}\,\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \sum_{j=m_n}^{\infty} a^{j/(2q)} \\
&\leq 2\,\widetilde{C}_{\boldsymbol{\lambda},q}\,\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \frac{a^{m_n/(2q)}}{1-a^{1/(2q)}}\,.
\end{aligned}$$

This proves (1.66). □

Lemma 1.4.9 *Let assumptions (A1), (A3), and (A5) be fulfilled. Let* $a \in [0,1)$ *and* $q \in (2,\infty)$ *be as in (A5). Then for any* $\boldsymbol{\lambda} \in \mathbb{R}^d$ *and* $\boldsymbol{x} \in \mathbb{R}^d$

$$\Big\| \sum_{i=1}^{n} Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - \sum_{i=1}^{n} Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_q = o\big((nb_n)^{1/2}\big).$$

Proof Since

$$\begin{aligned}
&Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) \\
&\quad= \mathbb{E}\big[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big|\boldsymbol{\epsilon}_i\big] \\
&\quad= \sum_{r=0}^{k} \Big(\mathbb{E}\big[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big|\boldsymbol{\epsilon}_{i-r}\big] - \mathbb{E}\big[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big|\boldsymbol{\epsilon}_{i-r-1}\big]\Big) \\
&\qquad+\mathbb{E}\big[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big|\boldsymbol{\epsilon}_{i-k-1}\big] \\
&\quad= \sum_{r=0}^{k} P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big) \\
&\qquad+\mathbb{E}\big[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big|\boldsymbol{\epsilon}_{i-k-1}\big]
\end{aligned}$$

holds for every $k \in \mathbb{N}$ and $\lim_{k\to\infty} \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{i-k-1}] = \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})] = 0$ $\mathbb{P}$-a.s. (by Corollary 11.1.4 in [18]), we obtain

$$\begin{aligned}
&\Big\| \sum_{i=1}^{n} \big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big) \Big\|_q \\
&\quad= \Big\| \sum_{i=1}^{n} \sum_{r=0}^{\infty} P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big) \Big\|_q \\
&\quad\leq \sum_{r=0}^{\infty} \Big\| \sum_{i=1}^{n} P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big) \Big\|_q, \qquad (1.72)
\end{aligned}$$

where the summands $P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\big)$ form a martingale difference sequence in i with respect to $\sigma(\boldsymbol{\epsilon}_{i-r})$. By Burkholder's inequality (in the form of part (iii) of Corollary 1.4.5), there exists a constant $C_q > 0$ such that

$$\begin{aligned}
&\sum_{r=0}^{\infty} \Big\| \sum_{i=1}^{n} P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\big) \Big\|_q \\
&\quad\le \sum_{r=0}^{\infty} C_q \Big(\sum_{i=1}^{n} \big\|P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\big)\big\|_q^2 \Big)^{1/2} \\
&\quad\le C_q \widetilde{C}_{\boldsymbol{\lambda},q} \sum_{r=0}^{\infty} a^{\max\{r,m_n\}/(2q)} \Big(\sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\Big)^{1/2},
\end{aligned}$$

where the last step is valid by Lemma 1.4.8. Since $(nb_n)^{-1}\sum_{i=1}^n \kappa\big(\frac{i-i_{p,n}}{nb_n}\big)^2$ converges to $\int_{-1}^1 \kappa(u)^2\,du =: \kappa_2$ as $n\to\infty$ due to assumption (A1), we have

$$\begin{aligned}
&\sum_{r=0}^{\infty} \Big\| \sum_{i=1}^{n} P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\big) \Big\|_q \\
&\quad\le C_q\,\widetilde{C}_{\boldsymbol{\lambda},q} \Big(\sum_{r=0}^{m_n-1} a^{m_n/(2q)} + \sum_{r=m_n}^{\infty} a^{r/(2q)} \Big) \Big(nb_n \int_{-1}^{1} \kappa(u)^2\,du + \mathcal{O}(1)\Big)^{1/2} \\
&\quad\le C_q\,\widetilde{C}_{\boldsymbol{\lambda},q} \big(m_n\, a^{m_n/(2q)} + a^{m_n/(2q)}(1-a^{1/(2q)})^{-1}\big)\big(\sqrt{\kappa_2\,(nb_n)} + \mathcal{O}(1)\big) \\
&\quad\le \sqrt{\kappa_2}\, C_q\,\widetilde{C}_{\boldsymbol{\lambda},q}\,(m_n+1)\,(1-a^{1/(2q)})^{-1}\, a^{m_n/(2q)}\big(\sqrt{nb_n} + \mathcal{O}(1)\big). \qquad (1.73)
\end{aligned}$$

Set $C_{\boldsymbol{\lambda},q} := \sqrt{\kappa_2}\, C_q\,\widetilde{C}_{\boldsymbol{\lambda},q}\,(1-a^{1/(2q)})^{-1}$. Then (1.72) and (1.73) imply

$$\Big\| \sum_{i=1}^{n} \big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\big)\Big\|_q \le C_{\boldsymbol{\lambda},q}\,(m_n+1)\,a^{m_n/(2q)}\big(\sqrt{nb_n} + \mathcal{O}(1)\big).$$

Thus $\lim_{n\to\infty}(nb_n)^{-1/2}\|\sum_{i=1}^n(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda}))\|_q = 0$, because $m_n = \lceil\log(n)\rceil$ tends to infinity as $n\to\infty$. □

Lemma 1.4.10 *Let assumptions (A1), (A3), and (A5) be fulfilled. Let $a \in [0,1)$ and $q \in (2,\infty)$ be as in (A5). Then for any $\boldsymbol{\lambda} \in \mathbb{R}^d$ and $\boldsymbol{x} \in \mathbb{R}^d$ there exist constants $C_{\boldsymbol{\lambda},q} > 0$ (depending on $\boldsymbol{\lambda}$ and q) and $n_* \in \mathbb{N}$ such that for $n \ge n_*$*

(i) $\|R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\|_q^2 \le C_{\boldsymbol{\lambda},q}\sum_{i=1}^{l_n}\kappa\big(\frac{(j-1)s_n+i-i_{p,n}}{nb_n}\big)^2$ *for* $j = 1,\dots,\lceil n/s_n\rceil$.

(ii) $\sum_{j=1}^{\lceil n/s_n\rceil}\|R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\|_q^q \le C_{\boldsymbol{\lambda},q}^{q/2}\,\frac{2nb_n-l_n}{s_n}\,\big(\|\kappa\|_\infty^2\, l_n + \mathcal{O}(1)\big)^{q/2}$.

(iii) $\|\sum_{j=1}^{\lceil n/s_n\rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\|_q^2 \le C_{\boldsymbol{\lambda},q}\,\frac{2nb_n-(s_n-l_n)}{s_n}\big(s_n - l_n + \mathcal{O}(1)\big)$.

(iv) $\|\sum_{j=1}^{\lceil n/s_n\rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\|_q^2 \le C_{\boldsymbol{\lambda},q}\frac{2nb_n-l_n}{s_n}\,\big(l_n + \mathcal{O}(1)\big)$.

Here $R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})$ and $r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})$ are defined as in (1.50) and (1.51), respectively.

Proof (i): Let $j \in \{1,\ldots,\lceil n/s_n\rceil\}$. For any $k \in \mathbb{N}$, we clearly have $R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) = \sum_{s=-k}^{l_n} P_{(j-1)s_n+s}(R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})) + \mathbb{E}[R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{(j-1)s_n-k-1}]$ and therefore

$$\begin{aligned}\|R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\|_q^2 &\leq 2\Big\|\sum_{s=-k}^{l_n} P_{(j-1)s_n+s}(R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}))\Big\|_q^2 + 2\big\|\mathbb{E}[R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{(j-1)s_n-k-1}]\big\|_q^2\\ &\leq C_q^2\sum_{s=-k}^{l_n}\big\|P_{(j-1)s_n+s}\big(R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\big)\big\|_q^2 + 2\big\|\mathbb{E}[R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{(j-1)s_n-k-1}]\big\|_q^2,\end{aligned}$$

where we used Burkholder's inequality (in the form of Corollary 1.4.5(iii)) applied to the martingale difference sequence $(P_{(j-1)s_n+s}(R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})))_{s=-k,\ldots,l_n}$ for the second step. Since $\lim_{k\to\infty}\mathbb{E}[R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{(j-1)s_n-k-1}] = \mathbb{E}[R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})] = 0$ $\mathbb{P}$-a.s. by Corollary 11.1.4 in [18], and the sequence $(|\mathbb{E}[R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{(j-1)s_n-k-1}]|)_{k\in\mathbb{N}}$ is bounded by a finite constant (this follows from the same property of the sequence $(|R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})|)_{k\in\mathbb{N}}$), the dominated convergence theorem ensures that the second summand of the bound above converges to 0 as $k\to\infty$. It follows that $\|R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\|_q^2 \leq C_q^2\sum_{s=-\infty}^{l_n}\|P_{(j-1)s_n+s}(R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}))\|_q^2$, and therefore

$$\begin{aligned}&\|R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\|_q^2\\ &\leq C_q^2\sum_{s=-\infty}^{l_n}\Big\|\sum_{i=1}^{l_n}P_{(j-1)s_n+s}(Y^{\{m_n\}}_{n,(j-1)s_n+i}(\boldsymbol{x},\boldsymbol{\lambda}))\Big\|_q^2\\ &= C_q^2\sum_{r=0}^{\infty}\Big\|\sum_{i=1}^{l_n}P_{(j-1)s_n+l_n-r}\big(Y^{\{m_n\}}_{n,(j-1)s_n+i}(\boldsymbol{x},\boldsymbol{\lambda})\big)\Big\|_q^2\\ &= C_q^2\sum_{r=0}^{\infty}\Big\|\sum_{i=1}^{l_n}P_{(j-1)s_n+i-r}\big(Y^{\{m_n\}}_{n,(j-1)s_n+i}(\boldsymbol{x},\boldsymbol{\lambda})\big)\Big\|_q^2\\ &\leq C_q^4\sum_{r=0}^{\infty}\sum_{i=1}^{l_n}\big\|P_{(j-1)s_n+i-r}\big(Y^{\{m_n\}}_{n,(j-1)s_n+i}(\boldsymbol{x},\boldsymbol{\lambda})\big)\big\|_q^2\\ &= C_q^4\sum_{r=0}^{\infty}\sum_{i=1}^{l_n}\big\|\mathbb{E}\big[P_{(j-1)s_n+i-r}\big(Y_{n,(j-1)s_n+i}(\boldsymbol{x},\boldsymbol{\lambda})\big)|\boldsymbol{\epsilon}_{(j-1)s_n+i:(j-1)s_n+i-m_n}\big]\big\|_q^2\\ &\leq C_q^4\sum_{r=0}^{\infty}\sum_{i=1}^{l_n}\big\|P_{(j-1)s_n+i-r}\big(Y_{n,(j-1)s_n+i}(\boldsymbol{x},\boldsymbol{\lambda})\big)\big\|_q^2,\end{aligned}\tag{1.74}$$

where the fourth step is valid by part (iii) of Corollary 1.4.5 applied to the martingale difference sequence $(P_{(j-1)s_n+i-r}(Y^{\{m_n\}}_{n,(j-1)s_n+i}(\boldsymbol{x},\boldsymbol{\lambda})))_{i=1,\ldots,l_n}$, the fifth step is valid by an analogous argumentation as in (1.68), and the last step is ensured by the conditional

Jensen inequality. By (1.69) and Lemma 1.4.7 we have for any $n \geq n_*$

$$\begin{aligned}
&\sum_{r=0}^{\infty}\sum_{i=1}^{l_n}\big\|P_{(j-1)s_n+i-r}\big(Y_{n,(j-1)s_n+i}(\boldsymbol{x},\boldsymbol{\lambda})\big)\big\|_q^2 \\
&\quad\leq \sum_{r=0}^{\infty}\sum_{i=1}^{l_n}\big\|Y_{n,(j-1)s_n+i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^*_{n,(j-1)s_n+i;(j-1)s_n+i-r}(\boldsymbol{x},\boldsymbol{\lambda})\big\|_q^2 \\
&\quad\leq \widetilde{C}^2_{\boldsymbol{\lambda},q}\sum_{r=0}^{\infty} a^{r/q}\sum_{i=1}^{l_n}\kappa\Big(\frac{(j-1)s_n+i-i_{p,n}}{nb_n}\Big)^2 \\
&\quad= \widetilde{C}^2_{\boldsymbol{\lambda},q}\,\frac{1}{1-a^{1/q}}\sum_{i=1}^{l_n}\kappa\Big(\frac{(j-1)s_n+i-i_{p,n}}{nb_n}\Big)^2 \qquad (1.75)
\end{aligned}$$

with n_* as in Lemma 1.4.7. Now (1.74) and (1.75) yield (i) with $C_{\boldsymbol{\lambda},q} := \widetilde{C}^2_{\boldsymbol{\lambda},q}\,C_q^4\,(1 - a^{1/q})^{-1}$.

(ii): By assertion (i), we have for any $n \geq n_*$

$$\begin{aligned}
&\sum_{j=1}^{\lceil n/s_n\rceil}\big\|R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\big\|_q^q \\
&\quad\leq \sum_{j=1}^{\lceil n/s_n\rceil}\Big(C_{\boldsymbol{\lambda},q}\sum_{i=1}^{l_n}\kappa\Big(\frac{(j-1)s_n+i-i_{p,n}}{nb_n}\Big)^2\Big)^{q/2} \\
&\quad= \sum_{j=\lceil (i_{p,n}-nb_n)/s_n+1\rceil}^{\lfloor (i_{p,n}+nb_n-l_n)/s_n+1\rfloor} C^{q/2}_{\boldsymbol{\lambda},q}\Big(nb_n\int_{((j-1)s_n-i_{p,n})/(nb_n)}^{((j-1)s_n+l_n-i_{p,n})/(nb_n)}\kappa(u)^2\,du + \mathcal{O}(1)\Big)^{q/2} \\
&\quad\leq \sum_{j=\lceil (i_{p,n}-nb_n)/s_n+1\rceil}^{\lfloor (i_{p,n}+nb_n-l_n)/s_n+1\rfloor} C^{q/2}_{\boldsymbol{\lambda},q}\Big(nb_n\|\kappa\|^2_\infty\frac{l_n}{nb_n} + \mathcal{O}(1)\Big)^{q/2} \\
&\quad\leq C^{q/2}_{\boldsymbol{\lambda},q}\,\frac{2nb_n-l_n}{s_n}\,\big(\|\kappa\|^2_\infty l_n + \mathcal{O}(1)\big)^{q/2},
\end{aligned}$$

where we used in the second step that the kernel function κ has support on $[-1,1]$ (recall (A1)) and we therefore sum over less than $\lceil n/s_n\rceil$ many summands.

(iii): For any $k \in \mathbb{N}$ we have

$$\sum_{j=1}^{\lceil n/s_n\rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) = \sum_{s=-k}^{\lceil n/s_n\rceil s_n} P_s\Big(\sum_{j=1}^{\lceil n/s_n\rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\Big) + \mathbb{E}\Big[\sum_{j=1}^{\lceil n/s_n\rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\Big|\boldsymbol{\epsilon}_{-k-1}\Big],$$

where $\lim_{k\to\infty}\mathbb{E}[\sum_{j=1}^{\lceil n/s_n\rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{-k-1}] = \mathbb{E}[\sum_{j=1}^{\lceil n/s_n\rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})] = 0$ $\mathbb{P}$-a.s. by Corollary 11.1.4 in [18], and where $(P_s(\sum_j r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}))_{s=-k,\dots,\lceil n/s_n\rceil s_n}$ is a martingale difference sequence. By the same application of Burkholder's inequality as at the beginning of the proof of (i) we obtain

$$\begin{aligned}
\Big\| \sum_{j=1}^{\lceil n/s_n \rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_q^2 &\le C_q^2 \sum_{s=-\infty}^{\lceil n/s_n \rceil s_n} \Big\| P_s\Big(\sum_{j=1}^{\lceil n/s_n \rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big) \Big\|_q^2 \\
&= C_q^2 \sum_{r=s_n}^{\infty} \Big\| \sum_{j=1}^{\lceil n/s_n \rceil} P_{(\lceil n/s_n \rceil -1)s_n - r}\big(r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\big) \Big\|_q^2 \\
&= C_q^2 \sum_{r=s_n}^{\infty} \Big\| \sum_{j=1}^{\lceil n/s_n \rceil} P_{(j-1)s_n - r}\big(r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\big) \Big\|_q^2,
\end{aligned}$$

where the projections $P_{(j-1)s_n-r}(r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}))$ form again a martingale difference sequence in j for fixed r. Due to part (iii) of Corollary 1.4.5

$$\begin{aligned}
\Big\| \sum_{j=1}^{\lceil n/s_n \rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_q^2 &\le C_q^4 \sum_{r=s_n}^{\infty} \sum_{j=1}^{\lceil n/s_n \rceil} \big\| P_{(j-1)s_n - r}\big(r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\big) \big\|_q^2 \\
&= C_q^4 \sum_{r=0}^{\infty} \sum_{j=1}^{\lceil n/s_n \rceil} \Big\| \sum_{i=l_n+1}^{s_n} P_{(j-1)s_n + s_n - r}\big(Y_{n,(j-1)s_n+i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\big) \Big\|_q^2 \\
&= C_q^4 \sum_{r=0}^{\infty} \sum_{j=1}^{\lceil n/s_n \rceil} \Big\| \sum_{i=l_n+1}^{s_n} P_{(j-1)s_n + i - r}\big(Y_{n,(j-1)s_n+i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\big) \Big\|_q^2.
\end{aligned}$$

Applying one more time Burkholder's inequality (in form of Corollary 1.4.5 (iii)) to the martingale difference sequence $(P_{(j-1)s_n+i-r}(Y_{n,(j-1)s_n+i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})))_{i=l_n+1,\dots,s_n}$ yields

$$\begin{aligned}
&\Big\| \sum_{j=1}^{\lceil n/s_n \rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_q^2 \\
&\le C_q^6 \sum_{r=0}^{\infty} \sum_{j=1}^{\lceil n/s_n \rceil} \sum_{i=l_n+1}^{s_n} \big\| P_{(j-1)s_n+i-r}\big(Y_{n,(j-1)s_n+i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\big) \big\|_q^2 \\
&= C_q^6 \sum_{r=0}^{\infty} \sum_{j=1}^{\lceil n/s_n \rceil} \sum_{i=l_n+1}^{s_n} \big\| \mathbb{E}\big[P_{(j-1)s_n+i-r}\big(Y_{n,(j-1)s_n+i}(\boldsymbol{x},\boldsymbol{\lambda})\big) \big| \boldsymbol{\epsilon}_{(j-1)s_n+i:(j-1)s_n+i-m_n} \big] \big\|_q^2 \\
&\le C_q^6 \sum_{r=0}^{\infty} \sum_{j=1}^{\lceil n/s_n \rceil} \sum_{i=l_n+1}^{s_n} \big\| P_{(j-1)s_n+i-r}\big(Y_{n,(j-1)s_n+i}(\boldsymbol{x},\boldsymbol{\lambda})\big) \big\|_q^2,
\end{aligned}$$

where the second step is valid by the same argumentation as in (1.68) and the third step by the conditional Jensen inequality. By (1.69) and Lemma 1.4.7, we now obtain for any $n \ge n_*$

$$\Big\| \sum_{j=1}^{\lceil n/s_n \rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_q^2$$

$$
\begin{aligned}
&\le C_q^6 \sum_{r=0}^{\infty} \sum_{j=1}^{\lceil n/s_n \rceil} \sum_{i=l_n+1}^{s_n} \big\| Y_{n,(j-1)s_n+i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^*_{n,(j-1)s_n+i;(j-1)s_n+i-r}(\boldsymbol{x},\boldsymbol{\lambda}) \big\|_q^2 \\
&\le C_q^6 \, \widetilde{C}^2_{\boldsymbol{\lambda},q} \sum_{r=0}^{\infty} a^{r/q} \sum_{j=1}^{\lceil n/s_n \rceil} \sum_{i=l_n+1}^{s_n} \kappa\Big(\frac{(j-1)s_n + i - i_{p,n}}{nb_n}\Big)^2 \\
&= C_q^6 \, \widetilde{C}^2_{\boldsymbol{\lambda},q} \frac{1}{1-a^{1/q}} \sum_{j=\lceil (i_{p,n}-nb_n-l_n)/s_n+1 \rceil}^{\lfloor (i_{p,n}+nb_n)/s_n \rfloor} \Big(nb_n \int_{((j-1)s_n+l_n-i_{p,n})/(nb_n)}^{(js_n-i_{p,n})/(nb_n)} \kappa(u)^2 \, du + \mathcal{O}(1) \Big) \\
&\le C_q^6 \, \widetilde{C}^2_{\boldsymbol{\lambda},q} \frac{1}{1-a^{1/q}} \sum_{j=\lceil (i_{p,n}-nb_n-l_n)/s_n+1 \rceil}^{\lfloor (i_{p,n}+nb_n)/s_n \rfloor} \Big(nb_n \, \|\kappa\|_\infty^2 \frac{s_n - l_n}{nb_n} + \mathcal{O}(1) \Big) \\
&\le C_{\boldsymbol{\lambda},q} \frac{2nb_n + l_n - s_n}{s_n} \big(s_n - l_n + \mathcal{O}(1) \big)
\end{aligned}
$$

with $C_{\boldsymbol{\lambda},q} := C_q^6 \, \widetilde{C}^2_{\boldsymbol{\lambda},q} (1-a^{1/q})^{-1}(\|\kappa\|_\infty^2 + 1)$, where the third step is valid because κ has support on $[-1, 1]$. This proves (iii).

(iv): The fourth assertion can be verified by the same steps as in the proof of (iii). □

Remark 1.4.11 Recall $l_n = \lceil \sqrt{nb_n} \rceil$ and $s_n = l_n + \lceil (\log n)^2 \rceil$. Thus if in addition to (A1), (A3), (A5) also condition (A2) holds true, then parts (iii) and (iv) of Lemma 1.4.10 imply

$$
\Big\| \sum_{j=1}^{\lceil n/s_n \rceil} r_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_q^2 = o(nb_n) \qquad \text{and} \qquad \Big\| \sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_q^2 = \mathcal{O}(nb_n), \tag{1.76}
$$

respectively. In particular,

$$
\Big\| \sum_{i=1}^{n} Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda}) - \sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_q^2 = o(nb_n), \tag{1.77}
$$

because the left-hand side of (1.77) coincides with $\| \sum_{j=1}^{\lceil n/s_n \rceil} r_{n,j}(\boldsymbol{x}\boldsymbol{\lambda}) \|_q^2$. ◇

Lemma 1.4.12 *Let $m_n = \lceil \log(n) \rceil$ and let assumptions (A1)–(A5) and (A9) be fulfilled. Then for every $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $\boldsymbol{\lambda} := (\lambda_1, \ldots, \lambda_d) \in \mathbb{R}^d$*

(i) $\big| \mathbb{V}\mathrm{ar}\big[\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) \big] - \mathbb{V}\mathrm{ar}\big[\sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda}) \big] \big| = o(nb_n)$.

(ii) $\big| \mathbb{V}\mathrm{ar}\big[\sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda}) \big] - \mathbb{V}\mathrm{ar}\big[\sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \big] \big| = o(nb_n)$.

(iii) $\mathbb{V}\mathrm{ar}\big[\sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \big] = \mathcal{O}(nb_n)$.

(iv) $\lim_{n\to\infty} \mathbb{V}\mathrm{ar}\big[c_n\sqrt{nb_n}\sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\big] = \mathbb{V}\mathrm{ar}\big[\sum_{k=1}^d \lambda_k B_p(x_k)\big]$.

Proof (i): Since $\mathbb{E}[\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})] = 0$ and $\mathbb{E}[\sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})] = 0$, we have

$$\begin{aligned}
&\Big|\mathbb{V}\mathrm{ar}\Big[\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\Big] - \mathbb{V}\mathrm{ar}\Big[\sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\Big]\Big| \\
&= \Big|\Big\|\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_2^2 - \Big\|\sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_2^2\Big| \\
&\le \Big\|\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - \sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_2\Big(\Big\|\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_2 + \Big\|\sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_2\Big).
\end{aligned}$$

By Lemma 1.4.9, $\|\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - \sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\|_2 = o((nb_n)^{1/2})$. For the proof of (i), it thus suffices to show that

$$\Big\|\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_q + \Big\|\sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_q = \mathcal{O}\big((nb_n)^{1/2}\big) \tag{1.78}$$

for some $q \in [2,\infty)$. We let q be as in (A5). For the proof of (1.78), we note that $Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})$ and $Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})$ can be written as telescoping sums $\sum_{r=0}^\infty P_{i-r}(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}))$ and $\sum_{r=0}^\infty P_{i-r}(Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda}))$, respectively, because we have $\lim_{k\to\infty}\mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{-k}] = \mathbb{E}[Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})] = 0$ $\mathbb{P}$-a.s. and $\lim_{k\to\infty}\mathbb{E}[Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})|\boldsymbol{\epsilon}_{-k}] = \mathbb{E}[Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})] = 0$ $\mathbb{P}$-a.s. by Corollary 11.1.4 in [18]. Then

$$\begin{aligned}
&\Big\|\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_q + \Big\|\sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_q \\
&\le \sum_{r=0}^\infty \Big\|\sum_{i=1}^n P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big)\Big\|_q + \sum_{r=0}^\infty \Big\|\sum_{i=1}^n P_{i-r}\big(Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\big)\Big\|_q \\
&\le C_q \sum_{r=0}^\infty \Big(\sum_{i=1}^n \big\|P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big)\big\|_q^2\Big)^{1/2} + C_q \sum_{r=0}^\infty \Big(\sum_{i=1}^n \big\|P_{i-r}\big(Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\big)\big\|_q^2\Big)^{1/2}
\end{aligned}$$

for some constant $C_q > 0$, where we applied Burkholder's inequality in the form of Corollary 1.4.5(iii) to the martingale difference sequences $(P_{i-r}(Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})))_{i=1,\dots,n}$ and $(P_{i-r}(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})))_{i=1,\dots,n}$, respectively, in the second step. Since $\|P_{i-r}(Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda}))\|_q = \|\mathbb{E}\big[P_{i-r}(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}))|\boldsymbol{\epsilon}_{i:i-m_n}\big]\|_q \le \|P_{i-r}(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}))\|_q$ by the conditional Jensen inequality, we further obtain by means of (1.69) and Lemma 1.4.7

$$\begin{aligned}
&\Big\|\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_q + \Big\|\sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_q \\
&\le 2\,C_q \sum_{r=0}^\infty \Big(\sum_{i=1}^n \big\|P_{i-r}\big(Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\big)\big\|_q^2\Big)^{1/2}
\end{aligned}$$

$$
\begin{aligned}
&\le 2\,C_q \sum_{r=0}^{\infty} \Big(\sum_{i=1}^{n} \big\| Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - Y^*_{n,i;i-r}(\boldsymbol{x},\boldsymbol{\lambda}) \big\|_q^2 \Big)^{1/2} \\
&\le 2\,C_q \sum_{r=0}^{\infty} \Big(C^2_{\boldsymbol{\lambda},q}\, a^{r/q} \sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2 \Big)^{1/2} \\
&\le 2\,C_q\, C_{\boldsymbol{\lambda},q} \sum_{r=0}^{\infty} a^{r/2q} \Big(nb_n \int_{-1}^{1} \kappa(u)^2\, du + \mathcal{O}(1) \Big)^{1/2}
\end{aligned}
$$

for any $n \ge n_*$ (with n_* as in Lemma 1.4.7). In view of $\kappa_2 := \int_{-1}^1 \kappa(u)^2\, du < \infty$, this implies (1.78).

(ii): To prove the second assertion, we observe

$$
\begin{aligned}
&\Big| \mathbb{V}\mathrm{ar}\Big[\sum_{i=1}^{n} Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) \Big] - \mathbb{V}\mathrm{ar}\Big[\sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big] \Big| \\
&= \Big| \Big\| \sum_{i=1}^{n} Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_2^2 - \Big\| \sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_2^2 \Big| \\
&\le \Big\| \sum_{i=1}^{n} Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - \sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_2 \\
&\qquad \cdot \Big(\Big\| \sum_{i=1}^{n} Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_2 + \Big\| \sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_2 \Big).
\end{aligned}
$$

For $q \in (2,\infty)$ as in (A5) we have $\| \sum_{i=1}^n Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - \sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \|_q = o((nb_n)^{1/2})$ by (1.77), and $\| \sum_{i=1}^n Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) \|_q + \| \sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \|_q = \mathcal{O}((nb_n)^{1/2})$ by (1.78) and (1.76). Along with $\|\cdot\|_2 \le \|\cdot\|_q$ this gives assertion (ii).

(iii): The third assertion follows directly from

$$
\mathbb{V}\mathrm{ar}\Big[\sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big] = \Big\| \sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_2^2 \le \Big\| \sum_{j=1}^{\lceil n/s_n \rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda}) \Big\|_q^2
$$

(for any $q \in (2,\infty)$) and (1.76).

(iv): For the proof of the fourth assertion we observe

$$
\begin{aligned}
&\mathbb{V}\mathrm{ar}\Big[c_n \sqrt{nb_n} \sum_{i=1}^{n} Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) \Big] \\
&\le c_n^2 (nb_n) \Big| \mathbb{V}\mathrm{ar}\Big[\sum_{i=1}^{n} Y^{\{m_n\}}_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) \Big] - \mathbb{V}\mathrm{ar}\Big[\sum_{i=1}^{n} Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) \Big] \Big| \\
&\quad + \mathbb{V}\mathrm{ar}\Big[c_n \sqrt{nb_n} \sum_{i=1}^{n} Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) \Big].
\end{aligned}
$$

The first summand converges to 0 as $n \to \infty$ by part (i) and (1.12). For the second summand we have $\mathbb{V}\mathrm{ar}[c_n\sqrt{nb_n}\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})] = \sum_{k=1}^d \sum_{l=1}^d \lambda_k\lambda_l \mathbb{E}[\widetilde{\mathcal{E}}_{p,n}(x_k)\widetilde{\mathcal{E}}_{p,n}(x_l)]$. Thus it converges to $\sum_{k=1}^d \sum_{l=1}^d \lambda_k\lambda_l\, \gamma_p(x_k,x_l) = \mathbb{V}\mathrm{ar}[\sum_{k=1}^d \lambda_k B_p(x_k)]$ as $n \to \infty$ by Lemma 1.2.2. This finishes the proof. □

Lemma 1.4.13 *Let assumptions (A1)–(A5) and (A9) be fulfilled. Then for any* $\boldsymbol{x} = (x_1,\ldots,x_d) \in \mathbb{R}^d$ *and* $\boldsymbol{\lambda} := (\lambda_1,\ldots,\lambda_d) \in \mathbb{R}^d$ *with* $\sum_{k=1}^d \sum_{l=1}^d \lambda_k\lambda_l\gamma_p(x_k,x_l) \neq 0$ *there exist constants* $c_{\boldsymbol{\lambda}} > 0$ *and* $n_0 \in \mathbb{N}$ *such that for any* $n \geq n_0$

$$\frac{1}{nb_n}\mathbb{V}\mathrm{ar}\Big[\sum_{j=1}^{\lceil n/s_n\rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\Big] \geq c_{\boldsymbol{\lambda}}. \tag{1.79}$$

Proof Let $\boldsymbol{x},\boldsymbol{\lambda} \in \mathbb{R}^d$. The limit $\lim_{n\to\infty}\mathbb{V}\mathrm{ar}[(nb_n)^{-1/2}\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})]$ exists and equals $\sum_{k=1}^d\sum_{l=1}^d \lambda_k\lambda_l\gamma_p(x_k,x_l)$ $(= \mathbb{V}\mathrm{ar}[\sum_{j=1}^n \lambda_j B_p(x_j)] \geq 0)$ by Lemma 1.2.2. By assumption the latter expression is distinct from zero, so that we can find constants $c_{\boldsymbol{\lambda}} > 0$ and $n_1 \in \mathbb{N}$ such that $\frac{1}{nb_n}\mathbb{V}\mathrm{ar}[\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})] \geq 4c_{\boldsymbol{\lambda}}$ for any $n \geq n_1$. Thus

$$\begin{aligned}
4c_{\boldsymbol{\lambda}} &\leq \frac{1}{nb_n}\mathbb{V}\mathrm{ar}\Big[\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\Big] \;=\; \frac{1}{nb_n}\Big\|\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_2^2 \\
&\leq \frac{2}{nb_n}\Big\|\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - \sum_{j=1}^{\lceil n/s_n\rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_2^2 + \frac{2}{nb_n}\Big\|\sum_{j=1}^{\lceil n/s_n\rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_2^2 \\
&=: S_1(n;\boldsymbol{x},\boldsymbol{\lambda}) + S_2(n;\boldsymbol{x},\boldsymbol{\lambda})
\end{aligned}$$

for any $n \geq n_1$. For the first summand we have

$$\begin{aligned}
S_1(n;\boldsymbol{x},\boldsymbol{\lambda}) &\leq \frac{4}{nb_n}\Big\|\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda}) - \sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_q^2 \\
&\quad + \frac{4}{nb_n}\Big\|\sum_{i=1}^n Y_{n,i}^{\{m_n\}}(\boldsymbol{x},\boldsymbol{\lambda}) - \sum_{j=1}^{\lceil n/s_n\rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\Big\|_q^2
\end{aligned} \tag{1.80}$$

for any $q \in [2,\infty)$. Letting $q \in (2,\infty)$ be as in (A5), Lemma 1.4.9 and (1.77) ensure that both summands on the right-hand side of (1.80) converge to 0 as $n \to \infty$. Therefore, we can find $n_0 \geq n_1$ such that $2c_{\boldsymbol{\lambda}} \leq S_2(n;\boldsymbol{x},\boldsymbol{\lambda})$ for any $n \geq n_0$. Since $\mathbb{V}\mathrm{ar}[\sum_{j=1}^{\lceil n/s_n\rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})] = \|\sum_{j=1}^{\lceil n/s_n\rceil} R_{n,j}(\boldsymbol{x},\boldsymbol{\lambda})\|_2^2$, this gives (1.79). □

Auxiliary results for the proof of assumption (b) in Theorem 1.4.1

Choose $\gamma \in (2\lambda+1,\infty)$ in such a way that condition (A3) is fulfilled. Moreover let $q \in (2,(\gamma-1)/\lambda)$.

Lemma 1.4.14 *Let assumptions (A1)–(A2) and (A6) be fulfilled. Then there exist constants $C_{1,q}, C_{2,q} > 0$ and $n_* \in \mathbb{N}$ such that for any $x, y \in \mathbb{R}$ with $x < y$ and $n \geq n_*$*

$$\begin{aligned}
&\|H_{p,n}(y) - H_{p,n}(x)\|_q^q \\
&\leq \; C_{1,q}(nb_n)^{-q/2}\Big(\sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^{\min\{q,4\}}\big(\mathbb{P}\big[X_{n,i} \leq y\big] - \mathbb{P}\big[X_{n,i} \leq x\big]\big)^{\min\{1,4/q\}}\Big)^{\max\{1,q/4\}} \\
&\quad + C_{2,q}\,(y-x)^{(q-2)/2}\int_x^y \max_{i \in I_{n;p}} \|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,du\,(1+\mathcal{O}((nb_n)^{-q/2})). \qquad (1.81)
\end{aligned}$$

Proof For notational simplicity we set $d_{n,i}(x,y) := \mathbb{1}_{\{x<X_{n,i}\leq y\}} - \mathbb{E}[\mathbb{1}_{\{x<X_{n,i}\leq y\}}|\boldsymbol{\epsilon}_{i-1}]$ for $x, y \in \mathbb{R}$ with $x < y$ and $i = 1, \ldots, n$. Since $\kappa(\frac{i-i_{p,n}}{nb_n})\,d_{n,i}(x,y)$ form a martingale difference sequence in i, we may apply Burkholder's inequality (in the form of part (ii) of Corollary 1.4.5) to obtain

$$\begin{aligned}
&\big\|H_{p,n}(y) - H_{p,n}(x)\big\|_q^q \\
&= \; \big(c_n\sqrt{nb_n}\big)^q\Big(\Big\|\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)d_{n,i}(x,y)\Big\|_q^2\Big)^{q/2} \\
&\leq \; \big(c_n\sqrt{nb_n}\big)^q\Big(C_q^2\Big\|\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2 d_{n,i}(x,y)^2\Big\|_{q/2}\Big)^{q/2} \\
&\leq \; 2^{q/2-1}\,C_q^q\big(c_n\sqrt{nb_n}\big)^q\Big\|\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\big(d_{n,i}(x,y)^2 - \mathbb{E}[d_{n,i}(x,y)^2|\boldsymbol{\epsilon}_{i-1}]\big)\Big\|_{q/2}^{q/2} \\
&\quad + 2^{q/2-1}C_q^q\,\big(c_n\sqrt{nb_n}\big)^q\Big\|\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\mathbb{E}[d_{n,i}(x,y)^2|\boldsymbol{\epsilon}_{i-1}]\Big\|_{q/2}^{q/2} \\
&=: \; 2^{q/2-1}\,C_q^q\,\big(c_n\sqrt{nb_n}\big)^q\big(S_1(n,x,y) + S_2(n,x,y)\big). \qquad (1.82)
\end{aligned}$$

We note that $\kappa(\frac{i-i_{p,n}}{nb_n})^2(d_{n,i}(x,y)^2 - \mathbb{E}[d_{n,i}(x,y)^2|\boldsymbol{\epsilon}_{i-1}])$ in the first summand form a martingale difference sequence in i. Applying Burkholder's inequality in the form of part (iii) of Corollary 1.4.5 yields

$$\begin{aligned}
&S_1(n,x,y) \\
&= \; \Big(\Big\|\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\big(d_{n,i}(x,y)^2 - \mathbb{E}[d_{n,i}(x,y)^2|\boldsymbol{\epsilon}_{i-1}]\big)\Big\|_{q/2}^{\min\{q/2,2\}}\Big)^{q/(2\min\{q/2,2\})} \\
&\leq \; C_{q/2}^{q/2}\Big(\sum_{i=1}^{n}\Big\|\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\big(d_{n,i}(x,y)^2 - \mathbb{E}[d_{n,i}(x,y)^2|\boldsymbol{\epsilon}_{i-1}]\big)\Big\|_{q/2}^{\min\{q/2,2\}}\Big)^{\max\{1,q/4\}} \\
&\leq \; C_{q/2}^{q/2}\Big(\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^{\min\{q,4\}}2^{\min\{q/2,2\}-1} \\
&\qquad \cdot\Big\{\big\|d_{n,i}(x,y)^2\big\|_{q/2}^{\min\{q/2,2\}} + \big\|\mathbb{E}[d_{n,i}(x,y)^2|\boldsymbol{\epsilon}_{i-1}]\big\|_{q/2}^{\min\{q/2,2\}}\Big\}\Big)^{\max\{1,q/4\}}.
\end{aligned}$$

$$
\begin{aligned}
&\le C_{q/2}^{q/2}\Big(\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^{\min\{q,4\}}2^{\min\{q/2,2\}}\big\|d_{n,i}(x,y)^2\big\|_{q/2}^{\min\{q/2,2\}}\Big)^{\max\{1,q/4\}} \\
&= 2^{q/2}\,C_{q/2}^{q/2}\Big(\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^{\min\{q,4\}}\big\|d_{n,i}(x,y)\big\|_{q}^{\min\{q,4\}}\Big)^{\max\{1,q/4\}}, \qquad (1.83)
\end{aligned}
$$

where we used $\|\mathbb{E}[d_{n,i}(x,y)^2|\boldsymbol{\epsilon}_{i-1}]\|_{q/2}^{\min\{q/2,2\}}\le\|d_{n,i}(x,y)^2\|_{q/2}^{\min\{q/2,2\}}$ (ensured by the conditional Jensen inequality) for the second-last step. By the conditional Jensen inequality

$$
\begin{aligned}
\big\|d_{n,i}(x,y)\big\|_{q}^{\min\{q,4\}} &\le 2^{\min\{q,4\}-1}\big(\|\mathbb{1}_{\{x<X_{n,i}\le y\}}\|_{q}^{\min\{q,4\}}+\|\mathbb{E}\big[\mathbb{1}_{\{x<X_{n,i}\le y\}}|\boldsymbol{\epsilon}_{i-1}\big]\|_{q}^{\min\{q,4\}}\big) \\
&\le 2^{\min\{q,4\}}\|\mathbb{1}_{\{x<X_{n,i}\le y\}}\|_{q}^{\min\{q,4\}} = 2^{\min\{q,4\}}\mathbb{E}\big[|\mathbb{1}_{\{x<X_{n,i}\le y\}}|\big]^{\min\{1,4/q\}} \\
&= 2^{\min\{q,4\}}\big(\mathbb{P}\big[X_{n,i}\le y\big]-\mathbb{P}\big[X_{n,i}\le x\big]\big)^{\min\{1,4/q\}}.
\end{aligned}
$$

Thus, in view of (1.83),

$$
\begin{aligned}
&2^{q/2-1}\,C_q^q\,(c_n\sqrt{nb_n})^q\,S_1(n,x,y) \qquad (1.84)\\
&\le C_{1,q}(nb_n)^{-q/2}\Big(\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^{\min\{q,4\}}\big(\mathbb{P}\big[X_{n,i}\le y\big]-\mathbb{P}\big[X_{n,i}\le x\big]\big)^{\min\{1,4/q\}}\Big)^{\max\{1,q/4\}}
\end{aligned}
$$

with $C_{1,q}:=2^{2q-1}\,C_q^q\,C_{q/2}^{q/2}C^q$ for some suitable constant $C>0$ (such that $c_n\sqrt{nb_n}\le C(nb_n)^{-1/2}$; recall (1.12)).

For the second summand, we expand $d_{n,i}(x,y)^2$ and obtain

$$
\begin{aligned}
S_2(n,x,y) &= \Big\|\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\,\mathbb{E}[\mathbb{1}_{\{x<X_{n,i}\le y\}}|\boldsymbol{\epsilon}_{i-1}]\big(1-\mathbb{E}[\mathbb{1}_{\{x<X_{n,i}\le y\}}|\boldsymbol{\epsilon}_{i-1}]\big)\Big\|_{q/2}^{q/2} \\
&\le \Big\|\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\,\mathbb{E}[\mathbb{1}_{\{x<X_{n,i}\le y\}}|\boldsymbol{\epsilon}_{i-1}]\Big\|_{q/2}^{q/2}.
\end{aligned}
$$

By assumption (A6), Hölder inequality, and Fubini's theorem we can conclude that for any $n\ge n_*$ (with n_* as in the proof of Lemma 1.4.7)

$$
\begin{aligned}
S_2(n,x,y) &= \Big\|\int_x^y\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\,\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\,du\Big\|_{q/2}^{q/2} \\
&\le \Big\|(y-x)^{(q-2)/q}\Big(\int_x^y\Big|\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\,\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\Big|^{q/2}du\Big)^{2/q}\Big\|_{q/2}^{q/2} \\
&= (y-x)^{(q-2)/2}\int_x^y\Big\|\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\,\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\Big\|_{q/2}^{q/2}du \\
&\le (y-x)^{(q-2)/2}\int_x^y\Big(\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\max_{j\in I_{n;p}}\|\mathfrak{f}_{n,j}(u,\boldsymbol{\epsilon}_{j-1})\|_{q/2}\Big)^{q/2}du.
\end{aligned}
$$

By (A1) we have $\lim_{n\to\infty}(nb_n)^{-1}\sum_{i=1}^n \kappa(\frac{i-i_{p,n}}{nb_n})^2 = \int_{-1}^1 \kappa(u)^2\,du =: \kappa_2$, and (1.12) implies $(nb_n)^{q/2} \le C\big(c_n\sqrt{nb_n}\big)^{-q}$ for some constant $C>0$. Thus

$$
\begin{aligned}
&S_2(n,x,y)\\
&\le\ (y-x)^{(q-2)/2}\int_x^y \Big(\Big(nb_n\int_{-1}^1 \kappa(u)^2\,du + \mathcal{O}(1)\Big)\max_{j\in I_{n;p}}\|\mathfrak{f}_{n,j}(u,\boldsymbol{\epsilon}_{j-1})\|_{q/2}\Big)^{q/2} du\\
&\le\ 2^{q/2-1}\kappa_2^{q/2}(y-x)^{(q-2)/2}\big((nb_n)^{q/2}+\mathcal{O}(1)\big)\int_x^y \max_{j\in I_{n;p}}\|\mathfrak{f}_{n,j}(u,\boldsymbol{\epsilon}_{j-1})\|_{q/2}^{q/2}\,du\\
&\le\ 2^{q/2-1}\kappa_2^{q/2}C\,(y-x)^{(q-2)/2}\big(c_n\sqrt{nb_n}\big)^{-q}\big(1+\mathcal{O}((nb_n)^{-q/2})\big)\\
&\quad\cdot\int_x^y \max_{j\in I_{n;p}}\|\mathfrak{f}_{n,j}(u,\boldsymbol{\epsilon}_{j-1})\|_{q/2}^{q/2}\,du.
\end{aligned}
$$

Along with (1.82) and (1.84) this implies (1.81) with $C_{2,q} := 2^{q-2}\kappa_2^{q/2}CC_q^q$. □

Lemma 1.4.15 *There exists a constant $C_q>0$ such that for any $x\in\mathbb{R}$ and $n\in\mathbb{N}$*

$$
\|H_{p,n}(x)\|_q^q \le C_q^q\big(c_n\sqrt{nb_n}\big)^q\Big\{\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\mathbb{P}[X_{n,i}\le x]^{2/q}\big(1-\mathbb{P}[X_{n,i}\le x]\big)^{2/q}\Big\}^{q/2}. \tag{1.85}
$$

In particular, if conditions (A1)–(A3) hold true, $\sup_{n\in\mathbb{N}}\sum_{k=1}^\infty 2^{kq\lambda}\|H_{p,n}(2^k)\|_q^q<\infty$.

Proof For any $x\in\mathbb{R}$ and $n\in\mathbb{N}$

$$
\begin{aligned}
&\big\|H_{p,n}(x)\big\|_q^q \qquad\qquad (1.86)\\
&=\ \big(c_n\sqrt{nb_n}\big)^q\Big(\Big\|\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\Big(\mathbb{1}_{[X_{n,i},\infty)}(x)-\mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}\big]\Big)\Big\|_q^2\Big)^{q/2}\\
&\le\ \big(c_n\sqrt{nb_n}\big)^q\Big(C_q^2\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\big\|\mathbb{1}_{[X_{n,i},\infty)}(x)-\mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}\big]\big\|_q^2\Big)^{q/2}\\
&\le\ C_q^q\,\big(c_n\sqrt{nb_n}\big)^q\Big(\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\mathbb{E}\Big[\big|\mathbb{1}_{[X_{n,i},\infty)}(x)-\mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}\big]\big|^2\Big]^{2/q}\Big)^{q/2},
\end{aligned}
$$

where we used Burkholder's inequality in the form of part (iii) of Corollary 1.4.5 for the first $\le$ (note that $(\kappa(\frac{i-i_{p,n}}{nb_n})(\mathbb{1}_{\{x<X_{n,i}\le y\}}-\mathbb{E}\big[\mathbb{1}_{\{x<X_{n,i}\le y\}}|\boldsymbol{\epsilon}_{i-1}\big]))_{i=1,\dots,n}$ is a martingale difference sequence) and $|\mathbb{1}_{[X_{n,i},\infty)}(x)-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}]|\le 1$ for the second $\le$. By the tower property of the conditional expectation and Jensen's inequality we further obtain

$$
\begin{aligned}
&\mathbb{E}\Big[\big|\mathbb{1}_{[X_{n,i},\infty)}(x)-\mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}\big]\big|^2\Big]\\
&=\ \mathbb{E}\Big[\mathbb{E}\big[\mathbb{1}^2_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}\big]-\mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}\big]^2\Big]\\
&\le\ \mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x)\big]-\mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x)\big]^2\ =\ \mathbb{P}[X_{n,i}\le x]\,\big(1-\mathbb{P}[X_{n,i}\le x]\big).
\end{aligned}
$$

In view of (1.86), this proves (1.85).

For the proof of the second assertion, we observe that by (1.85)

$$\begin{aligned}&\sum_{k=1}^{\infty} 2^{kq\lambda} \left\| H_{p,n}(2^k) \right\|_q^q \\ &\le C_q^q \big(c_n \sqrt{nb_n}\big)^q \sum_{k=1}^{\infty} 2^{kq\lambda} \Big\{ \sum_{i=1}^{n} \kappa \Big(\frac{i - i_{p,n}}{nb_n} \Big)^2 \big(1 - \mathbb{P}[X_{n,i} \le 2^k]\big)^{2/q} \Big\}^{q/2} \\ &= C_q^q \big(c_n \sqrt{nb_n}\big)^q \sum_{k=1}^{\infty} 2^{kq\lambda} \Big\{ \sum_{i=1}^{n} \kappa \Big(\frac{i - i_{p,n}}{nb_n} \Big)^2 \Big(\int_{2^k}^{\infty} f_{n,i}(y)\, dy \Big)^{2/q} \Big\}^{q/2}.\end{aligned}$$

Since $M := \sup_{n\in\mathbb{N}} \max_{1\le i\le n} \|f_{n,i}\|_{(\gamma)} < \infty$ by assumption (A3), we have

$$\begin{aligned}&\sum_{k=1}^{\infty} 2^{kq\lambda} \big\| H_{p,n}(2^k) \big\|_q^q \\ &\le M C_q^q \big(c_n \sqrt{nb_n}\big)^q \sum_{k=1}^{\infty} 2^{kq\lambda} \Big\{ \sum_{i=1}^{n} \kappa \Big(\frac{i - i_{p,n}}{nb_n} \Big)^2 \Big\}^{q/2} \int_{2^k}^{\infty} \phi_{-\gamma}(y)\, dy \\ &\le (\gamma - 1)^{-1} M C_q^q \big(c_n \sqrt{nb_n}\big)^q \Big(nb_n \int_{-1}^{1} \kappa(u)^2\, du + \mathcal{O}(1) \Big)^{q/2} \sum_{k=1}^{\infty} 2^{kq\lambda}\, \phi_{-\gamma+1}(2^k) \\ &\le 2^{q/2-1} (\gamma - 1)^{-1} M C_q^q \,\big(c_n nb_n\big)^q \big(\kappa_2 + \mathcal{O}\big((nb_n)^{-q/2}\big)\big) \sum_{k=1}^{\infty} 2^{k(q\lambda - \gamma + 1)}\end{aligned}$$

with $\kappa_2 := \int \kappa(u)^2\, du < \infty$ (recall assumption (A1)). The latter expression is bounded above uniformly in $n \in \mathbb{N}$ since $\sum_{k=1}^{\infty} 2^{k(q\lambda-\gamma+1)} < \infty$ (note that $q\lambda - \gamma + 1 < 0$ by our assumptions on γ, λ, q) and $c_n nb_n = \mathcal{O}(1)$ by (1.12). □

Lemma 1.4.16 *Let $\nu_q := q/2 - \max\{1, q/4\}$ and let assumptions (A1)–(A3) and (A6) be fulfilled. Then there exist constants $C_{1,q}, C_{2,q} > 0$ and $n_* \in \mathbb{N}$ such that for any $x \in \mathbb{R}$, $y > 0$, and $n \ge n_*$*

$$\begin{aligned}&\mathbb{E}\Big[\sup_{0\le s<y} |H_{p,n}(x+s) - H_{p,n}(x)|^q \Big] \\ &\le C_{1,q} (nb_n)^{-\nu_q} \big(1 + \log(nb_n)\big)^q \big(1 + \mathcal{O}\big((nb_n)^{-\max\{1,q/4\}}\big)\big) \int_x^{x+y} \max_{i\in I_{n;p}} f_{n,i}(u)\, du \\ &\quad + \big(C_{2,q} + \mathcal{O}\big((nb_n)^{-q/2}\big)\big) y^{q/2-1} \int_x^{x+y} \max_{i\in I_{n;p}} \|\mathfrak{f}_{n,i}(u, \boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\, du. \qquad (1.87)\end{aligned}$$

Proof Let $d_n := 1 + \lfloor q \log(nb_n)/((q-2)\log(2)) \rfloor$ and $h_n = h_n(y) := y2^{-d_n}$ for $n \in \mathbb{N}$ and $y > 0$; the particular choice of d_n will be used only in the last step of the proof (see (1.90) below). By the monotonicity of the involved indicator functions,

$$\big(\mathbb{1}_{[X_{n,i},\infty)}(x+s) - \mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+s)|\boldsymbol{\epsilon}_{i-1}]\big) - \big(\mathbb{1}_{[X_{n,i},\infty)}(x) - \mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}]\big)$$

$$\begin{aligned}
&\leq \big(\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n+1\rfloor)-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n+1\rfloor)|\boldsymbol{\epsilon}_{i-1}]\big)\\
&\quad -\big(\mathbb{1}_{[X_{n,i},\infty)}(x)-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}]\big)\\
&\quad +\big(\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n+1\rfloor)|\boldsymbol{\epsilon}_{i-1}]-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n\rfloor)|\boldsymbol{\epsilon}_{i-1}]\big)
\end{aligned}$$

and

$$\begin{aligned}
&\big(\mathbb{1}_{[X_{n,i},\infty)}(x+s)-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+s)|\boldsymbol{\epsilon}_{i-1}]\big)-\big(\mathbb{1}_{[X_{n,i},\infty)}(x)-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}]\big)\\
&\geq \big(\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n\rfloor)-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n\rfloor)|\boldsymbol{\epsilon}_{i-1}]\big)\\
&\quad -\big(\mathbb{1}_{[X_{n,i},\infty)}(x)-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}]\big)\\
&\quad -\big(\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n+1\rfloor)|\boldsymbol{\epsilon}_{i-1}]-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n\rfloor)|\boldsymbol{\epsilon}_{i-1}]\big)
\end{aligned}$$

for any $x\in\mathbb{R}$ and $s\geq 0$. Thus

$$\begin{aligned}
&\big|\big(\mathbb{1}_{[X_{n,i},\infty)}(x+s)-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+s)|\boldsymbol{\epsilon}_{i-1}]\big)-\big(\mathbb{1}_{[X_{n,i},\infty)}(x)-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}]\big)\big|\\
&\leq \big|\big(\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n+1\rfloor)-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n+1\rfloor)|\boldsymbol{\epsilon}_{i-1}]\big)\\
&\quad -\big(\mathbb{1}_{[X_{n,i},\infty)}(x)-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}]\big)\big|\\
&\quad +\big|\big(\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n\rfloor)-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n\rfloor)|\boldsymbol{\epsilon}_{i-1}]\big)\\
&\quad -\big(\mathbb{1}_{[X_{n,i},\infty)}(x)-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x)|\boldsymbol{\epsilon}_{i-1}]\big)\big|\\
&\quad +\big(\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n+1\rfloor)|\boldsymbol{\epsilon}_{i-1}]-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+h_n\lfloor s/h_n\rfloor)|\boldsymbol{\epsilon}_{i-1}]\big)
\end{aligned}$$

for any $x\in\mathbb{R}$ and $s\geq 0$. It immediately follows that for any $x\in\mathbb{R}$, $y>0$, and $n\in\mathbb{N}$

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{0\leq s<y}\big|H_{p,n}(x+s)-H_{p,n}(x)\big|^q\Big]\\
&\leq 3^{q-1}\mathbb{E}\Big[\sup_{0\leq s<y}\big|H_{p,n}\big(x+h_n\lfloor s/h_n+1\rfloor\big)-H_{p,n}(x)\big|^q\\
&\quad +3^{q-1}\,\mathbb{E}\Big[\sup_{0\leq s<y}\big|H_{p,n}\big(x+h_n\lfloor s/h_n\rfloor\big)-H_{p,n}(x)\big|^q\Big]+3^{q-1}\,\mathbb{E}\Big[\max_{j\leq 2^{d_n}}B_{n,j}(x)^q\Big]\\
&\leq 2\cdot 3^{q-1}\mathbb{E}\Big[\max_{j\leq 2^{d_n}}\big|H_{p,n}\big(x+jh_n\big)-H_{p,n}(x)\big|^q\Big]+3^{q-1}\mathbb{E}\Big[\max_{j\leq 2^{d_n}}B_{n,j}(x)^q\Big]\\
&=: 2\cdot 3^{q-1}S_1(n,x,y)+3^{q-1}S_2(n,x,y) \qquad (1.88)
\end{aligned}$$

with

$$\begin{aligned}
B_{n,j}(x) &:= c_n\sqrt{nb_n}\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\\
&\qquad\cdot\big(\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+jh_n)|\boldsymbol{\epsilon}_{i-1}]-\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+(j-1)h_n)|\boldsymbol{\epsilon}_{i-1}]\big).
\end{aligned}$$

On the one hand, for any $n\geq n_*$ (with n_* as in the proof of Lemma 1.4.7)

$$S_1(n,x,y) \qquad (1.89)$$

$$
\begin{aligned}
&\le \Big(\sum_{r=0}^{d_n}\Big\{\sum_{m=1}^{2^{d_n-r}} \big\|H_{p,n}(x+2^r m h_n)-H_{p,n}(x+2^r(m-1)h_n)\big\|_q^q\Big\}^{1/q}\Big)^q \\
&\le \Big(\sum_{r=0}^{d_n}\Big\{\sum_{m=1}^{2^{d_n-r}} \widetilde{C}_{1,q}\,(nb_n)^{-q/2} \\
&\qquad \cdot\Big(\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^{\min\{q,4\}}\Big(\int_{x+2^r(m-1)h_n}^{x+2^r m h_n} f_{n,i}(u)\,du\Big)^{\min\{1,4/q\}}\Big)^{\max\{1,q/4\}} \\
&\quad + \widetilde{C}_{2,q}\,(2^r h_n)^{(q-2)/2}\big(1+\mathcal{O}((nb_n)^{-q/2})\big)\int_{x+2^r(m-1)h_n}^{x+2^r m h_n}\max_{i\in I_{n;p}}\|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,du\Big\}^{1/q}\Big)^q \\
&\le \Big(\sum_{r=0}^{d_n}\Big\{\widetilde{C}_{1,q}(nb_n)^{-q/2}\Big(\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^{\min\{q,4\}}\Big)^{\max\{1,q/4\}}\int_x^{x+2^{d_n}h_n}\max_{i\in I_{n;p}} f_{n,i}(u)\,du \\
&\quad + \widetilde{C}_{2,q}\,(2^r h_n)^{(q-2)/2}\,\big(1+\mathcal{O}((nb_n)^{-q/2})\big)\int_x^{x+2^{d_n}h_n}\max_{i\in I_{n;p}}\|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,du\Big\}^{1/q}\Big)^q \\
&\le \Big(\sum_{r=0}^{d_n}\Big\{\widetilde{C}_{1,q}\,(nb_n)^{-q/2}\Big(\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^{\min\{q,4\}}\Big)^{\max\{1,q/4\}}\int_x^{x+2^{d_n}h_n}\max_{i\in I_{n;p}} f_{n,i}(u)\,du\Big\}^{1/q} \\
&\quad + \sum_{r=0}^{d_n} 2^{r(q-2)/(2q)} \\
&\qquad \cdot\Big\{\widetilde{C}_{2,q}h_n^{(q-2)/2}\big(1+\mathcal{O}((nb_n)^{-q/2})\big)\int_x^{x+2^{d_n}h_n}\max_{i\in I_{n;p}}\|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,du\Big\}^{1/q}\Big)^q \\
&\le 2^{q-1}d_n^q\,\widetilde{C}_{1,q}\,(nb_n)^{-q/2}\Big(nb_n\int_{-1}^{1}\kappa(u)^{\min\{q,4\}}\,du+\mathcal{O}(1)\Big)^{\max\{1,q/4\}} \\
&\qquad \cdot\int_x^{x+2^{d_n}h_n}\max_{i\in I_{n;p}} f_{n,i}(u)\,du \\
&\quad + 2^{q-1}\widetilde{C}_{2,q}(2^{(q-2)/(2q)}-1)^{-q}\big(2^{d_n+1}h_n\big)^{(q-2)/2}\big(1+\mathcal{O}((nb_n)^{-q/2})\big) \\
&\qquad \cdot\int_x^{x+2^{d_n}h_n}\max_{i\in I_{n;p}}\|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,du \\
&\le 2^{q-1}\,2^{\max\{1,q/4\}-1}\,\widetilde{C}_{1,q}(nb_n)^{-\nu_q}(1+\log(nb_n))^q \\
&\qquad \cdot\big(C_q+\mathcal{O}\big((nb_n)^{-\max\{1,q/4\}}\big)\big)\int_x^{x+y}\max_{i\in I_{n;p}} f_{n,i}(u)\,du \\
&\quad + \widetilde{C}_{2,q}\,\frac{2^{q-1}\,2^{(q-2)/2}}{(2^{(q-2)/(2q)}-1)^q}\,\big(1+\mathcal{O}((nb_n)^{-q/2})\big)\,y^{(q-2)/2}\int_x^{x+y}\max_{i\in I_{n;p}}\|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,du
\end{aligned}
$$

with $C_q := (\int_{-1}^1 \kappa(u)^{\min\{q,4\}}\,du)^{\max\{1,q/4\}} < \infty$ (recall (A1)), where we used Proposition 1(i) in [77] (for the first step), Lemma 1.4.14 (for the second step), and the choice of d_n and h_n (for the last step).

On the other hand, for the second summand in (1.88) we obtain for any $n \geq n_*$

(with n_* as in the proof of Lemma 1.4.7)

$$\begin{aligned}
S_2(n,x,y) &\leq \sum_{j=1}^{2^{d_n}} \mathbb{E}\big[B_{n,j}(x)^q\big] \;=\; \sum_{j=1}^{2^{d_n}} \mathbb{E}\big[B_{n,j}(x)^{q/2} B_{n,j}(x)^{q/2}\big] \\
&\leq (nb_n)^{q/2} \sum_{j=1}^{2^{d_n}} \mathbb{E}\Big[\Big\{c_n \sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\cdot 1\Big\}^{q/2} \cdot \Big\{c_n \sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \\
&\qquad \cdot \big(\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+jh_n)|\boldsymbol{\epsilon}_{i-1}] - \mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+(j-1)h_n)|\boldsymbol{\epsilon}_{i-1}]\big)\Big\}^{q/2}\Big] \\
&= (nb_n)^{q/2} \sum_{j=1}^{2^{d_n}} \mathbb{E}\Big[\Big\{c_n \sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \\
&\qquad \cdot \big(\mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+jh_n)|\boldsymbol{\epsilon}_{i-1}] - \mathbb{E}[\mathbb{1}_{[X_{n,i},\infty)}(x+(j-1)h_n)|\boldsymbol{\epsilon}_{i-1}]\big)\Big\}^{q/2}\Big] \\
&= (c_n nb_n)^{q/2} \sum_{j=1}^{2^{d_n}} \mathbb{E}\Big[\Big(\int_{x+(j-1)h_n}^{x+jh_n} \sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\,du\Big)^{q/2}\Big] \\
&\leq (c_n nb_n)^{q/2} \sum_{j=1}^{2^{d_n}} \mathbb{E}\Big[\Big(\Big\{\int_{x+(j-1)h_n}^{x+jh_n} 1^{q/(q-2)}\,du\Big\}^{(q-2)/q} \\
&\qquad \cdot \Big\{\int_{x+(j-1)h_n}^{x+jh_n} \Big(\sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\Big)^{q/2} du\Big\}^{2/q}\Big)^{q/2}\Big] \\
&= (c_n nb_n)^{q/2} h_n^{(q-2)/2} \sum_{j=1}^{2^{d_n}} \int_{x+(j-1)h_n}^{x+jh_n} \Big\|\sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\Big\|_{q/2}^{q/2} du \\
&\leq (c_n nb_n)^{q/2} h_n^{q/2-1} \int_x^{x+2^{d_n}h_n} \Big(\sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}\Big)^{q/2} du \\
&\leq (c_n nb_n)^{q/2} h_n^{q/2-1} \int_x^{x+2^{d_n}h_n} c_n^{-q/2} \max_{i\in I_{n;p}} \|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\, du \\
&= 2^{-d_n(q/2-1)} (nb_n)^{q/2}\, y^{q/2-1} \int_x^{x+y} \max_{i\in I_{n;p}} \|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\, du \\
&\leq y^{q/2-1} \int_x^{x+y} \max_{i\in I_{n;p}} \|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\, du, \qquad (1.90)
\end{aligned}$$

where we used Hölder's inequality (for the sixth step), Fubini's theorem (for the seventh step), $h_n = y2^{-d_n}$ (for the second-last step), and $2^{-d_n(q/2-1)}(nb_n)^{q/2} \leq 1$ (for the last step). Now (1.88), (1.89), and (1.90) imply (1.87) with $C_{2,q} := 2^q 3^q 2^{(q-2)/2}(2^{(q-2)/(2q)} - 1)^{-q}\widetilde{C}_{2,q}$ and $C_{1,q} := 2^{q+\max\{1,q/4\}-1} 3^q C_q \widetilde{C}_{1,q}$. □

Lemma 1.4.17 *Let assumptions (A1)–(A3), (A6), and (A7) be fulfilled. Then the following assertions hold.*

(i) $\sup_{n\in\mathbb{N}} \mathbb{E}\big[\sup_{x\in\mathbb{R}} |H_{p,n}(x)|^q \phi_{q\lambda}(x)\big] < \infty$.

(ii) $\lim_{w\to\infty} \sup_{n\in\mathbb{N}} \mathbb{P}\big[\sup_{|x|\ge w} |H_{p,n}(x)|\phi_\lambda(x) \ge \delta\big] = 0$ *for all* $\delta > 0$.

(iii) *For every* $\epsilon > 0$, $\delta > 0$, *and* $w > 0$ *there exist a number* $m \in \mathbb{N}$ *and a partition* $-w = x_0 < x_1 < \cdots < x_{m+1} = w$ *of the interval* $[-w, w]$ *such that* $\limsup_{n\to\infty} \mathbb{P}\big[\max_{1\le i\le m} \sup_{x\in[x_i,x_{i+1})} \big|H_{p,n}(x)\phi_\lambda(x) - H_{p,n}(x_i)\phi_\lambda(x_i)\big| \ge \delta\big] \le \epsilon$.

Proof (i): We will only prove $\sup_{n\in\mathbb{N}} \mathbb{E}[\sup_{x\ge 0} |H_{p,n}(x)|^q \phi_{q\lambda}(x)] < \infty$. The analogue with $x \le 0$ can be proven in the same way. We have

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{x\ge 0} |H_{p,n}(x)|^q \phi_{q\lambda}(x)\Big] \qquad (1.91)\\
&\le \sum_{k=1}^{\infty} \mathbb{E}\Big[\sup_{x\in[2^k,2^{k+1})} |H_{p,n}(x)|^q\, \phi_{q\lambda}(x)\Big] + \mathbb{E}\Big[\sup_{x\in[0,2)} |H_{p,n}(x)|^q \phi_{q\lambda}(x)\Big]\\
&\le 2^{q-1}\Big(\sum_{k=1}^{\infty} \mathbb{E}\Big[\sup_{x\in[2^k,2^{k+1})} |H_{p,n}(x) - H_{p,n}(2^k)|^q\, \phi_{q\lambda}(x)\Big]\\
&\quad + \sum_{k=1}^{\infty} \mathbb{E}\Big[\sup_{x\in[2^k,2^{k+1})} |H_{p,n}(2^k)|^q\, \phi_{q\lambda}(x)\Big]\Big) + 3^{q\lambda}\mathbb{E}\Big[\sup_{x\in[0,2)} |H_{p,n}(x)|^q\Big]\\
&=: 2^{q-1}(S_1(n) + S_2(n)) + 3^{q\lambda} S_3(n).
\end{aligned}
$$

It suffices to prove $\sup_{n\in\mathbb{N}} S_i(n) < \infty$ for $i = 1, 2, 3$. For the first summand we have for every $n \ge n_*$ (with n_* as in Lemma 1.4.7)

$$
\begin{aligned}
&S_1(n)\\
&\le \sum_{k=1}^{\infty} \mathbb{E}\Big[\sup_{x\in[2^k,2^{k+1})} |H_{p,n}(x) - H_{p,n}(2^k)|^q (2|x|)^{q\lambda}\Big]\\
&\le 2^{2q\lambda} \sum_{k=1}^{\infty} 2^{kq\lambda}\mathbb{E}\Big[\sup_{x\in[2^k,2^{k+1})} |H_{p,n}(x) - H_{p,n}(2^k)|^q\Big]\\
&\le 2^{2q\lambda}\, C_{1,q}\\
&\quad \cdot (nb_n)^{-\nu_q}(1 + \log(nb_n))^q\big(1 + \mathcal{O}\big((nb_n)^{-\max\{1,q/4\}}\big)\big) \sum_{k=1}^{\infty} 2^{kq\lambda} \int_{2^k}^{2^{k+1}} \max_{i\in I_{n;p}} f_{n,i}(u)\, du\\
&\quad + 2^{2q\lambda}\left(C_{2,q} + \mathcal{O}\big((nb_n)^{-q/2}\big)\right) \sum_{k=1}^{\infty} 2^{kq\lambda} 2^{k(q/2-1)} \int_{2^k}^{2^{k+1}} \max_{i\in I_{n;p}} \|\mathfrak{f}_{n,i}(u, \boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\, du\\
&\le 2^{2q\lambda}\, C_{1,q}(nb_n)^{-\nu_q}(1 + \log(nb_n))^q\big(1 + \mathcal{O}\big((nb_n)^{-\max\{1,q/4\}}\big)\big) \sum_{k=1}^{\infty} 2^{kq\lambda}\, \phi_{-q\lambda}(2^k)\\
&\quad \cdot \int_{2^k}^{2^{k+1}} \max_{i\in I_{n;p}} f_{n,i}(u)\phi_{q\lambda}(u)\, du + 2^{2q\lambda}\left(C_{2,q} + \mathcal{O}\big((nb_n)^{-q/2}\big)\right)
\end{aligned}
$$

$$\cdot \sum_{k=1}^{\infty} 2^{k(q\lambda-1+q/2)}\,\phi_{-q\lambda+1-q/2}(2^k)\int_{2^k}^{2^{k+1}} \max_{i\in I_{n;p}} \|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,\phi_{q\lambda-1+q/2}(u)\,du$$

$$\leq \; 2^{2q\lambda}C_{1,q}(nb_n)^{-\nu_q}(1+\log(nb_n))^q\big(1+\mathcal{O}\big((nb_n)^{-\max\{1,q/4\}}\big)\big)\int_2^{\infty} \max_{i\in I_{n;p}} f_{n,i}(u)\,\phi_{q\lambda}(u)\,du$$

$$+\,2^{2q\lambda}\big(C_{2,q}+\mathcal{O}\big((nb_n)^{-q/2}\big)\big)\int_2^{\infty} \max_{i\in I_{n;p}} \|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,\phi_{q\lambda-1+q/2}(u)\,du$$

$$\leq \; 2^{2q\lambda}MC_{1,q}\,(nb_n)^{-\nu_q}\,(1+\log(nb_n))^q\big(1+\mathcal{O}\big((nb_n)^{-\max\{1,q/4\}}\big)\big)\int_2^{\infty} \phi_{q\lambda-\gamma}(u)\,du$$

$$+\,2^{2q\lambda}\big(C_{2,q}+\mathcal{O}\big((nb_n)^{-q/2}\big)\big)\sup_{n\in\mathbb{N}}\int_{-\infty}^{\infty} \max_{i\in I_{n;p}} \|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,\phi_{q\lambda-1+q/2}(u)\,du$$

with $M := \sup_{n\in\mathbb{N}}\max_{i\in I_{n;p}}\|f_{n,i}\|_{(\gamma)}$, where we used Lemma 1.4.16 (with $x := 2^k$ and $y := 2^k$) for the second step. The latter expression is bounded in n, because $M < \infty$ by (A3), $\lim_{n\to\infty}(nb_n)^{-\nu_q}\log(nb_n)^q = 0$, $\int_2^\infty \phi_{q\lambda-\gamma}(u)\,du < \infty$ (since $q\lambda-\gamma < -1$ by $q < (\gamma-1)/\lambda$) and assumption (A7) holds. Hence, $\sup_{n\in\mathbb{N}} S_1(n) < \infty$.

For the second summand we have $S_2(n) \leq \sum_{k=1}^\infty \mathbb{E}[\sup_{x\in[2^k,2^{k+1})}|H_{p,n}(2^k)|^q(2|x|)^{q\lambda}] \leq 2^{2q\lambda}\sum_{k=1}^\infty 2^{kq\lambda}\|H_{p,n}(2^k)\|_q^q$. This expression is bounded above in n by the second assertion in Lemma 1.4.15.

For the third summand, we obtain by Lemma 1.4.15 (assertion (1.85)) and Lemma 1.4.16 (with $x := 0$ and $y := 2$) that for any $n \geq n_*$

$$S_3(n)$$

$$\leq \; 2^{q-1}\mathbb{E}\Big[\sup_{x\in[0,2)}|H_{p,n}(x)-H_{p,n}(0)|^q\Big] + 2^{q-1}\|H_{p,n}(0)\|_q^q$$

$$\leq \; 2^{q-1}C_{1,q}(nb_n)^{-\nu_q}(1+\log(nb_n))^q\big(1+\mathcal{O}\big((nb_n)^{-\max\{1,q/4\}}\big)\big)\int_0^2 \max_{i\in I_{n;p}} f_{n,i}(u)\,du$$

$$+\,2^{q-1}\big(C_{2,q}+\mathcal{O}\big((nb_n)^{-q/2}\big)\big)2^{q/2-1}\int_0^2 \max_{i\in I_{n;p}}\|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,du$$

$$+\,2^{q-1}C_q^q\big(c_n\sqrt{nb_n}\big)^q\Big(\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\Big)^{q/2}$$

$$\leq \; 2^qMC_{1,q}(nb_n)^{-\nu_q}(1+\log(nb_n))^q\big(1+\mathcal{O}\big((nb_n)^{-\max\{1,q/4\}}\big)\big)$$
$$+\,2^{3q/2-2}N\big(C_{2,q}+\mathcal{O}\big((nb_n)^{-q/2}\big)\big) + 2^{q-1}2^{q/2-1}\,C_q^q\big(c_n nb_n\big)^q\big(\kappa_2^{q/2}+\mathcal{O}\big((nb_n)^{-q/2}\big)\big),$$

with $M := \sup_{n\in\mathbb{N}}\max_{i\in I_{n;p}}\|f_{n,i}\|_\infty$, $N := \sup_{n\in\mathbb{N}}\int_0^2\max_{i\in I_{n;p}}\|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,du$, and $\kappa_2 := \int_{-1}^1\kappa(u)^2\,du$. Now κ_2, M, and N are finite by (A1), (A3), and (A7), respectively. Moreover, $\lim_{n\to\infty}(nb_n)^{-\nu_q}\log(nb_n)^q = 0$ and $\sup_{n\in\mathbb{N}} c_n nb_n < \infty$ by (1.12). Thus $S_3(n)$ is bounded above in n. This finishes the proof of (i).

(ii): We first observe that for $w \geq 2$

$$\mathbb{E}\Big[\sup_{|x|\geq w}|H_{p,n}(x)|^q\phi_{q\lambda}(x)\Big] \tag{1.92}$$

$$
\begin{aligned}
&\leq \; 2^{q-1} \sum_{k=\lfloor \log_2(w) \rfloor}^{\infty} \mathbb{E}\Big[\sup_{x \in [2^k, 2^{k+1}) \cup (-2^{k+1}, -2^k]} |H_{p,n}(x) - H_{p,n}(2^k)|^q \, \phi_{q\lambda}(x) \Big] \\
&\quad + 2^{q-1} \sum_{k=\lfloor \log_2(w) \rfloor}^{\infty} \mathbb{E}\Big[\sup_{x \in [2^k, 2^{k+1}) \cup (-2^{k+1}, -2^k]} |H_{p,n}(2^k)|^q \, \phi_{q\lambda}(x) \Big] \\
&\leq \; 2^{q-1} 2^{2q\lambda} \sum_{k=\lfloor \log_2(w) \rfloor}^{\infty} 2^{kq\lambda} \mathbb{E}\Big[\sup_{x \in [2^k, 2^{k+1}) \cup (-2^{k+1}, -2^k]} |H_{p,n}(x) - H_{p,n}(2^k)|^q \Big] \\
&\quad + 2^{q-1} 2^{2q\lambda} \sum_{k=\lfloor \log_2(w) \rfloor}^{\infty} 2^{kq\lambda} \| H_{p,n}(2^k) \|^q \\
&\leq \; 2^{q-1} 2^{2q\lambda} M C_{1,q} \, (nb_n)^{-\nu_q} \, (1 + \log(nb_n))^q \big(1 + \mathcal{O}\big((nb_n)^{-\max\{1, q/4\}} \big) \big) \\
&\qquad \cdot \int_{\{|u| \geq w\}} \phi_{q\lambda - \gamma}(u) \, du \\
&\quad + 2^{q-1} 2^{2q\lambda} \big(C_{2,q} + \mathcal{O}\big((nb_n)^{-q/2} \big) \big) \int_{\{|u| \geq w\}} \max_{i \in I_{n;p}} \| \mathfrak{f}_{n,i}(u, \boldsymbol{\epsilon}_{i-1}) \|_{q/2}^{q/2} \, \phi_{q\lambda - 1 + q/2}(u) \, du \\
&\quad + 2^{q-1} 2^{2q\lambda} 2^{q/2-1} (\gamma - 1)^{-1} M C_q^q \, \big(c_n n b_n \big)^q \big(\kappa_2 + \mathcal{O}\big((nb_n)^{-q/2} \big) \big) \sum_{k=\lfloor \log_2(w) \rfloor}^{\infty} 2^{k(q\lambda - \gamma + 1)}
\end{aligned}
$$

for some constants $C_{1,q}, \kappa_2, C_q, C_{2,q} > 0$ and $M := \sup_{n \in \mathbb{N}} \max_{1 \leq i \leq n} \| f_{n,i} \|_{(\gamma)}$, which can be shown by using the same arguments as in the proof of (i) and the proof of the second assertion in Lemma 1.4.15. By Markov's inequality and (1.92), we obtain for any $w \geq 2$

$$
\begin{aligned}
&\sup_{n \in \mathbb{N}} \mathbb{P}\Big[\sup_{|x| \geq w} |H_{p,n}(x)| \phi_{\lambda}(x) \geq \delta \Big] \\
&\quad \leq \; \sup_{n \in \mathbb{N}} \frac{1}{\delta^q} \mathbb{E}\Big[\sup_{|x| \geq w} |H_{p,n}(x)|^q \phi_{q\lambda}(x) \Big] \\
&\quad \leq \; \frac{1}{\delta^q} 2^{q-1} 2^{2q\lambda} M C_{1,q} \int_{\{|u| \geq w\}} \phi_{q\lambda - \gamma}(u) \, du \\
&\qquad\quad \cdot \sup_{n \in \mathbb{N}} \Big\{ (nb_n)^{-\nu_q} \, (1 + \log(nb_n))^q \big(1 + \mathcal{O}\big((nb_n)^{-\max\{1, q/4\}} \big) \big) \Big\} \\
&\qquad + \frac{1}{\delta^q} 2^{q-1} 2^{2q\lambda} \\
&\qquad\quad \cdot \sup_{n \in \mathbb{N}} \Big\{ \big(C_{2,q} + \mathcal{O}\big((nb_n)^{-q/2} \big) \big) \int_{\{|u| \geq w\}} \max_{i \in I_{n;p}} \| \mathfrak{f}_{n,i}(u, \boldsymbol{\epsilon}_{i-1}) \|_{q/2}^{q/2} \, \phi_{q\lambda - 1 + q/2}(u) \, du \Big\} \\
&\qquad + \frac{1}{\delta^q} 2^{q-1} 2^{2q\lambda} 2^{q/2-1} (\gamma - 1)^{-1} M C_q^q \sum_{k=\lfloor \log_2(w) \rfloor}^{\infty} 2^{k(q\lambda - \gamma + 1)} \\
&\qquad\quad \cdot \sup_{n \in \mathbb{N}} \Big\{ (c_n n b_n)^q \, \big(\kappa_2 + \mathcal{O}\big((nb_n)^{-q/2} \big) \big) \Big\}. \qquad (1.93)
\end{aligned}
$$

Now $M < \infty$ (by (A3)), $\sup_{n \in \mathbb{N}} (nb_n)^{-\nu_q} \log(nb_n)^q < \infty$, and $\sup_{n \in \mathbb{N}} (c_n n b_n)^q < \infty$ (by (1.12)). Along with (A7), $\int_{\{|u| \geq w\}} \phi_{q\lambda - \gamma}(u) \, du < \infty$, and $\sum_{k=\lfloor \log_2(w) \rfloor}^{\infty} 2^{k(q\lambda - \gamma + 1)} < \infty$

(note that $q\lambda - \gamma + 1 < 0$), we can conclude that the right-hand side of (1.93) converges to 0 as $w \to \infty$.

(iii): Let $\epsilon, \delta, w > 0$ be fixed. For the moment let also $z \in (0,1)$ be fixed (it will be specified later on). By the subadditivity of $\mathbb{P}$, it suffices to show that

$$\limsup_{n\to\infty} \sum_{j=-\lfloor w/z\rfloor-1}^{\lfloor w/z\rfloor} \mathbb{P}\Big[\sup_{x\in[jz,(j+1)z]} \big|H_{p,n}(x)\phi_\lambda(x) - H_{p,n}(jz)\phi_\lambda(jz)\big| \geq \delta\Big] \leq \epsilon. \tag{1.94}$$

By Markov's inequality

$$\begin{aligned}
&\limsup_{n\to\infty} \sum_{j=-\lfloor w/z\rfloor-1}^{\lfloor w/z\rfloor} \mathbb{P}\Big[\sup_{x\in[jz,(j+1)z]} \big|H_{p,n}(x)\phi_\lambda(x) - H_{p,n}(jz)\phi_\lambda(jz)\big| \geq \delta\Big] \qquad (1.95)\\
&\leq \delta^{-q} \limsup_{n\to\infty} \sum_{j=-\lfloor w/z\rfloor-1}^{\lfloor w/z\rfloor} \mathbb{E}\Big[\sup_{x\in[jz,(j+1)z]} \big|H_{p,n}(x)\phi_\lambda(x) - H_{p,n}(jz)\phi_\lambda(jz)\big|^q\Big]\\
&\leq C_{w,\lambda,q}\,\delta^{-q}\\
&\quad \limsup_{n\to\infty} \sum_{j=-\lfloor w/z\rfloor-1}^{\lfloor w/z\rfloor} \Big(\mathbb{E}\Big[\sup_{x\in[jz,(j+1)z]} \big|H_{p,n}(x) - H_{p,n}(jz)\big|^q\Big] + z^q\,\mathbb{E}\Big[\sup_{u\in\mathbb{R}} |H_{p,n}(u)|^q\Big]\Big)
\end{aligned}$$

with $C_{w,\lambda,q} := 2^{q-1}(2+w)^{\lambda q}(\lambda+1)^q$, where we used in the second step that for all $x \in [jz,(j+1)z]$ with $j \in \{-\lfloor w/z\rfloor - 1, \ldots, \lfloor w/z\rfloor\}$ we have

$$\begin{aligned}
&\big|H_{p,n}(x)\phi_\lambda(x) - H_{p,n}(jz)\phi_\lambda(jz)\big|^q\\
&\leq 2^{q-1}\big|\big(H_{p,n}(x) - H_{p,n}(jz)\big)(1+|x|)^\lambda\big|^q + 2^{q-1}\big|H_{p,n}(jz)\big((1+|x|)^\lambda - (1+|jz|)^\lambda\big)\big|^q\\
&\leq 2^{q-1}(2+w)^{\lambda q}\big|H_{p,n}(x) - H_{p,n}(jz)\big|^q + 2^{q-1}\sup_{u\in\mathbb{R}}|H_{p,n}(u)|^q \cdot \Big|\int_{jz}^x \lambda(1+|y|)^{\lambda-1}\,dy\Big|^q\\
&\leq 2^{q-1}(2+w)^{\lambda q}\big|H_{p,n}(x) - H_{p,n}(jz)\big|^q + 2^{q-1}\lambda^q(2+w)^{q(\lambda-1)}z^q\sup_{u\in\mathbb{R}}|H_{p,n}(u)|^q.
\end{aligned}$$

Applying Lemma 1.4.16 with $a := jz$ and $b := z$ (recall that $z \in (0,1)$) yields for any $n \geq n_*$

$$\begin{aligned}
&\sum_{j=-\lfloor w/z\rfloor-1}^{\lfloor w/z\rfloor} \mathbb{E}\Big[\sup_{x\in[jz,(j+1)z]} \big|H_{p,n}(x) - H_{p,n}(jz)\big|^q\Big]\\
&\leq C_{1,q}\,(nb_n)^{-\nu_q}(1+\log(nb_n))^q\big(1+\mathcal{O}\big((nb_n)^{-\max\{1,q/4\}}\big)\big)\\
&\qquad\cdot \sum_{j=-\lfloor w/z\rfloor-1}^{\lfloor w/z\rfloor} \int_{jz}^{(j+1)z} \max_{i\in I_{n;p}} f_{n,i}(u)\,du\\
&\quad + \big(C_{2,q} + \mathcal{O}\big((nb_n)^{-q/2}\big)\big) \sum_{j=-\lfloor w/z\rfloor-1}^{\lfloor w/z\rfloor} z^{q/2-1} \int_{jz}^{(j+1)z} \max_{i\in I_{n;p}} \|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,du
\end{aligned}$$

$$
\begin{aligned}
&\leq\ C_{1,q}\,(nb_n)^{-\nu_q}\,(1+\log(nb_n))^q\big(1+\mathcal{O}\big((nb_n)^{-\max\{1,q/4\}}\big)\big)\int_{-w-1}^{w+1}\max_{i\in I_{n;p}} f_{n,i}(u)\,du\\
&\quad +\big(C_{2,q}+\mathcal{O}\big((nb_n)^{-q/2}\big)\big)\,z^{q/2-1}\sum_{j=-\lfloor w/z\rfloor-1}^{\lfloor w/z\rfloor}\int_{jz}^{(j+1)z}\max_{i\in I_{n;p}}\|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,du\\
&\leq\ (2w+2)MC_{1,q}(nb_n)^{-\nu_q}(1+\log(nb_n))^q\big(1+\mathcal{O}\big((nb_n)^{-\max\{1,q/4\}}\big)\big)\\
&\quad +N\big(C_{2,q}+\mathcal{O}\big((nb_n)^{-q/2}\big)\big)z^{q/2-1}
\end{aligned}
$$

with $N:=\sup_{n\in\mathbb{N}}\int_{-w-1}^{w+1}\max_{i\in I_{n;p}}\|\mathfrak{f}_{n,i}(u,\boldsymbol{\epsilon}_{i-1})\|_{q/2}^{q/2}\,du$ and $M:=\sup_{n\in\mathbb{N}}\max_{i\in I_{n;p}}\|f_{n,i}\|_\infty$. Since the constants N and M are finite by assumptions (A7) and (A3), respectively, and $\lim_{n\to\infty}(nb_n)^{-\nu_q}\log(nb_n)^q=0$, this and (1.95) together imply

$$
\begin{aligned}
&\limsup_{n\to\infty}\sum_{j=-\lfloor w/z\rfloor-1}^{\lfloor w/z\rfloor}\mathbb{P}\Big[\sup_{x\in[jz,(j+1)z]}\big|H_{p,n}(x)\phi_\lambda(x)-H_{p,n}(jz)\phi_\lambda(jz)\big|\geq\delta\Big]\\
&\leq\ 2C_{w,\lambda,q}\frac{1}{\delta^q}\,NC_{2,q}\,z^{q/2-1}+C_{w,\lambda,q}\frac{1}{\delta^q}\Big(2\Big\lfloor\frac{w}{z}\Big\rfloor+2\Big)z^q\sup_{n\in\mathbb{N}}\mathbb{E}\Big[\sup_{u\in\mathbb{R}}|H_{p,n}(u)|^q\Big]\\
&\leq\ C_{1,w,\lambda,q}\frac{1}{\delta^q}\,z^{q/2-1}+C_{2,w,\lambda,q}\frac{1}{\delta^q}\,z^{q-1}
\end{aligned}
$$

with $C_{2,w,\lambda,q}:=C_{w,\lambda,q}(2w+2)\sup_{n\in\mathbb{N}}\mathbb{E}[\sup_{u\in\mathbb{R}}|H_{p,n}(u)|^q]$ (which is finite by (i)) and $C_{1,w,\lambda,q}:=2C_{w,\lambda,q}NC_{2,q}$. Now we may choose z ($\in(0,1)$) so small so that the latter bound is $\leq\epsilon$. This proves (1.94). □

Lemma 1.4.18 *Let assumptions (A1)–(A2) be fulfilled. Then there exist constants $C>0$ and $n_*\in\mathbb{N}$ such that for any $\nu\in\mathbb{R}$, $n\geq n_*$, $A\in\mathcal{B}(\mathbb{R})$, and $(\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R})^{\otimes\mathbb{N}},\mathcal{B}(\mathbb{R}))$-measurable maps $S_{n,i}:\mathbb{R}\times\mathbb{R}^{\mathbb{N}}\to\mathbb{R}$ ($n\in\mathbb{N}$, $i=1,\ldots,n$)*

$$
\begin{aligned}
&\Big\{\int_A\Big\|c_n\sqrt{nb_n}\sum_{i=1}^n\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\big(S_{n,i}(x,\boldsymbol{\epsilon}_i)-\mathbb{E}[S_{n,i}(x,\boldsymbol{\epsilon}_i)]\big)\Big\|_2^2\phi_\nu(x)\,dx\Big\}^{1/2}\\
&\leq\ C\big(1+\mathcal{O}((nb_n)^{-1/2})\big)\sum_{r=0}^\infty\Big\{\max_{i\in I_{n;p}}\int_A\|P_{i-r}(S_{n,i}(x,\boldsymbol{\epsilon}_i))\|_2^2\,\phi_\nu(x)\,dx\Big\}^{1/2}.
\end{aligned}
$$

Proof Since $\lim_{k\to\infty}\mathbb{E}[S_{n,i}(x,\boldsymbol{\epsilon}_i)|\boldsymbol{\epsilon}_{i-k-1}]=\mathbb{E}[S_{n,i}(x,\boldsymbol{\epsilon}_i)]$ $\mathbb{P}$-a.s. by Corollary 11.1.4 in [18], we may write $S_{n,i}(x,\boldsymbol{\epsilon}_i)-\mathbb{E}[S_{n,i}(x,\boldsymbol{\epsilon}_i)]=\sum_{r=0}^\infty P_{i-r}(S_{n,i}(x,\boldsymbol{\epsilon}_i))$. Thus, letting $A_r=A_r(n,\nu):=\{\max_{i\in I_{n;p}}\int_A\|P_{i-r}(S_{n,i}(x,\boldsymbol{\epsilon}_i))\|_2^2\,\phi_\nu(x)\,dx\}^{1/2}$,

$$
\begin{aligned}
&\int_A\Big\|c_n\sqrt{nb_n}\sum_{i=1}^n\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\big(S_{n,i}(x,\boldsymbol{\epsilon}_i)-\mathbb{E}[S_{n,i}(x,\boldsymbol{\epsilon}_i)]\big)\Big\|_2^2\phi_\nu(x)\,dx\\
&=\ c_n^2(nb_n)\int_A\Big\|\sum_{r=0}^\infty\sum_{i=1}^n\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)P_{i-r}\big(S_{n,i}(x,\boldsymbol{\epsilon}_i)\big)\Big\|_2^2\phi_\nu(x)\,dx
\end{aligned}
$$

$$
\begin{aligned}
&= c_n^2(nb_n)\int_A \mathbb{E}\Big[\Big\{\sum_{r=0}^{\infty} A_r^{1/2}\Big(\frac{1}{A_r^{1/2}}\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)P_{i-r}\big(S_{n,i}(x,\boldsymbol{\epsilon}_i)\big)\Big)\Big\}^2\Big]\phi_\nu(x)\,dx \\
&\le c_n^2(nb_n)\int_A \mathbb{E}\Big[\Big(\sum_{r=0}^{\infty} A_r\Big)\Big(\sum_{r=0}^{\infty}\frac{1}{A_r}\Big\{\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)P_{i-r}\big(S_{n,i}(x,\boldsymbol{\epsilon}_i)\big)\Big\}^2\Big)\Big]\phi_\nu(x)\,dx \\
&= c_n^2(nb_n)\Big(\sum_{r=0}^{\infty} A_r\Big)\sum_{r=0}^{\infty}\frac{1}{A_r}\int_A \Big\|\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)P_{i-r}\big(S_{n,i}(x,\boldsymbol{\epsilon}_i)\big)\Big\|_2^2\phi_\nu(x)\,dx \\
&\le C_2^2\, c_n^2(nb_n)\Big(\sum_{r=0}^{\infty} A_r\Big)\sum_{r=0}^{\infty}\frac{1}{A_r}\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\int_A \big\|P_{i-r}\big(S_{n,i}(x,\boldsymbol{\epsilon}_i)\big)\big\|_2^2\phi_\nu(x)\,dx
\end{aligned} \tag{1.96}
$$

for some constant $C_2 > 0$, where used Hölder's inequality for the third step and Burkholder's inequality (in form of Corollary 1.4.5 (iii)) applied to the martingale difference sequence $(\kappa(\frac{i-i_{p,n}}{nb_n})P_{i-r}(S_{n,i}(x,\boldsymbol{\epsilon}_i)))_{i=1,\dots,n}$ for the last step. By assumption (A1) and the definition of A_r

$$
\begin{aligned}
&\sum_{r=0}^{\infty}\frac{1}{A_r}\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\int_A \big\|P_{i-r}\big(S_{n,i}(x,\boldsymbol{\epsilon}_i)\big)\big\|_2^2\,\phi_\nu(x)\,dx \\
&\quad\le \Big(nb_n\int_{-1}^{1}\kappa(u)^2\,du + \mathcal{O}(1)\Big)\sum_{r=0}^{\infty}\frac{1}{A_r}\max_{i\in I_{n;p}}\int_A \big\|P_{i-r}\big(S_{n,i}(x,\boldsymbol{\epsilon}_i)\big)\big\|_2^2\,\phi_\nu(x)\,dx \\
&\quad\le \big(nb_n\kappa_2 + \mathcal{O}(1)\big)\sum_{r=0}^{\infty} A_r
\end{aligned}
$$

for any $n \ge n_*$ (with n_* as in Lemma 1.4.7), where $\kappa_2 := \int_{-1}^{1}\kappa(u)^2\,du < \infty$. Since $c_n^2(nb_n) \le \widetilde{C}(nb_n)^{-1}$ for some constant $\widetilde{C} > 0$ by (1.12), we may conclude in view of (1.96)

$$
\begin{aligned}
&\int_A \Big\|c_n\sqrt{nb_n}\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\big(S_{n,i}(x,\boldsymbol{\epsilon}_i) - \mathbb{E}[S_{n,i}(x,\boldsymbol{\epsilon}_i)]\big)\Big\|_2^2\,\phi_\nu(x)\,dx \\
&\quad\le C^2\big(1+\mathcal{O}((nb_n)^{-1})\big)\Big(\sum_{r=0}^{\infty}\Big\{\max_{i\in I_{n;p}}\int_A \|P_{i-r}(S_{n,i}(x,\boldsymbol{\epsilon}_i))\|_2^2\phi_\nu(x)\,dx\Big\}^{1/2}\Big)^2
\end{aligned}
$$

with $C := (\kappa_2\,\widetilde{C})^{1/2}C_2$. □

Lemma 1.4.19 *Let assumptions (A1)–(A2) and (A8) be fulfilled. Then the following assertions hold.*

(i) $\sup_{n\in\mathbb{N}}\mathbb{E}\big[\sup_{x\in\mathbb{R}}|Q_{p,n}(x)|^2\phi_{2\lambda}(x)\big] < \infty$.

(ii) $\lim_{w\to\infty}\sup_{n\in\mathbb{N}}\mathbb{P}\big[\sup_{|x|\ge w}|Q_{p,n}(x)|\phi_\lambda(x) \ge \delta\big] = 0$ *for all* $\delta > 0$.

(iii) *For every* $\epsilon > 0$, $\delta > 0$, *and* $w > 0$ *there exist a number* $m \in \mathbb{N}$ *and a partition* $-w = x_0 < x_1 < \cdots < x_{m+1} = w$ *of the interval* $[-w, w]$ *such that* $\limsup_{n\to\infty} \mathbb{P}\big[\max_{1\le i\le m} \sup_{x\in[x_i,x_{i+1})} \big|Q_{p,n}(x)\phi_\lambda(x) - Q_{p,n}(x_i)\phi_\lambda(x_i)\big| \ge \delta\big] \le \epsilon$.

Proof (i): Let $w \ge 0$ and $\alpha \in [0,1]$. By Lemma 1 in [78] and Fubini's theorem, we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{|x|\ge w} |Q_{p,n}(x)|^2 \phi_{2\lambda}(x)\Big]^{1/2} \\
&\le \Big\{C_{\lambda,\alpha} \int_{\{|y|\ge w\}} \|Q_{p,n}(y)\|_2^2\, \phi_{2\lambda-\alpha}(y)\, dy + C_{\lambda,\alpha} \int_{\{|y|\ge w\}} \|Q'_{p,n}(y)\|_2^2\, \phi_{2\lambda+\alpha}(y)\, dy\Big\}^{1/2}
\end{aligned}$$

for some positive constant $C_{\lambda,\alpha}$ depending on λ and α. Applying Lemma 1.4.18 with $S_{n,i}(x, \boldsymbol{\epsilon}_i) := \mathfrak{F}_{n,i}(x, \boldsymbol{\epsilon}_{i-1})$, $\nu := 2\lambda - \alpha$ and $S_{n,i}(x, \boldsymbol{\epsilon}_i) := \mathfrak{f}_{n,i}(x, \boldsymbol{\epsilon}_{i-1})$, $\nu := 2\lambda + \alpha$, respectively, yields for every $n \ge n_*$ (with n_* as in Lemma 1.4.7)

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{|x|\ge w} |Q_{p,n}(x)|^2 \phi_{2\lambda}(x)\Big]^{1/2} \\
&\le C_{\lambda,\alpha}^{1/2} C\big(1 + \mathcal{O}((nb_n)^{-1})\big)\Big(\sum_{r=0}^{\infty} \Big\{\max_{i\in I_{n;p}} \int_{\{|y|\ge w\}} \|P_{i-r}(\mathfrak{F}_{n,i}(y, \boldsymbol{\epsilon}_{i-1}))\|_2^2\, \phi_{2\lambda-\alpha}(y)\, dy\Big\}^{1/2} \\
&\quad + \sum_{r=0}^{\infty} \Big\{\max_{i\in I_{n;p}} \int_{\{|y|\ge w\}} \|P_{i-r}(\mathfrak{f}_{n,i}(y, \boldsymbol{\epsilon}_{i-1}))\|_2^2\, \phi_{2\lambda+\alpha}(y)\, dy\Big\}^{1/2}\Big).
\end{aligned}$$

Note that $\|P_{i-r}(\mathfrak{F}_{n,i}(y, \boldsymbol{\epsilon}_{i-1}))\|_2^2 = 0$ and $\|P_{i-r}(\mathfrak{f}_{n,i}(y, \boldsymbol{\epsilon}_{i-1}))\|_2^2 = 0$ for $r = 0$. By the same arguments as in (1.69) we now obtain

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{|x|\ge w} |Q_{p,n}(x)|^2 \phi_{2\lambda}(x)\Big]^{1/2} \\
&\le C_{\lambda,\alpha}^{1/2} C\big(1 + \mathcal{O}((nb_n)^{-1})\big) \\
&\quad \cdot\Big(\sum_{r=1}^{\infty} \Big\{\max_{i\in I_{n;p}} \int_{\{|y|\ge w\}} \|\mathfrak{F}_{n,i}(y, \boldsymbol{\epsilon}_{i-1}) - \mathfrak{F}_{n,i}(y, \boldsymbol{\epsilon}^*_{i-1,i-r})\|_2^2\, \phi_{2\lambda-\alpha}(y)\, dy\Big\}^{1/2} \\
&\quad + \sum_{r=1}^{\infty} \Big\{\max_{i\in I_{n;p}} \int_{\{|y|\ge w\}} \|\mathfrak{f}_{n,i}(y, \boldsymbol{\epsilon}_{i-1}) - \mathfrak{f}_{n,i}(y, \boldsymbol{\epsilon}^*_{i-1,i-r})\|_2^2\, \phi_{2\lambda+\alpha}(y)\, dy\Big\}^{1/2}\Big) \\
&\le C_{\lambda,\alpha}^{1/2} C\big(1 + \mathcal{O}((nb_n)^{-1})\big)\Big(\sup_{n\in\mathbb{N}} \sum_{r=1}^{\infty} \Big\{\max_{i\in I_{n;p}} \int_{\{|y|\ge w\}} \delta^2_{\boldsymbol{\epsilon},r-1;2}(\mathfrak{F}_{n,i}; y)\, \phi_{2\lambda-\alpha}(y)\, dy\Big\}^{1/2} \\
&\quad + \sup_{n\in\mathbb{N}} \sum_{r=1}^{\infty} \Big\{\max_{i\in I_{n;p}} \int_{\{|y|\ge w\}} \delta^2_{\boldsymbol{\epsilon},r-1;2}(\mathfrak{f}_{n,i}; y)\, \phi_{2\lambda+\alpha}(y)\, dy\Big\}^{1/2}\Big). \qquad (1.97)
\end{aligned}$$

In view of assumption (A8), we arrive at (i) by setting $w := 0$.

(ii): By Markov's inequality, we have for any $n \ge n_*$

$$\sup_{n\in\mathbb{N}} \mathbb{P}\Big[\sup_{|x|\ge w} |Q_{p,n}(x)|\phi_\lambda(x) \ge \delta\Big]$$

$$
\begin{aligned}
&\le \sup_{n\in\mathbb{N}} \frac{1}{\delta^2}\,\mathbb{E}\Big[\sup_{|x|\ge w} |Q_{p,n}(x)|^2\phi_{2\lambda}(x)\Big] \\
&\le C_{\lambda,\alpha}C^2\frac{1}{\delta^2}\sup_{n\in\mathbb{N}}\Big\{\big(1+\mathcal{O}((nb_n)^{-1})\big)\sum_{r=1}^{\infty}\Big\{\max_{i\in I_{n;p}}\int_{\{|y|\ge w\}}\delta^2_{\boldsymbol{\epsilon},r-1;2}(\mathfrak{F}_{n,i};y)\,\phi_{2\lambda-\alpha}(y)\,dy\Big\}^{1/2} \\
&\quad + \big(1+\mathcal{O}((nb_n)^{-1})\big)\sum_{r=1}^{\infty}\Big\{\max_{i\in I_{n;p}}\int_{\{|y|\ge w\}}\delta^2_{\boldsymbol{\epsilon},r-1;2}(\mathfrak{f}_{n,i};y)\,\phi_{2\lambda+\alpha}(y)\,dy\Big\}^{1/2}\Big\}^2,
\end{aligned}
$$

where the second step is valid by the same line of arguments as in (i). By assumption (A8) the latter bound tends to 0 as $w\to\infty$.

(iii): Let $\epsilon,\delta,w>0$ be fixed. Analogously to the proof of Lemma 1.4.17(iii), let $z\in(0,1)$ be fixed for the moment. By the subadditivity of $\mathbb{P}$, it suffices to show that

$$
\limsup_{n\to\infty}\sum_{j=-\lfloor w/z\rfloor-1}^{\lfloor w/z\rfloor}\mathbb{P}\Big[\sup_{x\in[jz,(j+1)z]}\big|Q_{p,n}(x)\phi_\lambda(x)-Q_{p,n}(jz)\phi_\lambda(jz)\big|\ge\delta\Big]\le\epsilon. \tag{1.98}
$$

Since for all $x\in[jz,(j+1)z]$ with $j\in\{-\lfloor w/z\rfloor-1,\dots,\lfloor w/z\rfloor\}$ we have

$$
\begin{aligned}
&\big|Q_{p,n}(x)\phi_\lambda(x)-Q_{p,n}(jz)\phi_\lambda(jz)\big|^2 \\
&\le 2\,\big|\big(Q_{p,n}(x)-Q_{p,n}(jz)\big)(1+|x|)^\lambda\big|^2+2\,\big|Q_{p,n}(jz)\big((1+|x|)^\lambda-(1+|jz|)^\lambda\big)\big|^2 \\
&\le 2\,(2+w)^{2\lambda}\Big|\int_{jz}^{x}Q'_{p,n}(y)\,dy\Big|^2+2\Big(\int_{jz}^{x}\lambda(1+|y|)^{\lambda-1}\,dy\Big)^2\sup_{u\in[-w-1,w+1]}\big|Q_{p,n}(u)\big|^2 \\
&\le 2\,(2+w)^{2\lambda}z^2\sup_{u\in[-w-1,w+1]}\big|Q'_{p,n}(u)\big|^2+2\lambda^2\,(2+w)^{2(\lambda-1)}\,z^2\sup_{u\in[-w-1,w+1]}\big|Q_{p,n}(u)\big|^2 \\
&\le C_{w,\lambda}\,z^2\sup_{u\in[-w-1,w+1]}\big|Q'_{p,n}(u)\big|^2+C_{w,\lambda}\,z^2\sup_{u\in[-w-1,w+1]}\big|Q_{p,n}(u)\big|^2
\end{aligned}
$$

(with $C_{w,\lambda}:=2(\lambda+1)^2(2+w)^{2\lambda}$), we obtain by Markov's inequality

$$
\begin{aligned}
&\limsup_{n\to\infty}\sum_{j=-\lfloor w/z\rfloor-1}^{\lfloor w/z\rfloor}\mathbb{P}\Big[\sup_{x\in[jz,(j+1)z]}\big|Q_{p,n}(x)\phi_\lambda(x)-Q_{p,n}(jz)\phi_\lambda(jz)\big|\ge\delta\Big] \qquad (1.99)\\
&\le \delta^{-2}\limsup_{n\to\infty}\sum_{j=-\lfloor w/z\rfloor-1}^{\lfloor w/z\rfloor}\mathbb{E}\Big[\sup_{x\in[jz,(j+1)z]}\big|Q_{p,n}(x)\phi_\lambda(x)-Q_{p,n}(jz)\phi_\lambda(jz)\big|^2\Big] \\
&\le C_{w,\lambda}\frac{1}{\delta^2}\big(2\Big\lfloor\frac{w}{z}\Big\rfloor+2\big)z^2 \\
&\quad \limsup_{n\to\infty}\Big(\mathbb{E}\Big[\sup_{u\in[-w-1,w+1]}\big|Q'_{p,n}(u)\big|^2\Big]+\mathbb{E}\Big[\sup_{u\in[-w-1,w+1]}\big|Q_{p,n}(u)\big|^2\Big]\Big) \\
&\le C_{w,\lambda}\frac{1}{\delta^2}(2w+2)z\limsup_{n\to\infty}\Big(\mathbb{E}\Big[\sup_{u\in[-w-1,w+1]}\big|Q'_{p,n}(u)\big|^2\Big]+\mathbb{E}\Big[\sup_{u\in[-w-1,w+1]}\big|Q_{p,n}(u)\big|^2\Big]\Big).
\end{aligned}
$$

In (i) we proved that $\mathbb{E}\big[\sup_{u\in\mathbb{R}}\big|Q_{p,n}(u)\big|^2\big]=\mathcal{O}(1)$. For the proof of (1.98) it thus remains to show that

$$
\mathbb{E}\Big[\sup_{u\in[-w-1,w+1]}\big|Q'_{p,n}(u)\big|^2\Big]=\mathcal{O}(1) \tag{1.100}
$$

since we can subsequently choose z ($\in (0,1)$) so small so that the expression in (1.99) is less or equal than ϵ.

To prove (1.100), we apply Lemma 4 in [75] with $t := -w-1$ and $\delta := 2w+2$ and Fubini's theorem and obtain for any $n \geq n_*$

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{u\in[-w-1,w+1]} \big|Q'_{p,n}(u)\big|^2\Big] \\
\leq\ & \frac{1}{w+1}\int_{-w-1}^{w+1} \|Q'_{p,n}(y)\|_2^2\, dy + 4(w+1)\int_{-w-1}^{w+1} \|Q''_{p,n}(y)\|_2^2\, dy \\
=\ & \frac{1}{w+1}\int_{-w-1}^{w+1} \|Q'_{p,n}(y)\|_2^2\, \phi_{2\lambda+\alpha}(y)\phi_{-2\lambda-\alpha}(y)\, dy \\
& +4(w+1)\int_{-w-1}^{w+1} \|Q''_{p,n}(y)\|_2^2\, \phi_{-\beta}(y)\phi_{\beta}(y)\, dy \\
\leq\ & \frac{1}{w+1}\int_{-w-1}^{w+1} \|Q'_{p,n}(y)\|_2^2\, \phi_{2\lambda+\alpha}(y)\, dy + 4(w+2)^{\beta+1}\int_{-w-1}^{w+1} \|Q''_{p,n}(y)\|_2^2\, \phi_{-\beta}(y)\, dy.
\end{aligned}$$

By Lemma 1.4.18 with $S_{n,i}(x,\boldsymbol{\epsilon}_i) := \mathfrak{f}_{n,i}(x,\boldsymbol{\epsilon}_{i-1})$, $\nu := 2\lambda+\alpha$ and $S_{n,i}(x,\boldsymbol{\epsilon}_i) := \mathfrak{f}'_{n,i}(x,\boldsymbol{\epsilon}_{i-1})$, $\nu := -\beta$,

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{u\in[-w-1,w+1]} \big|Q'_{p,n}(u)\big|^2\Big] \\
\leq\ & 2\,C^2\left(1+\mathcal{O}((nb_n)^{-1})\right) \\
& \Big(\frac{1}{w+1}\sum_{r=0}^{\infty}\Big\{\max_{i\in I_{n;p}}\int_{-w-1}^{w+1} \big\|P_{i-r}(\mathfrak{f}_{n,i}(y,\boldsymbol{\epsilon}_{i-1}))\big\|_2^2\, \phi_{2\lambda+\alpha}(y)\, dy\Big\}^{1/2} \\
& +4(w+2)^{\beta+1}\sum_{r=0}^{\infty}\Big\{\max_{i\in I_{n;p}}\int_{-w-1}^{w+1} \big\|P_{i-r}(\mathfrak{f}'_{n,i}(y,\boldsymbol{\epsilon}_{i-1}))\big\|_2^2\, \phi_{-\beta}(y)\, dy\Big\}^{1/2}\Big) \\
\leq\ & \Big(\frac{1}{w+1}\sum_{r=0}^{\infty}\Big\{\max_{i\in I_{n;p}}\int_{-w-1}^{w+1} \big\|\mathfrak{f}_{n,i}(y,\boldsymbol{\epsilon}_{i-1}) - \mathfrak{f}_{n,i}(y,\boldsymbol{\epsilon}^*_{i-1,i-r})\big\|_2^2\, \phi_{2\lambda+\alpha}(y)\, dy\Big\}^{1/2} \\
& +4(w+2)^{\beta+1}\sum_{r=0}^{\infty}\Big\{\max_{i\in I_{n;p}}\int_{-w-1}^{w+1} \big\|\mathfrak{f}'_{n,i}(y,\boldsymbol{\epsilon}_{i-1}) - \mathfrak{f}'_{n,i}(y,\boldsymbol{\epsilon}^*_{i-1,i-r})\big\|_2^2\, \phi_{-\beta}(y)\, dy\Big\}^{1/2}\Big) \\
\leq\ & 2C^2\big(1+\mathcal{O}((nb_n)^{-1})\big)\Big(\frac{1}{w+1}\sup_{n\in\mathbb{N}}\sum_{r=1}^{\infty}\Big\{\max_{i\in I_{n;p}}\int_{-\infty}^{\infty} \delta^2_{\boldsymbol{\epsilon},r-1;2}(\mathfrak{f}_{n,i};y)\, \phi_{2\lambda+\alpha}(y)\, dy\Big\}^{1/2} \\
& +4(w+2)^{\beta+1}\sup_{n\in\mathbb{N}}\sum_{r=1}^{\infty}\Big\{\max_{i\in I_{n;p}}\int_{-\infty}^{\infty} \delta^2_{\boldsymbol{\epsilon},r-1;2}(\mathfrak{f}'_{n,i};y)\, \phi_{-\beta}(y)\, dy\Big\}^{1/2}\Big), \qquad (1.101)
\end{aligned}$$

where the second step is valid by the same arguments as in (1.69). Because of our assumptions, the latter bound is finite for fixed w, which implies (1.100). □

1.5 Remaining proofs

1.5.1 Some auxiliary results

Lemma 1.5.1 *Let assumptions (A5) and (A9) be fulfilled. For every $x, y \in \mathbb{R}$ and $i, j \in \mathbb{Z}$*

$$\left|\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\right| \leq C_a \cdot a^{|i-j|/4}$$

for some $a \in [0,1)$, where C_a is a positive constant depending on a.

Proof For every $x, y \in \mathbb{R}$ and $i, j \in \mathbb{Z}$ we have

$$\begin{aligned}
&\left|\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\right| \\
&= \Big|\mathbb{E}\Big[\Big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)) - \mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big]\Big) \\
&\qquad \cdot\Big(\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j)) - \mathbb{E}\big[\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big]\Big)\Big]\Big| \\
&= \Big|\mathbb{E}\Big[\sum_{r=-\infty}^{\max\{i,j\}}\Big(\mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big|\boldsymbol{\epsilon}_r\big] - \mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big|\boldsymbol{\epsilon}_{r-1}\big]\Big) \\
&\qquad \cdot\sum_{s=-\infty}^{\max\{i,j\}}\Big(\mathbb{E}\big[\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big|\boldsymbol{\epsilon}_s\big] - \mathbb{E}\big[\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big|\boldsymbol{\epsilon}_{s-1}\big]\Big)\Big]\Big| \\
&= \Big|\mathbb{E}\Big[\sum_{r=-\infty}^{\max\{i,j\}}\sum_{s=-\infty}^{\max\{i,j\}} P_r\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)P_s\big(\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\Big]\Big|, \qquad (1.102)
\end{aligned}$$

where we used

$$\begin{aligned}
&\mathbb{E}\big[\mathbb{1}_{(-\infty,z]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))\big|\boldsymbol{\epsilon}_t\big] - \mathbb{E}\big[\mathbb{1}_{(-\infty,z]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))\big|\boldsymbol{\epsilon}_{t-1}\big] \qquad (1.103) \\
&\quad = \mathbb{1}_{(-\infty,z]}(G_{j_p}(p,\boldsymbol{\epsilon}_k)) - \mathbb{1}_{(-\infty,z]}(G_{j_p}(p,\boldsymbol{\epsilon}_k)) = 0 \quad \text{for } t > k
\end{aligned}$$

and $\lim_{t\to-\infty}\mathbb{E}[\mathbb{1}_{(-\infty,z]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))|\boldsymbol{\epsilon}_t] = \mathbb{E}[\mathbb{1}_{(-\infty,z]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))]$ $\mathbb{P}$-a.s. (see Theorem 7.4.3 in [32]) for the second-last step, and the definition of the projection operator for the last step. We also have

$$\mathbb{E}\big[P_s\big(\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\big|\boldsymbol{\epsilon}_r\big] = \mathbb{E}\big[\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big|\boldsymbol{\epsilon}_r\big] - \mathbb{E}\big[\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big|\boldsymbol{\epsilon}_r\big] = 0$$

for any $r, s \in \mathbb{Z}$ with $r < s$, which implies

$$\begin{aligned}
&\mathbb{E}\big[P_r\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)P_s\big(\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\big] \\
&\quad = \mathbb{E}\big[\mathbb{E}\big[P_r\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)P_s\big(\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\big|\boldsymbol{\epsilon}_r\big]\big] \\
&\quad = \mathbb{E}\big[P_r\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)\mathbb{E}\big[P_s\big(\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\big|\boldsymbol{\epsilon}_r\big]\big] = 0
\end{aligned}$$

for any $r, s \in \mathbb{Z}$ with $r < s$. Analogously $\mathbb{E}[P_r(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)))P_s(\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j)))] = 0$ for any $r, s \in \mathbb{Z}$ with $s < r$. Therefore we can conclude from (1.102) and (1.103)

$$\left|\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\right|$$

$$
\begin{aligned}
&= \Big| \sum_{r=-\infty}^{\min\{i,j\}} \mathbb{E}\big[P_r\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)P_r\big(\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\big]\Big| \\
&\leq \sum_{r=-\infty}^{\min\{i,j\}} \big\|P_r\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)P_r\big(\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\big\|_1 \\
&\leq \sum_{r=-\infty}^{\min\{i,j\}} \big\|P_r\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)\big\|_2 \big\|P_r\big(\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\big\|_2. \qquad (1.104)
\end{aligned}
$$

By the same line of arguments as in (1.69) and (1.61)–(1.63) we obtain

$$
\begin{aligned}
&\big\|P_r\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)\big\|_2 \\
&\quad\leq \big\|\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}^*_{i,r}))\big\|_2 \\
&\quad\leq \big\|\big\{\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}^*_{i,r}))\big\}\mathbb{1}_{\{|G_{j_p}(p,\boldsymbol{\epsilon}_i)-G_{j_p}(p,\boldsymbol{\epsilon}^*_{i,r})|>(\delta_{\boldsymbol{\epsilon},i-r;2}(G_{j_p}))^{1/2}\}}\big\|_2 \\
&\qquad+\big\|\big\{\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}^*_{i,r}))\big\}\mathbb{1}_{\{|G_{j_p}(p,\boldsymbol{\epsilon}_i)-G_{j_p}(p,\boldsymbol{\epsilon}^*_{i,r})|\leq(\delta_{\boldsymbol{\epsilon},i-r;2}(G_{j_p}))^{1/2}\}}\big\|_2 \\
&\quad\leq (\delta_{\boldsymbol{\epsilon},i-r;2}(G_{j_p}))^{-1/2}\,\|G_{j_p}(p,\boldsymbol{\epsilon}_i) - G_{j_p}(p,\boldsymbol{\epsilon}^*_{i,r})\|_2 + \Big(\int_{x-(\delta_{\boldsymbol{\epsilon},i-r;2}(G_{j_p}))^{1/2}}^{x+(\delta_{\boldsymbol{\epsilon},i-r;2}(G_{j_p}))^{1/2}} f_p(u)\,du\Big)^{1/2} \\
&\quad\leq (\delta_{\boldsymbol{\epsilon},i-r;2}(G_{j_p}))^{1/2} + (2\|f_p\|_\infty)^{1/2}(\delta_{\boldsymbol{\epsilon},i-r;2}(G_{j_p}))^{1/4} \;\leq\; C a^{(i-r)/4} \qquad (1.105)
\end{aligned}
$$

with $C := (2M\widetilde{C})^{1/2} + \widetilde{C}$, where $\delta_{\boldsymbol{\epsilon},i-r;2}(G_{j_p}) = \widetilde{C}a^{i-r/2}$ for some positive constant $\widetilde{C}$ by assumption (A5) and $M := \|f_p\|_\infty < \infty$ by assumption (A9). Therefore, we may conclude in view of (1.104)

$$
\begin{aligned}
&\big|\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\big| \\
&\quad\leq \sum_{r=-\infty}^{\min\{i,j\}} C^2 a^{(i-r)/4} a^{(j-r)/4} \;=\; C^2 a^{(i+j-2\min\{i,j\})/4} \sum_{r=0}^{\infty} a^{r/2} \\
&\quad= C_a\, a^{(\max\{i,j\}-\min\{i,j\})/4} \;=\; C_a\, a^{|i-j|/4}
\end{aligned}
$$

with $C_a := C^2\,(1-\sqrt{a})^{-1}$. □

Lemma 1.5.2 *Let assumptions (A1)–(A4) be fulfilled. Then for any* $x \in \mathbb{R}$

$$
\begin{aligned}
\lim_{n\to\infty} \Big\| c_n\sqrt{nb_n} \sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \Big(&\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)) \qquad (1.106) \\
&-\mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big]\Big)\Big\|_2 \;=\; 0.
\end{aligned}
$$

Proof We will proceed as for (1.72). To this end we regard the argument of the norm on the left-hand side of (1.106) as a telescoping sum, again using Theorem 7.4.3 in [32]. Then

$$
\Big\| c_n\sqrt{nb_n} \sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\Big\{\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))
$$

$$
\begin{aligned}
&-\mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big]\Big\}\Big\|_2 \\
= \;& \Big\| c_n\sqrt{nb_n}\sum_{r=0}^{\infty}\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)P_{i-r}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)\Big\|_2 \\
\le \;& c_n\sqrt{nb_n}\sum_{r=0}^{\infty}\Big\|\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)P_{i-r}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)\Big\|_2 \\
\le \;& C_1\, c_n\sqrt{nb_n} \qquad (1.107) \\
&\cdot\sum_{r=0}^{\infty}\Big(\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2\big\|P_{i-r}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)\big\|_2^2\Big)^{1/2}
\end{aligned}
$$

for some constant $C_1 > 0$, where we applied Burkholder's inequality (in form of Corollary 1.4.5 (iii)) to the martingale difference sequence $(\kappa(\frac{i-i_{p,n}}{nb_n})P_{i-r}(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))))_{i=1,\dots,n}$ in the last step. Below we will show that the following assertion holds true, where we use the same notation as in the proof of Lemma 1.4.7.

($\mathfrak{A}$) There exists a constant $C > 0$ such that for any $n \ge n_*$, $r \in \mathbb{N}$, $x \in \mathbb{R}$ and $i \in I_n^{++}$ ($\subseteq I_{n;p}$) we have $\|P_{i-r}(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)))\|_2^2 \le Ca^{r/4}b_n^{1/4}$.

In view of ($\mathfrak{A}$) and (1.107) we obtain for any $n \ge n_*$

$$
\begin{aligned}
&\Big\| c_n\sqrt{nb_n}\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\Big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)) \\
&-\mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big]\Big)\Big\|_2 \\
&\le \; C_1c_n\sqrt{nb_n}\sum_{r=0}^{\infty}\Big(\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2 Ca^{r/4}b_n^{1/4}\Big)^{1/2} \\
&\le \; C_1C^{1/2}c_n\sqrt{nb_n}\,b_n^{1/8}\sum_{r=0}^{\infty}a^{r/8}\Big(nb_n\int_{-1}^{1}\kappa(u)^2\,du + \mathcal{O}(1)\Big)^{1/2} \\
&\le \; C_ac_nnb_nb_n^{1/8}\big(\sqrt{\kappa_2} + \mathcal{O}((nb_n)^{-1/2})\big)
\end{aligned}
$$

with constants $C_a := C_1C^{1/2}(1-a^{1/8})^{-1}$ and $\kappa_2 := \int_{-1}^{1}\kappa(u)^2\,du$, where the second step is valid by assumption (A1). Since $c_n\,nb_n = \mathcal{O}(1)$ by (1.12), the latter bound tends to 0 as $n \to \infty$. This proves (1.106).

For the proof of ($\mathfrak{A}$), we show that $\|P_{i-r}(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)))\|_2^2$ may be bounded from above in two different ways. On the one hand, we have

$$
\begin{aligned}
&\big\|P_{i-r}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)\big\|_2^2 \\
\le \;& \Big(\big\|P_{i-r}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i))\big)\big\|_2 + \big\|P_{i-r}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)\big\|_2\Big)^2 \\
\le \;& \Big(\big\|\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_{i,i-r}^{*}))\big\|_2 + C_2\,a^{r/4}\Big)^2
\end{aligned}
$$

$$\leq \left(C_3\, a^{r/4} + C_2\, a^{r/4}\right)^2 \;\leq\; C_4\, a^{r/2} \tag{1.108}$$

for some constants $C_2, C_3 > 0$ and $C_4 := (C_2+C_3)^2$. Here, the second step is justified by the same arguments as for (1.70) and (1.105), and the third step is justified by (1.60).

On the other hand, Minkowski's inequality and the conditional Jensen inequality yield

$$\begin{aligned}
&\big\|P_{i-r}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)\big\|_2^2\\
&\quad\leq\; 2\big\|\mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))|\boldsymbol{\epsilon}_{i-r}\big]\big\|_2^2\\
&\qquad+2\big\|\mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))|\boldsymbol{\epsilon}_{i-r-1}\big]\big\|_2^2\\
&\quad\leq\; 4\big\|\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big\|_2^2\\
&\quad=\; 4\big\|\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big\|_1.
\end{aligned} \tag{1.109}$$

Note that for $n \geq n_*$ and $i \in I_n^{++}$ $(\subseteq I_{n;p})$, the random variables $G_{j_p}(i/n,\boldsymbol{\epsilon}_i)$ and $G_{j_p}(p,\boldsymbol{\epsilon}_i)$ have the same distribution as $G_{j_p}(i/n,\boldsymbol{\epsilon}_0)$ and $\xi_p = G_{j_p}(p,\boldsymbol{\epsilon}_0)$, respectively. Hence, for every $n \geq n_*$, $i \in I_n^{++}$, and $x \in \mathbb{R}$

$$\begin{aligned}
&\|\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\|_1\\
&\quad=\; \mathbb{E}\big[|\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_0)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_0))|\big]\\
&\quad\leq\; \mathbb{E}\big[|\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_0)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_0))|\mathbb{1}_{\{|G_{j_p}(i/n,\boldsymbol{\epsilon}_0)-G_{j_p}(p,\boldsymbol{\epsilon}_0)|\leq\delta_n\}}\big]\\
&\qquad+\mathbb{E}\big[|\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_0)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_0))|\mathbb{1}_{\{|G_{j_p}(i/n,\boldsymbol{\epsilon}_0)-G_{j_p}(p,\boldsymbol{\epsilon}_0)|>\delta_n\}}\big]\\
&\quad=:\; S_1(n,i,x) + S_2(n,i,x)
\end{aligned}$$

for any $\delta_n > 0$. For the first summand we have for every $n \geq n_*$, $i \in I_n^{++}$, and $x \in \mathbb{R}$

$$S_1(n,i,x) \;\leq\; \mathbb{P}\big[x-\delta_n \leq G_{j_p}(i/n,\boldsymbol{\epsilon}_0) \leq x+\delta_n\big] \;=\; \int_{x-\delta_n}^{x+\delta_n} f_{n,i}(y)\,dy \;\leq\; C_5\,\delta_n$$

with $C_5 := 2\sup_{n\in\mathbb{N}}\max_{i\in I_{p,n}}\|f_{n,i}\|_\infty < \infty$ by assumption (A3). Thus, $S_1(n,i,x) = \mathcal{O}(\delta_n)$ uniformly in $i \in I_n^{++}$ and $x \in \mathbb{R}$. In exactly the same way we obtain the analogue for every $x \in \mathbb{R}_-$. Hence $S_1(n,i,x) = \mathcal{O}(\delta_n)$ uniformly in $i \in I_n^{++}$ and $x \in \mathbb{R}$. For the second summand we have for any $n \geq n_*$, $i \in I_n^{++}$, and $x \in \mathbb{R}$

$$\begin{aligned}
S_2(n,i,x) \;&\leq\; \mathbb{P}\big[|G_{j_p}(i/n,\boldsymbol{\epsilon}_0) - G_{j_p}(p,\boldsymbol{\epsilon}_0)| > \delta_n\big] \;\leq\; \delta_n^{-1}\,\|G_{j_p}(i/n,\boldsymbol{\epsilon}_0) - G_{j_p}(p,\boldsymbol{\epsilon}_0)\|_1\\
&\leq\; \delta_n^{-1}\,C_6\,|i/n-p| \;\leq\; C_6C_7\,\delta_n^{-1}(b_n)
\end{aligned}$$

for some constant $C_6 > 0$, where we used Markov's inequality and assumption (A4). Recall that for any $i \in I_n^{++}$ we have $|i/n - p| \leq b_n + 1/n \leq C_7\, b_n$ for some positive constant C_7. Therefore $S_2(n,i,x) = \mathcal{O}(b_n\delta_n^{-1})$ uniformly in $i \in I_n^{++}$ and $x \in \mathbb{R}$, and altogether $\|\mathbb{1}_{(-\infty,x]}(X_{n,i}) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\|_1 = \mathcal{O}(\delta_n) + \mathcal{O}(b_n\delta_n^{-1})$ uniformly in $i \in I_n^{++}$ and $x \in \mathbb{R}$. Setting $\delta_n := b_n^{1/2}$ we obtain from (1.109)

$$\big\|P_{i-r}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)\big\|_2^2 \leq C_8 b_n^{1/2} \tag{1.110}$$

for some positive constant C_8. Now (1.108) and (1.110) together imply

$$\begin{aligned} &\big\|P_{i-r}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big)\big\|_2^2 \\ &\quad\le\ \min\{C_4\, a^{r/2}, C_8 b_n^{1/2}\} \\ &\quad\le\ C_4\, a^{r/2}\min\{1, C_4^{-1}C_8\, a^{-r/2}b_n^{1/2}\}\ \le\ C_4^{1/2}C_8^{1/2}\, a^{r/4}b_n^{1/4}, \end{aligned}$$

where we used $\min\{1,|t|\} \le |t|^{1/2}$ in the last step. Setting $C := C_4^{1/2}C_8^{1/2}$ implies $(\mathfrak{A})$. $\square$

Lemma 1.5.3 *If assumptions (A1)–(A3) and (A5) hold and in addition* $\sqrt{nb_n}\|F_{p,n}(\cdot) - \mathbb{E}[\widehat{F}_{p,n}(\cdot)]\|_{(0)} \to 0$*, then* $\lim_{n\to\infty}|\mathbb{E}[\mathcal{E}_{p,n}(x)\mathcal{E}_{p,n}(y)] - \mathbb{E}[\widetilde{\mathcal{E}}_{p,n}(x)\widetilde{\mathcal{E}}_{p,n}(y)]| = 0$ *for any* $x, y \in \mathbb{R}$.

Proof For every $x, y \in \mathbb{R}$ we have

$$\begin{aligned} &\|\mathcal{E}_{p,n}(x)\mathcal{E}_{p,n}(y) - \widetilde{\mathcal{E}}_{p,n}(x)\widetilde{\mathcal{E}}_{p,n}(y)\|_1 \\ &\quad\le\ \|\mathcal{E}_{p,n}(x)(\mathcal{E}_{p,n}(y) - \widetilde{\mathcal{E}}_{p,n}(y))\|_1 + \|\widetilde{\mathcal{E}}_{p,n}(y)(\mathcal{E}_{p,n}(x) - \widetilde{\mathcal{E}}_{p,n}(x))\|_1 \\ &\quad\le\ \|\mathcal{E}_{p,n}(x)\|_2\cdot\|\mathcal{E}_{p,n}(y) - \widetilde{\mathcal{E}}_{p,n}(y)\|_2 + \|\widetilde{\mathcal{E}}_{p,n}(y)\|_2\cdot\|\mathcal{E}_{p,n}(x) - \widetilde{\mathcal{E}}_{p,n}(x)\|_2 \\ &\quad=:\ S_1(n,x)\cdot S_2(n,y) + S_3(n,y)\cdot S_2(n,x) \end{aligned}$$

by Minkowski's inequality and the Cauchy-Schwarz inequality. Since $S_1(n,x) \le \|\mathcal{E}_{p,n}(x) - \widetilde{\mathcal{E}}_{p,n}(x)\|_2 + \|\widetilde{\mathcal{E}}_{p,n}(x)\|_2 = S_2(n,x) + S_3(n,x)$, it thus suffices to show that $\lim_{n\to\infty} S_2(n,x) = 0$ and $S_3(n,x) = \mathcal{O}(1)$ in n for every $x \in \mathbb{R}$. The latter assertion follows directly from $\|\sum_{i=1}^n Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\|_q = \mathcal{O}((nb_n)^{1/2})$ (see (1.78)) with $d := 1$ and $\lambda_1 := 1$ (recall $q \in (2,\infty)$). Moreover, in view of $\|\mathcal{E}_{p,n}(x) - \widetilde{\mathcal{E}}_{p,n}(x)\|_2 = \sqrt{nb_n}|\mathbb{E}[\widehat{F}_{p,n}(x)] - F_{p,n}(x)|$, the former assertion is valid. $\square$

1.5.2 Proof of Lemma 1.2.2

First of all note that the two-sided series on the right-hand side of (1.8) converges absolutely for every $x, y \in \mathbb{R}$. Indeed, by Lemma 1.5.1 we have

$$\sum_{k=-\infty}^{\infty}\big|\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_k)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_0))\big)\big|\ \le\ C_a\sum_{k=-\infty}^{\infty} a^{|k|/4}\ =\ \frac{2C_a}{1-a^{1/4}},$$

taking into account that $a \in [0,1)$.

Now, the mapping $(x,y) \mapsto \mathbb{E}[\widetilde{\mathcal{E}}_{p,n}(x)\widetilde{\mathcal{E}}_{p,n}(y)]$ is the covariance function of the L^2-process $\widetilde{\mathcal{E}}_{p,n}$. Thus it is symmetric and positive semi-definite. As these properties are preserved under the limit, the mapping $(x,y) \mapsto \gamma_p(x,y)$ is symmetric and positive semi-definite, provided we can show that

$$\lim_{n\to\infty}\mathbb{E}\big[\widetilde{\mathcal{E}}_{p,n}(x)\widetilde{\mathcal{E}}_{p,n}(y)\big] = \gamma_p(x,y) \tag{1.111}$$

holds for any $x, y \in \mathbb{R}$. Hence it remains to show (1.111).

Recall that for n sufficiently large the process $\widetilde{\mathcal{E}}_{p,n}$ depends only on the observations associated with those i for which i/n lies in $(p_{j_p}, p_{j_p+1}]$. Therefore, we may and do assume without loss of generality that $\ell = 0$ in the definition of $X_{n,i}$, so that $X_{n,i} = G_{j_p}(i/n, \boldsymbol{\epsilon}_i)$ for some $(\mathcal{B}((0,1]) \otimes \mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}}, \mathcal{B}(\mathbb{R}))$-measurable map $G_{j_p} : (0,1] \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$.

Let $x, y \in \mathbb{R}$ be arbitrary but fixed. Then

$$\begin{aligned}
&\big|\mathbb{E}\big[\widetilde{\mathcal{E}}_{p,n}(x)\widetilde{\mathcal{E}}_{p,n}(y)\big] - \gamma_p(x,y)\big| \\
&\quad \leq \ \Big|c_n^2\, nb_n \sum_{i,j=1}^{n} \kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big)\kappa\Big(\frac{j - i_{p,n}}{nb_n}\Big)\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n, \boldsymbol{\epsilon}_i)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(j/n, \boldsymbol{\epsilon}_j))\big) \\
&\qquad\quad -c_n^2\, nb_n \sum_{i,j=1}^{n} \kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big)\kappa\Big(\frac{j - i_{p,n}}{nb_n}\Big)\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p, \boldsymbol{\epsilon}_i)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p, \boldsymbol{\epsilon}_j))\big)\Big| \\
&\qquad\quad +\Big|c_n^2\, nb_n \sum_{i,j=1}^{n} \kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big)\kappa\Big(\frac{j - i_{p,n}}{nb_n}\Big)\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p, \boldsymbol{\epsilon}_i)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p, \boldsymbol{\epsilon}_j))\big) \\
&\qquad\qquad -\gamma_p(x,y)\Big| \\
&\quad =: \ S_1(n,x,y) + S_2(n,x,y).
\end{aligned} \tag{1.112}$$

We will now show in two steps that $S_1(n,x,y)$ and $S_2(n,x,y)$ converge to 0 as $n \to \infty$.

Step 1. For the first summand, we have

$$\begin{aligned}
S_1(n,x,y) \ &\leq \ \Big|c_n^2\, nb_n \sum_{i,j=1}^{n} \kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big)\kappa\Big(\frac{j - i_{p,n}}{nb_n}\Big) \\
&\qquad \cdot \mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n, \boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p, \boldsymbol{\epsilon}_i)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(j/n, \boldsymbol{\epsilon}_j))\big)\Big| \\
&\quad + \Big|c_n^2\, nb_n \sum_{i,j=1}^{n} \kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big)\kappa\Big(\frac{j - i_{p,n}}{nb_n}\Big) \\
&\qquad \cdot \mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p, \boldsymbol{\epsilon}_i)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(j/n, \boldsymbol{\epsilon}_j)) - \mathbb{1}_{(-\infty,y]}(G_{j_p}(p, \boldsymbol{\epsilon}_j))\big)\Big| \\
&=: \ S_{1,1}(n,x,y) + S_{1,2}(n,x,y).
\end{aligned}$$

By Hölder's inequality

$$\begin{aligned}
&S_{1,1}(n,x,y) \\
&\quad \leq \ \Big\|c_n\sqrt{nb_n}\sum_{i=1}^{n} \kappa\Big(\frac{i - i_{p,n}}{nb_n}\Big)\Big\{\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n, \boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p, \boldsymbol{\epsilon}_i)) \\
&\qquad\quad -\mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n, \boldsymbol{\epsilon}_i)) - \mathbb{1}_{(-\infty,x]}(G_{j_p}(p, \boldsymbol{\epsilon}_i))\big]\Big\} \\
&\qquad \cdot c_n\sqrt{nb_n}\sum_{j=1}^{n} \kappa\Big(\frac{j - i_{p,n}}{nb_n}\Big)\Big\{\mathbb{1}_{(-\infty,y]}(G_{j_p}(j/n, \boldsymbol{\epsilon}_j)) - \mathbb{E}[\mathbb{1}_{(-\infty,y]}(G_{j_p}(j/n, \boldsymbol{\epsilon}_j))]\Big\}\Big\|_1
\end{aligned}$$

$$
\begin{aligned}
\leq\ & \Big\| c_n\sqrt{nb_n}\sum_{i=1}^{n}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\Big\{\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i))-\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\\
&\quad -\mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(G_{j_p}(i/n,\boldsymbol{\epsilon}_i))-\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i))\big]\Big\}\Big\|_2\\
&\cdot\Big\| c_n\sqrt{nb_n}\sum_{j=1}^{n}\kappa\Big(\frac{j-i_{p,n}}{nb_n}\Big)\Big\{\mathbb{1}_{(-\infty,y]}(G_{j_p}(j/n,\boldsymbol{\epsilon}_j))-\mathbb{E}[\mathbb{1}_{(-\infty,y]}(G_{j_p}(j/n,\boldsymbol{\epsilon}_j))]\Big\}\Big\|_2\\
=:\ & S_{1,1,1}(n,x,y)\cdot S_{1,1,2}(n,x,y)
\end{aligned}
$$

On the one hand, the factor $S_{1,1,2}(n,x,y)$ is bounded above in n, which follows from $\|\sum_{i=1}^{n}Y_{n,i}(\boldsymbol{x},\boldsymbol{\lambda})\|_q=\mathcal{O}((nb_n)^{1/2})$ (see (1.78)) with $d:=1$ and $\lambda_1:=1$ (recall $q\in(2,\infty)$). On the other hand, the factor $S_{1,1,1}(n,x,y)$ converges to 0 as $n\to\infty$ by Lemma 1.5.2. As a consequence we have $\lim_{n\to\infty}S_{1,1}(n,x,y)=0$. Analogously one can show that $\lim_{n\to\infty}S_{1,2}(n,x,y)=0$. Hence $\lim_{n\to\infty}S_1(n,x,y)=0$.

Step 2. It remains to show that $\lim_{n\to\infty}S_2(n,x,y)=0$. Let $r_n:=-8\log(nb_n)/\log(a)$ and observe

$$
\begin{aligned}
S_2(n,x,y)\ \leq\ & \Big| c_n^2\,nb_n\sum_{i=1}^{n}\sum_{\{1\leq j\leq n;|j-i|>r_n\}}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\kappa\Big(\frac{j-i_{p,n}}{nb_n}\Big)\\
&\quad\cdot\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)),\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\\
&\quad-\kappa_2\sum_{\{k\in\mathbb{Z};|k|>r_n\}}\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_0)),\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))\big)\Big|\\
&+\Big| c_n^2\,nb_n\sum_{i=1}^{n}\sum_{\{1\leq j\leq n;|j-i|\leq r_n\}}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\kappa\Big(\frac{j-i_{p,n}}{nb_n}\Big)\\
&\quad\cdot\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)),\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\\
&\quad-\kappa_2\sum_{\{j\in\mathbb{Z};|k|\leq r_n\}}\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_0)),\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))\big)\Big|\\
=:\ & S_{2,1}(n,x,y)+S_{2,2}(n,x,y).
\end{aligned}
$$

In the remainder we will show that both summands $S_{2,1}(n,x,y)$ and $S_{2,2}(n,x,y)$ converge to 0 as $n\to\infty$.

For the first summand, we obtain by Lemma 1.5.1

$$
\begin{aligned}
& S_{2,1}(n,x,y)\\
\leq\ & \Big| c_n^2 nb_n\sum_{i=1}^{n}\sum_{\{1\leq j\leq n;|j-i|>r_n\}}\kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\kappa\Big(\frac{j-i_{p,n}}{nb_n}\Big)\\
&\cdot\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)),\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\Big|\\
&+\Big|\kappa_2\sum_{\{k\in\mathbb{Z};|k|>r_n\}}\mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_0)),\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))\big)\Big|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \; C_a c_n^2 nb_n \sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \sum_{\{1\leq j\leq n; |j-i|>r_n\}} \kappa\Big(\frac{j-i_{p,n}}{nb_n}\Big) a^{|i-j|/4} + C_a\kappa_2 \sum_{\{k\in\mathbb{Z};|k|>r_n\}} a^{|k|/4} \\
&\leq \; C_a c_n^2 nb_n \sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \sum_{j=1}^n \kappa\Big(\frac{j-i_{p,n}}{nb_n}\Big) a^{r_n/4} + 2C_a\kappa_2 \sum_{k=0}^\infty a^{(r_n+k)/4} \\
&= \; C_a nb_n a^{r_n/4} + 2C_a\kappa_2 a^{r_n/4} \sum_{k=0}^\infty a^{k/4},
\end{aligned}
$$

where the last step is valid by the definition of c_n. Since $a^{r_n/4} = (nb_n)^{-2}$, we obtain

$$
S_{2,1}(n,x,y) \;\leq\; C_a\,(nb_n)^{-1} + (nb_n)^{-2}\,\frac{2\,C_a\kappa_2}{1-a^{1/4}}\,,
$$

and the latter converges to 0 as $n\to\infty$.

For the second summand, we obtain

$$
\begin{aligned}
&S_{2,2}(n,x,y) \\
&\quad\leq \; \Big| c_n^2\, nb_n \sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \sum_{\{1\leq j\leq n;|j-i|\leq r_n\}} \Big\{\kappa\Big(\frac{j-i_{p,n}}{nb_n}\Big) - \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\Big\} \\
&\qquad \cdot \mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\Big| \\
&\qquad + \Big| c_n^2 nb_n \sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2 \sum_{j=i-r_n}^{i+r_n} \mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big) \\
&\qquad - \kappa_2 \sum_{\{k\in\mathbb{Z};|k|\leq r_n\}} \mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_0)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))\big)\Big| \\
&\quad =: \; S_{2,2,1}(n,x,y) + S_{2,2,2,}(n,x,y).
\end{aligned}
$$

On the one hand, by Lemma 1.5.1 we obtain

$$
\begin{aligned}
&S_{2,2,1}(n,x,y) \\
&\leq \; C_a \Big| c_n^2\, nb_n \sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \sum_{\{1\leq j\leq n;|j-i|\leq r_n\}} \Big\{\kappa\Big(\frac{j-i_{p,n}}{nb_n}\Big) - \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\Big\}\, a^{|i-j|/4}\Big| \\
&\leq \; C_a\, c_n^2\, nb_n \sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \sum_{\{1\leq j\leq n;|j-i|\leq r_n\}} \Big| \int_{(i-i_{p,n})/(nb_n)}^{(j-i_{p,n})/(nb_n)} \kappa'(u)\,du\Big| a^0 \\
&\leq \; C_a M c_n^2 nb_n \sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \sum_{\{1\leq j\leq n;|j-i|\leq r_n\}} \frac{|j-i|}{nb_n} \\
&\leq \; C_a M \frac{2\,r_n^2}{nb_n} c_n^2\, nb_n \sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big) \;\leq\; C_a M \frac{128}{\log^2(a)} \frac{\log^2(nb_n)}{nb_n} (c_n\, nb_n)
\end{aligned}
$$

with $M := \sup_{y\in\mathbb{R}} |\kappa'(y)| < \infty$ (by assumption (A1)). Since $c_n\, nb_n = \mathcal{O}(1)$ by (1.12) and $\log^2(nb_n)/(nb_n) \to 0$ as $n\to\infty$, we obtain $\lim_{n\to\infty} S_{2,2,1}(n,x,y) = 0$.

On the other hand, by Lemma 1.5.1 and (1.12) we obtain for n sufficiently large

$$
\begin{aligned}
& S_{2,2,2}(n,x,y) \\
= \ & \Big| c_n^2\, nb_n \sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2 \sum_{j=-r_n}^{r_n} \mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_i)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_{i+j}))\big) \\
& \quad - \kappa_2 \sum_{j=-r_n}^{r_n} \mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_0)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\Big| \\
= \ & \Big| \Big\{ c_n^2\, nb_n \sum_{i=1}^{n} \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)^2 - \kappa_2 \Big\} \\
& \quad \cdot \sum_{j=-r_n}^{r_n} \mathbb{C}\mathrm{ov}\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_0)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_j))\big)\Big| \\
\leq \ & C_a \Big| c_n^2\, nb_n \big(nb_n \int_{-1}^{1} \kappa(u)^2\, du + \mathcal{O}(1)\big) - \kappa_2 \Big| \sum_{j=-r_n}^{r_n} a^{|j|/4} \\
\leq \ & 2\, C_a \Big| \big(c_n^2\, (nb_n)^2 - 1\big)\kappa_2 + \mathcal{O}\big((nb_n)^{-1}\big)\Big| \frac{1-a^{(r_n+1)/4}}{1-a^{1/4}} \\
\leq \ & 2\, C_a \Big| \Big(\Big(\frac{nb_n}{nb_n+\mathcal{O}(1)}\Big)^2 - 1\Big)\kappa_2 + \mathcal{O}\big((nb_n)^{-1}\big)\Big| \frac{1-a^{1/4}(nb_n)^{-2}}{1-a^{1/4}} \\
\leq \ & \frac{2\, C_a}{1-a^{1/4}} \Big| \Big(\big(1+\mathcal{O}((nb_n)^{-1})\big)^{-2} - 1\Big)\kappa_2 + \mathcal{O}\big((nb_n)^{-1}\big)\Big|,
\end{aligned}
$$

where we used in the second-last step that $c_n = (nb_n \int \kappa(u)\, du + \mathcal{O}(1))^{-1}$ under (A1). The latter bound converges to 0 as $n \to \infty$. Hence $\lim_{n\to\infty} S_2(n,x,y) = 0$. □

1.5.3 Proof of Lemma 1.2.3

Any centered Gaussian process with covariance function γ_p possesses a continuous modification if

$$
\sup_{x\neq y} |\gamma_p(x,y) - \gamma_p(y,y)|/|x-y|^{\beta} < \infty \qquad \text{for some constant } \beta > 0. \tag{1.113}
$$

This is a well-known consequence of the Kolmogorov–Chentsov criterion. For instance, one can combine this criterion with Lemma 1.1 in [58], taking into account that for any centered Gaussian process $(B(t))_{t\in\mathbb{R}}$ with covariance function γ_p we have $\mathbb{E}[(B(t)-B(s))^2] = \gamma_p(t,t) - 2\gamma_p(s,t) + \gamma_p(s,s) \leq |\gamma_p(t,t) - \gamma_p(s,t)| + |\gamma_p(s,s) - \gamma_p(s,t)|$.

Let $x, y \in \mathbb{R}$ and assume without loss of generality $x < y$. Following the same steps as in (1.104) and applying (1.105) yields

$$
\begin{aligned}
& |\gamma_p(y,y) - \gamma_p(x,y)| \\
= \ & \Big|\kappa_2 \sum_{k=-\infty}^{\infty} \mathbb{C}\mathrm{ov}\big(\mathbb{1}_{[x,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k)), \mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_0))\big)\Big|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \kappa_2 \sum_{k=-\infty}^{\infty} \sum_{r=-\infty}^{\min\{k,0\}} \big\| P_r(\mathbb{1}_{[x,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))) \big\|_2 \big\| P_r(\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_0))) \big\|_2 \\
&\leq C_1\kappa_2 \sum_{k=-\infty}^{\infty} \sum_{r=-\infty}^{\min\{k,0\}} \big\| P_r(\mathbb{1}_{[x,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))) \big\|_2 \, a^{-r/4}. \qquad (1.114)
\end{aligned}
$$

On the one hand, we obtain by (1.105)

$$
\begin{aligned}
\big\| P_r(\mathbb{1}_{[x,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))) \big\|_2 &\leq \big\| P_r(\mathbb{1}_{(-\infty,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))) \big\|_2 + \big\| P_r(\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))) \big\|_2 \\
&\leq 2C_2 a^{(k-r)/4}.
\end{aligned}
$$

On the other hand, we may apply the conditional Jensen inequality to obtain

$$
\begin{aligned}
\big\| P_r(\mathbb{1}_{[x,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))) \big\|_2 &\leq \big\| \mathbb{E}\big[\mathbb{1}_{[x,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k)) | \boldsymbol{\epsilon}_r\big] \big\|_2 + \big\| \mathbb{E}\big[\mathbb{1}_{[x,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k)) | \boldsymbol{\epsilon}_{r-1}\big] \big\|_2 \\
&\leq 2\big\| \mathbb{1}_{[x,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k)) \big\|_2 \;=\; 2\mathbb{E}\big[\mathbb{1}_{[x,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))\big]^{1/2} \\
&= 2\mathbb{P}\big[G_{j_p}(p,\boldsymbol{\epsilon}_0) \in [x,y]\big]^{1/2} \;\leq\; 2\Big(\int_x^y f_p(u)\, du \Big)^{1/2} \\
&\leq 2C_3 |y-x|^{1/2}
\end{aligned}
$$

with $C_3 := \|f_p\|_\infty^{1/2}$ (which is finite by (A9)), where we used in the forth step that $G_{j_p}(p,\boldsymbol{\epsilon}_k)$ has the same distribution than $G_{j_p}(p,\boldsymbol{\epsilon}_0)$. Hence

$$
\begin{aligned}
\big\| P_r(\mathbb{1}_{[x,y]}(G_{j_p}(p,\boldsymbol{\epsilon}_k))) \big\|_2 &\leq 2C_3|y-x|^{1/2} \min\Big\{1, \frac{C_2 a^{(k-r)/4}}{C_3|y-x|^{1/2}}\Big\} \\
&\leq 2C_3^{1/2}|y-x|^{1/4} C_2^{1/2} a^{(k-r)/8},
\end{aligned}
$$

where we used the inequality $\min\{1,|t|\} \leq |t|^{1/2}$ for the second step. Together with (1.114) this implies

$$
\begin{aligned}
|\gamma_p(y,y) - \gamma_p(x,y)| &\leq 2\kappa_2 C_1 C_3^{1/2} C_2^{1/2} |y-x|^{1/4} \sum_{k=-\infty}^{\infty} \sum_{r=-\infty}^{\min\{k,0\}} a^{(k-r)/8} a^{-r/8} \\
&\leq 2\kappa_2 C_1 C_3^{1/2} C_2^{1/2} |y-x|^{1/4} \sum_{k=-\infty}^{\infty} a^{(k-2\min\{k,0\})/8} \sum_{r=0}^{\infty} a^{r/4} \\
&\leq \frac{2\kappa_2}{1-a^{1/4}} C_1 C_3^{1/2} C_2^{1/2} |y-x|^{1/4} \sum_{k=-\infty}^{\infty} a^{|k|/8} \;\leq M|y-x|^{1/4}
\end{aligned}
$$

with $M := 4\kappa_2(1-a^{1/8})^{-2} C_1 C_3^{1/2} C_2^{1/2}$. This proves (1.113) for $\beta := 1/4$. □

1.5.4 Proof of Lemma 1.2.6

Below we will show that the following assertion holds true, where we use the same notation as in the proof of Lemma 1.4.7.

($\mathfrak{B}$) There exist constants $C>0$ and $n_0 \geq n_*$ such that for any $n \geq n_0$ and $i \in I_n^{++}$ ($\subseteq I_{n;p}$) we have $\|F_{n,i}-F_{p,n}\|_{(\lambda)} \leq C(b_n)^{q/(q+1)}$ for some $q \in [\lambda,\infty)\cap(0,\infty)$.

Here $F_{n,i}$ denotes the distribution function of $X_{n,i}$. With the help of ($\mathfrak{B}$) one can easily verify that the claim of the lemma holds true. Indeed, let C and n_0 ($\geq n_*$) be as in ($\mathfrak{B}$). Then for any $n \geq n_0$

$$\begin{aligned}
\big\|\mathbb{E}[\widehat{F}_{p,n}]-F_{p,n}\big\|_{(\lambda)} &= \sup_{x\in\mathbb{R}} \big|\mathbb{E}[\widehat{F}_{p,n}(x)]-F_{p,n}(x)\big|\phi_\lambda(x)\\
&= \sup_{x\in\mathbb{R}} \Big|c_n\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\mathbb{E}\big[\mathbb{1}_{[X_{n,i},\infty)}(x)\big]-F_{p,n}(x)\Big|\phi_\lambda(x)\\
&= \sup_{x\in\mathbb{R}} \Big|c_n\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\big(F_{n,i}(x)-F_{p,n}(x)\big)\Big|\phi_\lambda(x)\\
&\leq c_n\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)\|F_{n,i}-F_{p,n}\|_{(\lambda)}\\
&\leq c_n\sum_{i=1}^n \kappa\Big(\frac{i-i_{p,n}}{nb_n}\Big)C(b_n)^{q/(q+1)} \leq C(b_n)^{q/(q+1)}.
\end{aligned}$$

Along with (B2) this gives the claim of the lemma.

It remains to show ($\mathfrak{B}$). For $n \geq n_*$ and $i \in I_n^{++}$ ($\subseteq I_{n;p}$), the random variables $\xi_{n,i} := G_{j_p}(i/n,\boldsymbol{\epsilon}_0)$ and $\xi_{p,n} := G_{j_p}(i_{p,n}/n,\boldsymbol{\epsilon}_0)$ have the same distribution as $X_{n,i}$ and $X_{n,i_{p,n}}$, respectively. Thus for any $x\in\mathbb{R}$ and $i \in I_n^{++}$

$$\begin{aligned}
|F_{n,i}(x)-F_{p,n}(x)| &= \big|\mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(X_{n,i})-\mathbb{1}_{(-\infty,x]}(X_{n,i_{p,n}})\big]\big|\\
&= \big|\mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(\xi_{n,i})-\mathbb{1}_{(-\infty,x]}(\xi_{p,n})\big]\big|\\
&\leq \big|\mathbb{E}\big[\big(\mathbb{1}_{(-\infty,x]}(\xi_{n,i})-\mathbb{1}_{(-\infty,x]}(\xi_{p,n})\big)\mathbb{1}_{\{|\xi_{n,i}-\xi_{p,n}|\leq x\delta_n\}}\big]\big|\\
&\quad +\big|\mathbb{E}\big[\big(\mathbb{1}_{(-\infty,x]}(\xi_{n,i})-\mathbb{1}_{(-\infty,x]}(\xi_{p,n})\big)\mathbb{1}_{\{|\xi_{n,i}-\xi_{p,n}|> x\delta_n\}}\big]\big|\\
&=: S_1(n,i,x)+S_2(n,i,x) \qquad (1.115)
\end{aligned}$$

for any $\delta_n>0$. For the first summand we have for any $x \geq 1$

$$\begin{aligned}
S_1(n,i,x) &\leq \mathbb{P}\big[x-x\delta_n \leq \xi_{n,i} \leq x+x\delta_n\big]\\
&= \mathbb{P}\big[x-x\delta_n \leq G_{j_p}(i_{p,n}/n,\boldsymbol{\epsilon}_0) \leq x+x\delta_n\big] = \int_{x-x\delta_n}^{x+x\delta_n} f_{n,i_{p,n}}(y)\,dy.
\end{aligned}$$

Assuming that δ_n is nonincreasing and tends to 0 as $n\to\infty$, we can choose $n_0\in\mathbb{N}$ with $n_0 \geq n_*$ such that $\delta_n \leq 1/2$ for all $n \geq n_0$. Then, for any $n\geq n_0$ and $x \geq 1$

$$\begin{aligned}
\phi_\lambda(x)S_1(n,x) &\leq \phi_\lambda(x)\int_{x(1-\delta_n)}^{x(1+\delta_n)} f_{n,i_{p,n}}(y)\,dy\\
&\leq \|f_{n,i_{p,n}}\|_{(\gamma)}\,\phi_\lambda(x)\int_{x(1-\delta_n)}^{x(1+\delta_n)} \phi_{-\gamma}(y)\,dy
\end{aligned}$$

$$
\begin{aligned}
&\le C_1\phi_\lambda(x)\,(2x\delta_n) \sup_{y\in(x(1-\delta_n),x(1+\delta_n))} \phi_{-\gamma}(y)\\
&\le 2C_1\delta_n\,\phi_{\lambda+1}(x) \sup_{y\in(x(1-\delta_n),x(1+\delta_n))} \phi_{-\gamma}(y)\\
&= 2C_1\delta_n\,\phi_{\lambda+1}(x)\phi_{-\gamma}(x(1-\delta_n)) \;\le\; 2C_1\delta_n\,\frac{(1+x)^{\lambda+1}}{(1+x/2)^\gamma}\\
&\le 2C_1\delta_n\Big(\frac{1}{1+x/2}+\frac{x}{1+x/2}\Big)^{\lambda+1} \;\le\; C_2\delta_n,
\end{aligned}
$$

where $C_2 := 2C_1 3^{\lambda+1}$ and $C_1 := \sup_{n\in\mathbb{N}}\max_{i\in I_{p,n}}\|f_{n,i}\|_{(\gamma)} < \infty$ (recall (A3)). Thus $\sup_{x\in[1,\infty)} S_1(n,i,x)\phi_\lambda(x) = \mathcal{O}(\delta_n)$ uniformly in $i\in I_n^{++}$. In the same way we obtain the analogue with "$x\in[1,\infty)$" replaced by "$x\in(-\infty,-1]$". Hence we have $\sup_{x\in\mathbb{R}\setminus[-1,1]} S_1(n,i,x)\phi_\lambda(x) = \mathcal{O}(\delta_n)$ uniformly in $i\in I_n^{++}$. For the second summand we have for any $x\in\mathbb{R}\setminus\{0\}$

$$
\begin{aligned}
S_2(n,i,x) &\le \mathbb{P}\big[|G_{j_p}(i/n,\boldsymbol{\epsilon}_0)-G_{j_p}(i_{p,n}/n,\boldsymbol{\epsilon}_0)|>x\delta_n\big]\\
&\le (x\delta_n)^{-q}\big\|G_{j_p}(i/n,\boldsymbol{\epsilon}_0)-G_{j_p}(i_{p,n}/n,\boldsymbol{\epsilon}_0)\big\|_q^q \;\le\; C_3(x\delta_n)^{-q}\,(b_n)^q
\end{aligned}
$$

for some constants $C_3>0$ and $q\in[\lambda,\infty)\cap(0,\infty)$, where we used Markov's inequality and (B4). Thus we have $\phi_\lambda(x)S_2(n,i,x)\le 2^\lambda C_3(b_n\delta_n^{-1})^q$ for any $x\in\mathbb{R}\setminus[-1,1]$ and $i\in I_n^{++}$, and therefore $\sup_{x\in\mathbb{R}\setminus[-1,1]}\phi_\lambda(x)S_2(n,i,x)=\mathcal{O}((b_n\delta_n^{-1})^q)$ uniformly in $i\in I_n^{++}$. Hence $\sup_{x\in\mathbb{R}\setminus[-1,1]}\phi_\lambda(x)|F_{n,i}(x)-F_{p,n}(x)| = \mathcal{O}(\delta_n)+\mathcal{O}((b_n\delta_n^{-1})^q)$ for all $i\in I_n^{++}$.

By the same line of arguments (but with $\le x\delta_n$ and $> x\delta_n$ in (1.115) replaced by $\le\delta_n$ and $>\delta_n$ respectively) we obtain $\sup_{x\in[-1,1]}|F_{n,i}(x)-F_{p,n}(x)| = \mathcal{O}(\delta_n)+\mathcal{O}((b_n\delta_n^{-1})^q)$ for any $i\in I_n^{++}$. Altogether, $\|F_{n,i}-F_{p,n}\|_{(\lambda)}=\mathcal{O}(\delta_n)+\mathcal{O}((b_n\delta_n^{-1})^q)$ for any $i\in I_n^{++}$. Choosing $\delta_n := b_n^{q/(q+1)}$ we arrive at $\|F_{n,i}-F_{p,n}\|_{(\lambda)}=\mathcal{O}(b_n^{q/(q+1)})$ for all $i\in I_n^{++}$. □

1.5.5 Proof of Corollary 1.2.7

First of all we note that in the specific setting of the PLS linear process in Subsection 1.2.3 the shape of $\mathfrak{F}_{n,i}$ is rather explicit. To see this, observe that for $i=1,\ldots,n$, $x\in\mathbb{R}$, and $\boldsymbol{x}=(x_1,x_2,\ldots)\in\mathbb{R}^{\mathbb{N}}$

$$
\begin{aligned}
\mathbb{P}[X_{n,i}\le x\|\boldsymbol{\epsilon}_{i-1}=\boldsymbol{x}] &= \mathbb{P}\Big[\sum_{j=0}^{\ell}\sum_{k=0}^{\infty}a_{j,k}\big(\tfrac{i}{n}\big)\varepsilon_{i-k}\mathbb{1}_{(p_j,p_{j+1}]}\big(\tfrac{i}{n}\big)\le x\Big\|\boldsymbol{\epsilon}_{i-1}=\boldsymbol{x}\Big]\\
&= \mathbb{P}\Big[\varepsilon_i+\sum_{j=0}^{\ell}\sum_{k=1}^{\infty}a_{j,k}\big(\tfrac{i}{n}\big)\varepsilon_{i-k}\mathbb{1}_{(p_j,p_{j+1}]}\big(\tfrac{i}{n}\big)\le x\Big\|\boldsymbol{\epsilon}_{i-1}=\boldsymbol{x}\Big]\\
&= \mathbb{P}\Big[\varepsilon_i+\sum_{j=0}^{\ell}\sum_{k=1}^{\infty}a_{j,k}\big(\tfrac{i}{n}\big)x_k\mathbb{1}_{(p_j,p_{j+1}]}\big(\tfrac{i}{n}\big)\le x\Big]\\
&= \mathbb{P}_{\varepsilon_i}\Big[\Big(-\infty,\,x-\sum_{j=0}^{\ell}\sum_{k=1}^{\infty}a_{j,k}\big(\tfrac{i}{n}\big)x_k\mathbb{1}_{(p_j,p_{j+1}]}\big(\tfrac{i}{n}\big)\Big]\Big]
\end{aligned}
$$

$$= \quad F_\varepsilon\Big(x - \sum_{j=0}^{\ell}\sum_{k=1}^{\infty} a_{j,k}\big(\tfrac{i}{n}\big) x_k \mathbb{1}_{(p_j,p_{j+1}]}\big(\tfrac{i}{n}\big)\Big),$$

where F_ε denotes the distribution function of ε_0. Hence, for $i = 1, \ldots, n$ we may define $\mathfrak{F}_{n,i}$ through

$$\mathfrak{F}_{n,i}(x, \boldsymbol{x}) := F_\varepsilon\big(x - \ell_{n,i}(\boldsymbol{x})\big) \tag{1.116}$$

with $\ell_{n,i}(\boldsymbol{x}) := \sum_{j=0}^{\ell}\sum_{k=1}^{\infty} a_{j,k}(\frac{i}{n}) x_k \mathbb{1}_{(p_j,p_{j+1}]}(\frac{i}{n})$. In particular,

$$\mathfrak{f}_{n,i}(x, \boldsymbol{x}) = f_\varepsilon\big(x - \ell_{n,i}(\boldsymbol{x})\big) \qquad \text{and} \qquad \mathfrak{f}'_{n,i}(x, \boldsymbol{x}) = f'_\varepsilon\big(x - \ell_{n,i}(\boldsymbol{x})\big). \tag{1.117}$$

Moreover,

$$f_{n,i}(x) \;:=\; \mathbb{E}[\mathfrak{f}_{n,i}(x, \boldsymbol{\epsilon}_{i-1})] \;=\; \mathbb{E}[f_\varepsilon(x - \ell_{n,i}(\boldsymbol{\epsilon}_{i-1}))] \;=\; \mathbb{E}[f_\varepsilon(x - Y_{n,i-1})] \tag{1.118}$$

provides a Lebesgue density $f_{n,i}$ of $X_{n,i}$ for any $i = 1, \ldots, n$, where $Y_{n,i-1} := \ell_{n,i}(\boldsymbol{\epsilon}_{i-1})$. With the latter definition of $Y_{n,i-1}$, we also have

$$\mathfrak{F}_{n,i}(x, \boldsymbol{\epsilon}_{i-1}) = F_\varepsilon(x - Y_{n,i-1}). \tag{1.119}$$

Part 1. The first assertion of Corollary 1.2.7 follows from Theorem 1.2.4, if we prove that assertions (A3)–(A9) hold true. We will frequently use the inequality

$$(1 + |u + v|) \leq (1 + |u|)(1 + |v|) \quad \text{ for } u, v \in \mathbb{R}. \tag{1.120}$$

(A3): Using (1.118) and (1.120), we obtain

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} \max_{1\leq i\leq n} \|f_{n,i}\|_{(2\lambda+4)} \\
&\quad= \sup_{n\in\mathbb{N}} \max_{1\leq i\leq n} \sup_{x\in\mathbb{R}} \Big| \int_{\mathbb{R}} f_\varepsilon(x-y)(1+|x|)^{2\lambda+4}\, \mathbb{P}_{Y_{n,i-1}}(dy)\Big| \\
&\quad\leq \sup_{n\in\mathbb{N}} \max_{1\leq i\leq n} \sup_{x\in\mathbb{R}} \Big| \int_{\mathbb{R}} f_\varepsilon(x-y)(1+|x-y|)^{2\lambda+4}(1+|y|)^{2\lambda+4}\, \mathbb{P}_{Y_{n,i-1}}(dy)\Big| \\
&\quad\leq N \sup_{n\in\mathbb{N}} \max_{1\leq i\leq n} \mathbb{E}\big[(1+|Y_{n,i-1}|)^{2\lambda+4}\big] \\
&\quad\leq 2^{2\lambda+3} N \big(1 + \sup_{n\in\mathbb{N}} \max_{1\leq i\leq n} \mathbb{E}\big[|Y_{n,i-1}|^{2\lambda+4}\big]\big),
\end{aligned} \tag{1.121}$$

where $N := \sup_{z\in\mathbb{R}} f_\varepsilon(z)(1+|z|)^{2\lambda+4} \leq \|f_\varepsilon\|_{(\gamma)}$ is finite by assumption (c) and $\gamma > 2\lambda+5$. By Minkowski's inequality we further have

$$\begin{aligned}
\sup_{n\in\mathbb{N}} \max_{1\leq i\leq n} \mathbb{E}\big[|Y_{n,i-1}|^{2\lambda+4}\big] \;&=\; \sup_{n\in\mathbb{N}} \max_{1\leq i\leq n} \Big\| \sum_{j=0}^{\ell}\sum_{k=1}^{\infty} a_{j,k}\big(\tfrac{i}{n}\big)\varepsilon_{i-k}\mathbb{1}_{(p_j,p_{j+1}]}\big(\tfrac{i}{n}\big)\Big\|_{2\lambda+4}^{2\lambda+4} \\
&\leq\; \sup_{n\in\mathbb{N}} \Big(\sum_{j=0}^{\ell}\sum_{k=1}^{\infty} \sup_{\pi\in(p_j,p_{j+1}]} |a_{j,k}(\pi)|\, \|\varepsilon_0\|_{2\lambda+4}\Big)^{2\lambda+4}
\end{aligned}$$

$$= \quad M^{2\lambda+4}\|\varepsilon_0\|_{2\lambda+4}^{2\lambda+4}, \tag{1.122}$$

where $M := \sum_{j=0}^{\ell}\sum_{k=1}^{\infty}\sup_{\pi\in(p_j,p_{j+1}]}|a_{j,k}(\pi)|$ is finite by assumption (a). Moreover,

$$\begin{aligned}\|\varepsilon_0\|_{2\lambda+4} &= \Big(\int |y|^{2\lambda+4} f_\varepsilon(y)\,dy\Big)^{1/(2\lambda+4)} \le \Big(\int \phi_{2\lambda+4}(y) f_\varepsilon(y)\,dy\Big)^{1/(2\lambda+4)}\\ &\le \|f_\varepsilon\|_{(\gamma)}^{1/(2\lambda+4)}\Big(\int \phi_{2\lambda+4-\gamma}(y)\,dy\Big)^{1/(2\lambda+4)}. \end{aligned} \tag{1.123}$$

By assumption (c) and $2\lambda+4-\gamma<-1$ (recall $\gamma>2\lambda+5$), the latter bound is finite. Together with (1.121) and (1.122), this proves (A3) (with $\gamma := 2\lambda+4$).

(A4): For any $\pi,\pi'\in(p_{j_p},p_{j_{p+1}}]$ with $\pi\le\pi'$ we have

$$\begin{aligned}\|G_{j_p}(\pi,\boldsymbol{\epsilon}_0)-G_{j_p}(\pi',\boldsymbol{\epsilon}_0)\|_1 &= \Big\|\sum_{k=0}^{\infty}\big(a_{j_p,k}(\pi)-a_{j_p,k}(\pi')\big)\varepsilon_{-k}\Big\|_1\\ &\le \sum_{k=0}^{\infty}\Big|\int_{\pi}^{\pi'} a'_{j_p,k}(y)\,dy\Big|\,\|\varepsilon_0\|_1 \le M\,|\pi-\pi'|, \end{aligned} \tag{1.124}$$

where $M := \|\varepsilon_0\|_1\sum_{k=0}^{\infty}\sup_{y\in(p_{j_p},p_{j_{p+1}}]}|a'_{j_p,k}(y)|$ is finite by assumption (b) and (1.123).

(A5): Assertion (A5) follows directly from assumption (a), (1.123), and Example 1.2.1.

(A6): Assertion (A6) is an immediate consequence of (1.116)–(1.117), because f_ε was assumed to be continuously differentiable.

(A7): Let $q_\lambda := (4\lambda+6)/(2\lambda+1)$ so that $\lambda q_\lambda - 1 + q_\lambda/2 = 2\lambda+2$. By (1.117) and (1.118),

$$\begin{aligned}&\sup_{n\in\mathbb{N}}\int_{\{|x|\ge w\}}\max_{i\in I_{n;p}}\|\mathfrak{f}_{n,i}(x,\boldsymbol{\epsilon}_{i-1})\|_{q_\lambda/2}^{q_\lambda/2}\,\phi_{\lambda q_\lambda-1+q_\lambda/2}(x)\,dx\\ &\quad\le \sup_{n\in\mathbb{N}}\int_{\{|x|\ge w\}}\max_{i\in I_{n;p}}\mathbb{E}\big[|f_\varepsilon(x-Y_{n,i-1})|\big]\,\phi_{2\lambda+2}(x)\,dx\,\sup_{z\in\mathbb{R}}|f_\varepsilon(z)|^{q_\lambda/2-1}\\ &\quad\le M^{q_\lambda/2-1}\sup_{n\in\mathbb{N}}\int_{\{|x|\ge w\}}\max_{i\in I_{n;p}}|f_{n,i}(x)|\,\phi_{2\lambda+2}(x)\,dx\\ &\quad\le 2^{2\lambda}M^{q_\lambda/2-1}N\int_{\{|x|\ge w\}}\phi_{-2}(x)\,dx,\end{aligned}$$

where $M := \sup_{z\in\mathbb{R}}|f_\varepsilon(z)|$ and $N := \sup_{n\in\mathbb{N}}\max_{i\in I_{n;p}}\|f_{n,i}\|_{(2\lambda+4)}$ are finite by (c) and (1.121), respectively. Since $\int_{-\infty}^{\infty}\phi_{-2}(x)\,dx<\infty$, the latter bound is finite for $w=0$ and converges to 0 as $w\to\infty$.

(A8): For the first, second and third assertion of (A8) it suffices to show that the following conditions ($\mathfrak{C}$), ($\mathfrak{D}$) and ($\mathfrak{E}$) (respectively) are satisfied for some constants $C_1,C_2,C_3\in(0,\infty)$.

($\mathfrak{C}$) $\int_{\{|x|\ge w\}}\big(\frac{\partial}{\partial u}F_\varepsilon(x-u)\big)^2\phi_{2\lambda}(x)\,dx \le C_1\,\phi_{2\lambda+2}(u)\int_{\{|x|\ge w\}}\phi_{-2}(x)\,dx.$

($\mathfrak{D}$) $\int_{\{|x|\geq w\}} \big(\frac{\partial}{\partial u} f_\varepsilon(x-u)\big)^2\, \phi_{2\lambda}(x)\,dx \leq C_2\, \phi_{2\lambda+2}(u) \int_{\{|x|\geq w\}} \phi_{-2}(x)\,dx.$

($\mathfrak{E}$) $\int_{\{|x|\geq w\}} \big(\frac{\partial}{\partial u} f'_\varepsilon(x-u)\big)^2\, \phi_{-2\lambda}(x)\,dx \;\leq\; C_3\, \phi_{|2\lambda-2|}(u) \int_{\{|x|\geq w\}} \phi_{-2}(x)\,dx.$

We will first show that ($\mathfrak{C}$) implies the first assertion of (A8); analogously one can prove that ($\mathfrak{D}$) and ($\mathfrak{E}$) imply the second and the third assertion of (A8) (we omit the corresponding details). Thereafter we will verify that ($\mathfrak{C}$), ($\mathfrak{D}$), and ($\mathfrak{E}$) hold true.

Assume that ($\mathfrak{C}$) holds true. By Fubini's theorem and (1.119), we have

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} \sum_{r=1}^{\infty} \Big\{ \max_{i\in I_{n;p}} \int_{\{|x|\geq w\}} \delta^2_{\boldsymbol{\epsilon},r-1;2}(\mathfrak{F}_{n,i};x)\phi_{2\lambda}(x)\,dx \Big\}^{1/2} \qquad (1.125)\\
&= \sup_{n\in\mathbb{N}} \sum_{r=1}^{\infty} \Big\{ \max_{i\in I_{n;p}} \int_{\{|x|\geq w\}} \mathbb{E}\big[(\mathfrak{F}_{n,i}(x,\boldsymbol{\epsilon}_{i-1}) - \mathfrak{F}_{n,i}(x,\boldsymbol{\epsilon}^*_{i-1,i-r}))^2\big]\phi_{2\lambda}(x)\,dx \Big\}^{1/2}\\
&= \sup_{n\in\mathbb{N}} \sum_{r=1}^{\infty} \Big\{ \max_{i\in I_{n;p}} \mathbb{E}\Big[\int_{\{|x|\geq w\}} \big(F_\varepsilon(x-Y_{n,i-1}) - F_\varepsilon(x-Y^*_{n,i-1;i-r})\big)^2 \phi_{2\lambda}(x)\,dx \Big] \Big\}^{1/2}.
\end{aligned}$$

Letting $\gamma_w(u) := \big\{ \int_{\{|x|\geq w\}} (\frac{\partial}{\partial u} F_\varepsilon(x-u))^2\, \phi_{2\lambda}(x)\,dx \big\}^{1/2}$, we obtain for any $y, y' \in \mathbb{R}$ (assuming without loss of generality $y \leq y'$)

$$\begin{aligned}
&\int_{\{|x|\geq w\}} \big(F_\varepsilon(x-y) - F_\varepsilon(x-y')\big)^2\, \phi_{2\lambda}(x)\,dx\\
&= \int_{\{|x|\geq w\}} \Big(\int_y^{y'} \frac{\partial}{\partial u} F_\varepsilon(x-u)\,du\Big)^2 \phi_{2\lambda}(x)\,dx\\
&= \int_{\{|x|\geq w\}} \Big(\int_y^{y'} \Big(\frac{\partial}{\partial u} F_\varepsilon(x-u)\Big) \frac{\sqrt{\gamma_w(u)}}{\sqrt{\gamma_w(u)}}\,du\Big)^2 \phi_{2\lambda}(x)\,dx\\
&\leq \int_{\{|x|\geq w\}} \Big(\int_y^{y'} \Big(\frac{\partial}{\partial u} F_\varepsilon(x-u)\Big)^2 \frac{1}{\gamma_w(u)}\,du\Big)\Big(\int_y^{y'} \gamma_w(u)\,du\Big)\phi_{2\lambda}(x)\,dx\\
&= \int_y^{y'} \frac{1}{\gamma_w(u)} \int_{\{|x|\geq w\}} \Big(\frac{\partial}{\partial u} F_\varepsilon(x-u)\Big)^2 \phi_{2\lambda}(x)\,dx\,du \int_y^{y'} \gamma_w(u)\,du\\
&= \Big(\int_y^{y'} \gamma_w(u)\,du\Big)^2, \qquad (1.126)
\end{aligned}$$

where we used the Cauchy-Schwarz inequality in the third step and Fubini's theorem in the second-last step. Now, by (1.126) and ($\mathfrak{C}$)

$$\begin{aligned}
&\int_{\{|x|\geq w\}} \big(F_\varepsilon(x-y) - F_\varepsilon(x-y')\big)^2 \phi_{2\lambda}(x)\,dx\\
&\leq \Big(\int_y^{y'} \Big\{C_1 \phi_{2\lambda+2}(u) \int_{\{|x|\geq w\}} \phi_{-2}(x)\,dx\Big\}^{1/2} du\Big)^2\\
&= C_1\Big(\int_y^{y'} \phi_{\lambda+1}(u)\,du\Big)^2 \int_{\{|x|\geq w\}} \phi_{-2}(x)\,dx
\end{aligned}$$

$$
\begin{aligned}
&\leq C_1|y-y'|^2\big(\sup_{y\in[y,y']}\phi_{\lambda+1}(y)\big)^2\int_{\{|x|\geq w\}}\phi_{-2}(x)\,dx\\
&\leq 2C_1|y-y'|^2\big(\phi_{2\lambda+2}(y')+\phi_{2\lambda+2}(y)\big)\int_{\{|x|\geq w\}}\phi_{-2}(x)\,dx.
\end{aligned}
$$

In view of (1.125), we therefore obtain

$$
\begin{aligned}
&\sup_{n\in\mathbb{N}}\sum_{r=1}^{\infty}\Big\{\max_{i\in I_{n;p}}\int_{\{|x|\geq w\}}\delta^2_{\boldsymbol{\epsilon},r-1;2}(\mathfrak{F}_{n,i};x)\phi_{2\lambda}(x)\,dx\Big\}^{1/2}\\
&\leq \sqrt{2C_1}\Big(\int_{\{|x|\geq w\}}\phi_{-2}(x)\,dx\Big)^{1/2}\\
&\quad\cdot\sup_{n\in\mathbb{N}}\sum_{r=1}^{\infty}\Big\{\max_{i\in I_{n;p}}\mathbb{E}\Big[|Y_{n,i-1}-Y^*_{n,i-1;i-r}|^2\big(\phi_{2\lambda+2}(Y_{n,i-1})+\phi_{2\lambda+2}(Y^*_{n,i-1;i-r})\big)\Big]\Big\}^{1/2}\\
&\leq \sqrt{2C_1}\Big(\int_{\{|x|\geq w\}}\phi_{-2}(x)\,dx\Big)^{1/2}\sup_{n\in\mathbb{N}}\sum_{r=1}^{\infty}\Big\{\max_{i\in I_{n;p}}\|(Y_{n,i-1}-Y^*_{n,i-1;i-r})^2\|_{\lambda+2}\\
&\quad\cdot\max_{i\in I_{n;p}}\big\|\big(1+|Y_{n,i-1}|\big)^{2\lambda+2}+\big(1+|Y^*_{n,i-1;i-r}|\big)^{2\lambda+2}\big\|_{(\lambda+2)/(\lambda+1)}\Big\}^{1/2}\\
&\leq \sqrt{2C_1}\Big(\int_{\{|x|\geq w\}}\phi_{-2}(x)\,dx\Big)^{1/2}\sup_{n\in\mathbb{N}}\sum_{r=1}^{\infty}\Big\{\max_{i\in I_{n;p}}\|Y_{n,i-1}-Y^*_{n,i-1;i-r}\|^2_{2\lambda+4}\\
&\quad\cdot 2^{2\lambda+1}\max_{i\in I_{n;p}}\Big(\big(1+\|Y_{n,i-1}\|^{2\lambda+2}_{2\lambda+4}\big)+\big(1+\|Y^*_{n,i-1;i-r}\|^{2\lambda+2}_{2\lambda+4}\big)\Big)\Big\}^{1/2}\\
&\leq 2^{\lambda+1}\sqrt{2C_1}\big(1+N^{(2\lambda+2)/(2\lambda+4)}\big)^{1/2}\Big(\int_{\{|x|\geq w\}}\phi_{-2}(x)\,dx\Big)^{1/2}\\
&\quad\cdot\sup_{n\in\mathbb{N}}\sum_{r=1}^{\infty}\max_{i\in I_{n;p}}\|Y_{n,i-1}-Y^*_{n,i-1;i-r}\|_{2\lambda+4}\\
&\leq 2^{\lambda+1}\sqrt{2C_1}\big(1+N^{(\lambda+1)/(2\lambda+4)}\big)\Big(\int_{\{|x|\geq w\}}\phi_{-2}(x)\,dx\Big)^{1/2}\\
&\quad\cdot\sup_{n\in\mathbb{N}}\sum_{r=1}^{\infty}\max_{i\in I_{n;p}}\Big(\sum_{j=0}^{\ell}\sup_{\pi\in(p_j,p_{j+1}]}|a_{j,r-1}(\pi)|\,\|\varepsilon_{i-r}-\varepsilon^*\|_{2\lambda+4}\mathbb{1}_{(p_j,p_{j+1}]}(i/n)\Big)\\
&\leq 2^{\lambda+2}\sqrt{2C_1}\big(1+N^{(\lambda+1)/(2\lambda+4)}\big)\Big(\int_{\{|x|\geq w\}}\phi_{-2}(x)\,dx\Big)^{1/2}\|\varepsilon_0\|_{2\lambda+4}\\
&\quad\cdot\sum_{r=1}^{\infty}\sup_{\pi\in(p_{j_p},p_{j_p+1}]}|a_{j,r-1}(\pi)|\\
&\leq 2^{\lambda+2}\sqrt{2C_1}\widetilde{C}\big(1+N^{(\lambda+1)/(2\lambda+4)}\big)\Big(\int_{\{|x|\geq w\}}\phi_{-2}(x)\,dx\Big)^{1/2}\|\varepsilon_0\|_{2\lambda+4}\sum_{r=1}^{\infty}a^r\\
&= 2^{\lambda+2}\sqrt{2C_1}\widetilde{C}(1-a)^{-1}\big(1+N^{(\lambda+1)/(2\lambda+4)}\big)\Big(\int_{\{|x|\geq w\}}\phi_{-2}(x)\,dx\Big)^{1/2}\|\varepsilon_0\|_{2\lambda+4},
\end{aligned}
$$

where the second step holds by Hölder's inequality, $N := \sup_{n\in\mathbb{N}} \max_{1\le i\le n} \mathbb{E}[|Y_{n,i-1}|^{2\lambda+4}]$ is finite by (1.122), and $\widetilde{C} \in (0,\infty)$ is chosen such that $\sup_{y\in(p_{j_p},p_{j_p+1}]} |a_{j_p,r-1}(y)| \le \widetilde{C} a^{r-1}$ (recall assumption (a)). Since $\|\varepsilon_0\|_{2\lambda+4} < \infty$ by (1.123) and $\int_{\mathbb{R}} \phi_{-2}(x)\,dx < \infty$, the latter bound is finite for $w = 0$ and converges to 0 as $w \to \infty$. Hence, we have shown that ($\mathfrak{C}$) implies the first assertion of (A8). Analogously one can show that ($\mathfrak{D}$) and ($\mathfrak{E}$) imply the remaining assertions of (A8).

It remains to show ($\mathfrak{C}$), ($\mathfrak{D}$), and ($\mathfrak{E}$). Using (1.120), we obtain

$$\begin{aligned}
&\int_{\{|x|\ge w\}} \Big(\frac{\partial}{\partial u} F_\varepsilon(x-u)\Big)^2 \phi_{2\lambda}(x)\,dx + \int_{\{|x|\ge w\}} \Big(\frac{\partial}{\partial u} f_\varepsilon(x-u)\Big)^2 \phi_{2\lambda}(x)\,dx \\
&\quad= \int_{\{|x|\ge w\}} \big(f_\varepsilon(x-u)^2 + f_\varepsilon'(x-u)^2\big)\phi_{2\lambda+2}(x)\phi_{-2}(x)\,dx \\
&\quad\le \int_{\{|x|\ge w\}} \big(f_\varepsilon(x-u)^2 + f_\varepsilon'(x-u)^2\big)\phi_{2\lambda+2}(x-u)\phi_{2\lambda+2}(u)\phi_{-2}(x)\,dx \\
&\quad\le (C_1+C_2)\,\phi_{2\lambda+2}(u) \int_{\{|x|\ge w\}} \phi_{-2}(x)\,dx,
\end{aligned}$$

where $C_1 := \|f_\varepsilon\|_{(\lambda+1)}^2$ is finite by (1.123), and $C_2 := \|f_\varepsilon'\|_{(\lambda+1)}^2$ is finite by assumption (d). This proves ($\mathfrak{C}$) and ($\mathfrak{D}$). Moreover we also have ($\mathfrak{E}$), because

$$\begin{aligned}
&\int_{\{|x|\ge w\}} \Big(\frac{\partial}{\partial u} f_\varepsilon'(x-u)\Big)^2 \phi_{-2\lambda}(x)\,dx \\
&\quad= \int_{\{|x|\ge w\}} f_\varepsilon''(x-u)^2\,\phi_{2-2\lambda}(x)\phi_{-2}(x)\,dx \\
&\quad\le \int_{\{|x|\ge w\}} f_\varepsilon''(x-u)^2\,\phi_{2-2\lambda}(x-u)\,\phi_{|2\lambda-2|}(u)\phi_{-2}(x)\,dx \\
&\quad\le C_3\,\phi_{|2\lambda-2|}(u) \int_{\{|x|\ge w\}} \phi_{-2}(x)\,dx,
\end{aligned}$$

where $C_3 := \|f_\varepsilon''\|_{(1-\lambda)}^2$ is finite by assumption (d). In the third step we used that by (1.120) we have $\phi_{2-2\lambda}(x) = \phi_{2\lambda-2}(u)(\phi_{2\lambda-2}(x)\phi_{2\lambda-2}(u))^{-1} \le \phi_{2\lambda-2}(u)(\phi_{2\lambda-2}(x-u))^{-1}$ if $2-2\lambda < 0$, and $\phi_{2-2\lambda}(x) \le \phi_{2-2\lambda}(x-u)\phi_{2-2\lambda}(u)$ if $2-2\lambda \ge 0$.

(A9): Analogously to (1.118) we obtain that $f_p(x) := \mathbb{E}[f_\varepsilon(x-Y_p)]$ with $Y_p := \xi_p - \varepsilon_0$ provides a Lebesgue density f_p of ξ_p. Thus, since

$$\|f_p\|_\infty = \sup_{x\in\mathbb{R}} \Big| \int f_\varepsilon(x-y)\,\mathbb{P}_{Y_p}(dy)\Big| \le \sup_{z\in\mathbb{R}} |f_\varepsilon(z)| < \infty$$

by assumption (c), the assertion of (A9) hold true.

Part 2. The second assertion of Corollary 1.2.7 follows directly from Lemma 1.2.6 and the second assertion of Theorem 1.2.4, provided we can show that assumption (B4) holds for $q := 2\lambda + 4$. But analogously to (1.124) we obtain

$$\|G_{j_p}(\pi,\boldsymbol{\epsilon}_0) - G_{j_p}(\pi',\boldsymbol{\epsilon}_0)\|_{2\lambda+4} \le |\pi-\pi'|\,\|\varepsilon_0\|_{2\lambda+4} \sum_{k=0}^{\infty} \sup_{\widetilde{\pi}\in(p_{j_p},p_{j_p+1}]} |a_{j_p,k}'(\widetilde{\pi})|$$

for any $\pi, \pi' \in (p_{j_p}, p_{j_{p+1}}]$. Assertion (B4) is therefore a direct consequence of assumption (b) and (1.123). $\square$

1.5.6 Proof of Lemma 1.2.8

(i): We observe that

$$\begin{aligned}
&\Big\| \sup_{\pi\in[p_j,p_{j+1}]} \sum_{r=0}^{\infty} \Big\{ \Big[\prod_{s=0}^{r} A_j(\pi, \varepsilon_{t-s}) \Big] \bar{b}_j(\pi, \varepsilon_{t-r-1}) \Big\}_{(1)} \Big\|_q \\
&\le \sum_{r=0}^{\infty} \Big\| \sup_{\pi\in[p_j,p_{j+1}]} \Big\{ \Big[\prod_{s=0}^{r} A_j(\pi, \varepsilon_{t-s}) \Big] \bar{b}_j(\pi, \varepsilon_{t-r-1}) \Big\}_{(1)} \Big\|_q \\
&\le \sum_{r=0}^{\infty} \Big\{ \Big|\Big\|\Big[\prod_{s=0}^{r} \sup_{\pi\in[p_j,p_{j+1}]} A_j(\pi, \varepsilon_{t-s}) \Big] \sup_{\pi\in[p_j,p_{j+1}]} \bar{b}_j(\pi, \varepsilon_{t-r-1}) \Big\|\Big|_q \Big\}_{(1)},
\end{aligned}$$

where for any vector v (resp. matrix A) we denote by $|||v|||_q$ (resp. $|||A|||_q$) the vector (resp. matrix) of the entry-wise L^q norms of v (resp. A), and $\sup_\pi$ in front of a vector (resp. matrix) refers to the vector (resp. matrix) obtained by taking entry-wise the supremum over π. The random variables $\{\varepsilon_i\}_{i\in\mathbb{Z}}$ and with that $\sup_{\pi\in[p_j,p_{j+1}]} A_j(\pi, \varepsilon_t), \ldots,$ $\sup_{\pi\in[p_j,p_{j+1}]} A_j(\pi, \varepsilon_{t-r}), \sup_{\pi\in[p_j,p_{j+1}]} \bar{b}_j(\pi, \varepsilon_{t-r-1})$ are independent so that

$$\begin{aligned}
&\Big|\Big\|\Big[\prod_{s=0}^{r} \sup_{\pi\in[p_j,p_{j+1}]} A_j(\pi, \varepsilon_{t-s}) \Big] \sup_{\pi\in[p_j,p_{j+1}]} \bar{b}_j(\pi, \varepsilon_{t-r-1}) \Big\|\Big|_q \\
&= \Big[\prod_{s=0}^{r} ||| \sup_{\pi\in[p_j,p_{j+1}]} A_j(\pi, \varepsilon_{t-s}) |||_q \Big] \Big|\Big\| \sup_{\pi\in[p_j,p_{j+1}]} \bar{b}_j(\pi, \varepsilon_{t-r-1}) \Big\|\Big|_q \\
&= \Big[\prod_{s=0}^{r} \sup_{\pi\in[p_j,p_{j+1}]} A_j(\pi, \|\varepsilon_{t-s}\|_q) \Big] \sup_{\pi\in[p_j,p_{j+1}]} \bar{b}_j(\pi, \|\varepsilon_{t-r-1}\|_q) \\
&= \Big[\sup_{\pi\in[p_j,p_{j+1}]} A_j(\pi, \|\varepsilon_0\|_q) \Big]^{r+1} \sup_{\pi\in[p_j,p_{j+1}]} \bar{b}_j(\pi, \|\varepsilon_0\|_q).
\end{aligned}$$

Thus we can find a finite constant $C > 0$ such that

$$\Big\| \sup_{\pi\in[p_j,p_{j+1}]} \sum_{r=0}^{\infty} \Big\{ \Big[\prod_{s=0}^{r} A_j(\pi, \varepsilon_{t-s}) \Big] \bar{b}_j(\pi, \varepsilon_{t-r-1}) \Big\}_{(1)} \Big\|_q \le C \sum_{r=0}^{\infty} \rho_j^{r+1}, \tag{1.127}$$

where ρ_j denotes the spectral radius of of $\sup_{\pi\in[p_j,p_{j+1}]} A_j(\pi, \|\varepsilon_0\|_q)$. For the proof of (i) it thus remains to show that $\rho_j < 1$.

The characteristic polynomial of the matrix $\sup_{\pi\in[p_j,p_{j+1}]} A_j(\pi, \|\varepsilon_0\|_q)$ is given by $p(\lambda) = (-1)^{\mathcal{P}} \lambda^{\mathcal{P}} \{1 - \sum_{s=1}^{\mathcal{P}} \sup_{\pi\in[p_j,p_{j+1}]} a_{j,s}(\pi) \|\varepsilon_0\|_q \lambda^{-s}\}$. We now prove by the way of contradiction that every eigenvalue λ fulfills $|\lambda| \le \big(\sum_{s=1}^{\mathcal{P}} \sup_{\pi\in[p_j,p_{j+1}]} a_{j,s}(\pi) \|\varepsilon_0\|_q \big)^{1/\mathcal{P}}$.

Assume that $|\lambda| > \big(\sum_{s=1}^{\mathcal{P}} \sup_{\pi\in[p_j,p_{j+1}]} a_{j,s}(\pi)\|\varepsilon_0\|_q\big)^{1/\mathcal{P}}$. Then we obtain by repeated application of the reverse triangle inequality that

$$\begin{aligned}
|p(\lambda)| &\geq |\lambda|^{\mathcal{P}}\Big\{1 - \sum_{s=1}^{\mathcal{P}} \sup_{\pi\in[p_j,p_{j+1}]} a_{j,s}(\pi)\|\varepsilon_0\|_q|\lambda|^{-s}\Big\} \\
&> |\lambda|^{\mathcal{P}}\Big\{1 - \sum_{s=1}^{\mathcal{P}} \sup_{\pi\in[p_j,p_{j+1}]} a_{j,s}(\pi)\|\varepsilon_0\|_q\Big(\sum_{s=1}^{\mathcal{P}} \sup_{\pi\in[p_j,p_{j+1}]} a_{j,s}(\pi)\|\varepsilon_0\|_q\Big)^{-s/\mathcal{P}}\Big\} \\
&\geq |\lambda|^{\mathcal{P}}\Big\{1 - \sum_{s=1}^{\mathcal{P}} \sup_{\pi\in[p_j,p_{j+1}]} a_{j,s}(\pi)\|\varepsilon_0\|_q\Big(\sum_{s=1}^{\mathcal{P}} \sup_{\pi\in[p_j,p_{j+1}]} a_{j,s}(\pi)\|\varepsilon_0\|_q\Big)^{-1}\Big\} \;=\; 0,
\end{aligned}$$

which means that there does not exist any eigenvalue. Hence $\rho_j \leq (\sum_{s=1}^{\mathcal{P}} \sup_{\pi\in[p_j,p_{j+1}]} a_{j,s}(\pi)\|\varepsilon_0\|_q)^{1/\mathcal{P}}$. But the latter bound is strictly smaller than 1 by the assumption of the lemma.

(ii): Let the function G_j be defined by (1.15), and set $\overline{G}_j(\pi,\boldsymbol{x}_i) := [G_j(\pi,\boldsymbol{x}_i), G_j(\pi,\boldsymbol{x}_{i-1}),\ldots,G_j(\pi,\boldsymbol{x}_{i-\mathcal{P}+1})]'$ ($\in \mathbb{R}^{\mathcal{P}}$) for any $\pi\in[p_j,p_{j+1}]$, $\boldsymbol{x}\in\mathbb{R}^{\mathbb{Z}}$, and $i\in\mathbb{N}$. Below we will show that for any $i\in\mathbb{N}$

$$\overline{G}_j(\pi,\boldsymbol{x}_i) = \bar{b}_j(\pi,x_i) + A_j(\pi,x_i)\overline{G}_j(\pi,\boldsymbol{x}_{i-1}) \quad \text{for any } \pi\in[p_j,p_{j+1}], \quad \mathbb{P}_{\boldsymbol{\epsilon}}\text{-a.e. } \boldsymbol{x}\in\mathbb{R}^{\mathbb{Z}}. \tag{1.128}$$

The first row of the vector equation in (1.128) is

$$G_j(\pi,\boldsymbol{x}_i) = a_{j,0}(\pi)x_i + \big[a_{j,1}(\pi)x_i, a_{j,2}(\pi)x_i,\ldots,a_{j,\mathcal{P}}(\pi)x_i\big]\overline{G}_j(\pi,\boldsymbol{x}_{i-1}),$$

which is just a restatement of the equation in (1.13).

To show (1.128), we note that for any $i\in\mathbb{N}$ and $\mathbb{P}_{\boldsymbol{\epsilon}}$-a.e. $\boldsymbol{x}\in\mathbb{R}^{\mathbb{Z}}$,

$$\overline{G}_j(\pi,\boldsymbol{x}_i) = \bar{b}_j(\pi,x_i) + \sum_{r=0}^{\infty}\Big\{\Big[\prod_{s=0}^{r} A_j(\pi,x_{i-s})\Big]\bar{b}_j(\pi,x_{i-r-1})\Big\} \quad \text{for any } \pi\in(p_j,p_{j+1}] \tag{1.129}$$

(this can be verified straightforwardly, using (1.15) and the definition of $\overline{G}_j$). Plugging (1.129) (with $i-1$ in place of i) in the right-hand side of (1.128) yields

$$\begin{aligned}
&\bar{b}_j(\pi,x_i) + A_j(\pi,x_i)\Big\{\bar{b}_j(\pi,x_{i-1}) + \sum_{r=0}^{\infty}\Big\{\Big[\prod_{s=0}^{r} A_j(\pi,x_{i-1-s})\Big]\bar{b}_j(\pi,x_{i-r-2})\Big\}\Big\} \\
&= \bar{b}_j(\pi,x_i) + A_j(\pi,x_i)\Big\{\bar{b}_j(\pi,x_{i-1}) + \sum_{r=1}^{\infty}\Big\{\Big[\prod_{s=1}^{r} A_j(\pi,x_{i-s})\Big]\bar{b}_j(\pi,x_{i-r-1})\Big\}\Big\} \\
&= \bar{b}_j(\pi,x_i) + A_j(\pi,x_i)\bar{b}_j(\pi,x_{i-1}) + \sum_{r=1}^{\infty}\Big\{\Big[\prod_{s=0}^{r} A_j(\pi,x_{i-s})\Big]\bar{b}_j(\pi,x_{i-r-1})\Big\} \\
&= \bar{b}_j(\pi,x_i) + \sum_{r=0}^{\infty}\Big\{\Big[\prod_{s=0}^{r} A_j(\pi,x_{i-s})\Big]\bar{b}_j(\pi,x_{i-r-1})\Big\} \;=\; \overline{G}_j(\pi,\boldsymbol{x}_i),
\end{aligned}$$

so that (1.128) indeed holds.

(iii): Let H_j be another solution of (1.13) with finite q-moments (as in assertion (iii)), and set $\overline{H}_j(\pi, \boldsymbol{x}_i) := [H_j(\pi, \boldsymbol{x}_i), H_j(\pi, \boldsymbol{x}_{i-1}), \ldots, H_j(\pi, \boldsymbol{x}_{i-\mathcal{P}+1})]'$ $(\in \mathbb{R}^{\mathcal{P}})$ for any $\pi \in [p_j, p_{j+1}]$, $\boldsymbol{x} \in \mathbb{R}^{\mathbb{Z}}$, and $i \in \mathbb{N}$. Let $n \in \mathbb{N}$ and $i \in \{1, \ldots, n\}$ with $i/n \in [p_j, p_{j+1}]$. In the following we will show that $\mathbb{P}$-a.s.

$$\overline{H}_j(i/n, \boldsymbol{\epsilon}_i) = \overline{b}_j(i/n, \varepsilon_i) + \sum_{r=0}^{\infty} \big\{ \big[\prod_{s=0}^{r} A_j(i/n, \varepsilon_{i-s}) \big] \overline{b}_j(i/n, \varepsilon_{i-r-1}) \big\}. \tag{1.130}$$

The first row of (1.130) shows that $H_j(i/n, \boldsymbol{\epsilon}_i) = G_j(i/n, \boldsymbol{\epsilon}_i)$ $\mathbb{P}$-a.s.

Since H_j solves (1.13), we get that (1.128) (with $\overline{G}_j$ replaced by $\overline{H}_j$) holds for any $i \in \mathbb{N}$. Performing this recursion $K \geq 1$ times, we obtain that $\mathbb{P}$-a.s.

$$\overline{H}_j(i/n, \boldsymbol{\epsilon}_i) = \overline{b}_j(i/n, \varepsilon_i) + \sum_{r=0}^{K-1} \big\{ \big[\prod_{s=0}^{r} A_j(i/n, \varepsilon_{i-s}) \big] \overline{b}_j(i/n, \varepsilon_{i-r-1}) \big\} + R_j(i/n, \boldsymbol{\epsilon}_{i-K}), \tag{1.131}$$

where $R_j(i/n, \boldsymbol{\epsilon}_{i-K}) := [\prod_{s=0}^{K} A_j(i/n, \varepsilon_{i-s})] \overline{H}_j(i/n, \boldsymbol{\epsilon}_{i-K-1})$. By part (i) the second summand on the right-hand side of (1.131) converges $\mathbb{P}$-a.s. to the second summand on the right-hand side of (1.130). For the proof of (1.130) it thus suffices to show that $\sum_{K=1}^{\infty} \mathbb{P}[\{|R_j(i/n, \boldsymbol{\epsilon}_{i-K})|\}_{(s)} > \eta] < \infty$ for any $\eta > 0$ and $s \in \{1, \ldots, \mathcal{P}\}$. In view of Markov's inequality, for this it in turn suffices to show that for any $s \in \{1, \ldots, \mathcal{P}\}$

$$\sum_{K=1}^{\infty} \big\| \big\{ R_j(i/n, \boldsymbol{\epsilon}_{i-K}) \big\}_{(s)} \big\|_q < \infty. \tag{1.132}$$

Since $A_j(i/n, \varepsilon_{i-s}), s = 0, \ldots, K$, and $\overline{H}_j(i/n, \boldsymbol{\epsilon}_{i-K-1})$ are independent, we have for any $K \in \mathbb{N}$

$$\begin{aligned} \big|\big\| R_j(i/n, \boldsymbol{\epsilon}_{i-K}) \big\|\big|_q &= \Big[\prod_{s=0}^{K} |\|A_j(i/n, \varepsilon_{i-s})\||_q \Big] \big|\big\| \overline{H}_j(i/n, \boldsymbol{\epsilon}_{i-K-1}) \big\|\big|_q \\ &= \big[A_j(i/n, \|\varepsilon_0\|_q) \big]^{K+1} \big|\big\| \overline{H}_j(i/n, \boldsymbol{\epsilon}_{i-K-1}) \big\|\big|_q . \end{aligned}$$

By the assumption that $\|\varepsilon_0\|_q \max_{j=0,\ldots,\ell} \sum_{s=0}^{\mathcal{P}} \sup_{\pi \in [p_j, p_{j+1}]} a_{j,s}(\pi) < 1$, the spectral radius ρ_j of $A_j(i/n, \|\varepsilon_0\|_q)$ is strictly smaller than 1 (as we have seen in the proof of part (i)). Since the q-th moments of H_j are finite by assumption, we can find a finite constant $C > 0$ such that

$$\big\{ \big|\big\| R_j(i/n, \boldsymbol{\epsilon}_{i-K}) \big\|\big|_q \big\}_{(s)} \leq C \rho_j^{K+1}$$

for any $K \in \mathbb{N}$ and $s \in \{1, \ldots, \mathcal{P}\}$. Since ρ_j^{K+1} goes to 0 exponentially fast as $K \to \infty$, (1.132) holds. □

1.5.7 Proof of Corollary 1.2.9

First of all, we give an explicit description of $\mathfrak{F}_{n,i}(x, \boldsymbol{x})$ in the specific setting of the PLS ARCH process in Subsection 1.2.3. We note that for $i = 1, \ldots, n$, $x \in \mathbb{R}$, and $\boldsymbol{x} = (x_1, x_2, \ldots) \in \mathbb{R}^{\mathbb{N}}$

$$\begin{aligned} &\mathbb{P}[X_{n,i} \leq x \| \boldsymbol{\epsilon}_{i-1} = \boldsymbol{x}] \\ &= \mathbb{P}\Big[\varepsilon_i \Big\{ \sum_{j=0}^{\ell} \Big(a_{j,0}\big(\tfrac{i}{n}\big) + \sum_{s=1}^{\mathcal{P}} a_{j,s}\big(\tfrac{i}{n}\big) G_j\big(\tfrac{i}{n}, \boldsymbol{\epsilon}_{i-s}\big)\Big) \mathbb{1}_{(p_j, p_{j+1}]}\big(\tfrac{i}{n}\big) \Big\} \leq x \Big\| \boldsymbol{\epsilon}_{i-1} = \boldsymbol{x}\Big] \\ &= \mathbb{P}\big[\varepsilon_i \leq x / \Lambda_{n,i}(\boldsymbol{x})\big] \;=\; F_\varepsilon\big(x/\Lambda_{n,i}(\boldsymbol{x})\big), \end{aligned}$$

where $\Lambda_{n,i}(\boldsymbol{x}) := \sum_{j=0}^{\ell} (a_{j,0}\big(\tfrac{i}{n}\big) + \sum_{s=1}^{\mathcal{P}} a_{j,s}\big(\tfrac{i}{n}\big) G_j\big(\tfrac{i}{n}, \boldsymbol{x}_{s+1}\big)) \mathbb{1}_{(p_j, p_{j+1}]}\big(\tfrac{i}{n}\big)$ and F_ε denotes the distribution function of ε_0. For $i = 1, \ldots, n$, we may thus define $\mathfrak{F}_{n,i}(x, \boldsymbol{x})$ through

$$\mathfrak{F}_{n,i}(x, \boldsymbol{x}) := F_\varepsilon\big(x/\Lambda_{n,i}(\boldsymbol{x})\big). \tag{1.133}$$

As a result

$$\mathfrak{f}_{n,i}(x, \boldsymbol{x}) = f_\varepsilon\big(x/\Lambda_{n,i}(\boldsymbol{x})\big)/\Lambda_{n,i}(\boldsymbol{x}) \quad \text{and} \quad \mathfrak{f}'_{n,i}(x, \boldsymbol{x}) = f'_\varepsilon\big(x/\Lambda_{n,i}(\boldsymbol{x})\big)/\Lambda^2_{n,i}(\boldsymbol{x}). \tag{1.134}$$

Moreover,

$$f_{n,i}(x) := \mathbb{E}[\mathfrak{f}_{n,i}(x, \boldsymbol{\epsilon}_{i-1})] = \mathbb{E}\big[f_\varepsilon\big(x/\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})\big)/\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})\big] \tag{1.135}$$

provides a Lebesgue density $f_{n,i}$ of $X_{n,i}$ for any $i = 1, \ldots, n$.

(A3): Let q and γ be as in conditions (a) and (b) of the corollary. Without loss of generality assume that $\gamma \leq q$. Since $\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1}) \geq \min_{j=0,\ldots,\ell} \inf_{\pi \in (p_j, p_{j+1}]} a_{j,0}(\pi) =: \beta$ for any $i = 1, \ldots, n$, and $\beta > 0$ by assumption (a), we obtain in view of (1.135)

$$\|f_{n,i}\|_{(\gamma)} \;=\; \big\|\mathbb{E}\big[f_\varepsilon\big(\,\cdot\,/\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})\big)/\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})\big]\big\|_{(\gamma)} \;\leq\; \frac{1}{\beta} \sup_{x \in \mathbb{R}} \mathbb{E}\big[f_\varepsilon\big(x/\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})\big)\phi_\gamma(x)\big]$$

for any $i = 1, \ldots, n$. By assumption (b), there exists a finite constant $c > 0$ such that $f_\varepsilon(x)\phi_\gamma(x) \leq c$ for all $x \in \mathbb{R}$. As a result we have for any $x \in \mathbb{R}$ and $i = 1, \ldots, n$

$$f_\varepsilon\big(x/\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})\big)\phi_\gamma(x) \;\leq\; c\,\phi_\gamma(x)\,\phi_{-\gamma}(x/\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})) \;\leq\; c\,\max\{1, \Lambda^\gamma_{n,i}(\boldsymbol{\epsilon}_{i-1})\}. \tag{1.136}$$

Consequently

$$\|f_{n,i}\|_{(\gamma)} \;\leq\; \frac{c}{\beta}\,\mathbb{E}\big[\max\{1, \Lambda^\gamma_{n,i}(\boldsymbol{\epsilon}_{i-1})\}\big] \;\leq\; \frac{c}{\beta}\big(1 + \|\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})\|^\gamma_\gamma\big).$$

Assertion (A3) follows, if we can show that $\sup_{n \in \mathbb{N}} \max_{1 \leq i \leq n} \|\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})\|^\gamma_\gamma < \infty$. Note that $\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})$ coincides with the sum $\sum_{j=0}^{\ell} G_j(\pi, \boldsymbol{\epsilon}_{i;1}) \mathbb{1}_{(p_j, p_{j+1}]}(i/n)$, where we set $\boldsymbol{\epsilon}_{i;1} := (1, \varepsilon_{i-1}, \varepsilon_{i-2}, \ldots)$ and $G_j(\pi, \boldsymbol{\epsilon}_{i;1})$ can be represented analogously to (1.16). It thus suffices to verify that $\max_{j=0,\ldots,\ell} \sup_{\pi \in (p_j, p_{j+1}]} \|G_j(\pi, \boldsymbol{\epsilon}_{i;1})\|_q < \infty$, because we assumed $\gamma \leq q$. The latter assertion can be shown in the same way as part (i) of Lemma 1.2.8.

(A4): By (1.13), we have for any $\pi, \pi' \in (p_{j_p}, p_{j_p+1}]$,

$$\begin{aligned} & G_{j_p}(\pi, \boldsymbol{\epsilon}_0) - G_{j_p}(\pi', \boldsymbol{\epsilon}_0) \\ = \; & \varepsilon_0\big(a_{j_p,0}(\pi) - a_{j_p,0}(\pi')\big) + \sum_{s=1}^{\mathcal{P}} \big(a_{j_p,s}(\pi) - a_{j_p,s}(\pi')\big) G_{j_p}(\pi, \boldsymbol{\epsilon}_{-s})\, \varepsilon_0 \\ & + \sum_{s=1}^{\mathcal{P}} a_{j_p,s}(\pi')\big(G_{j_p}(\pi, \boldsymbol{\epsilon}_{-s}) - G_{j_p}(\pi', \boldsymbol{\epsilon}_{-s})\big)\varepsilon_0. \end{aligned} \tag{1.137}$$

By the differentiability of $a_{j_p,k}(\pi)$ and the independence between $G_{j_p}(\pi, \boldsymbol{\epsilon}_{-s})$ and ε_0, there exist two finite constants $C_1 > 0$ and $C_2 > 0$ such that $\|\varepsilon_0\big(a_{j_p,0}(\pi) - a_{j_p,0}(\pi')\big)\|_1 \le C_1|\pi - \pi'|$ and $\| \sum_{s=1}^{\mathcal{P}} \big(a_{j_p,s}(\pi) - a_{j_p,s}(\pi')\big) G_{j_p}(\pi, \boldsymbol{\epsilon}_{-s})\, \varepsilon_0\|_1 \le C_2|\pi - \pi'|$. Furthermore, observe that, due to the stationarity of the process $\{\boldsymbol{\epsilon}_{-s}\}_{s=0}^{\mathcal{P}}$, we have $\|G_{j_p}(\pi, \boldsymbol{\epsilon}_{-s}) - G_{j_p}(\pi', \boldsymbol{\epsilon}_{-s})\|_1 = \|G_{j_p}(\pi, \boldsymbol{\epsilon}_0) - G_{j_p}(\pi', \boldsymbol{\epsilon}_0)\|_1$, $s = 1, 2, \ldots, \mathcal{P}$. Hence

$$\begin{aligned} \Big\| \sum_{s=1}^{\mathcal{P}} a_{j_p,s}(\pi')\big(G_{j_p}(\pi, \boldsymbol{\epsilon}_{-s}) - G_{j_p}(\pi', \boldsymbol{\epsilon}_{-s})\big)\, \varepsilon_0 \Big\|_1 & \le a \max_{1 \le s \le \mathcal{P}} \|G_{j_p}(\pi, \boldsymbol{\epsilon}_{-s}) - G_{j_p}(\pi', \boldsymbol{\epsilon}_{-s})\|_1 \\ & = a\|G_{j_p}(\pi, \boldsymbol{\epsilon}_0) - G_{j_p}(\pi', \boldsymbol{\epsilon}_0)\|_1, \end{aligned}$$

where $a = \|\varepsilon_0\|_1 \max_{j=0,\ldots,\ell} \sum_{s=0}^{\mathcal{P}} \sup_{\pi \in [p_j, p_{j+1}]} a_{j,s}(\pi) < 1$. Plugging the above results into (1.137), we have

$$\big\|G_{j_p}(\pi, \boldsymbol{\epsilon}_0) - G_{j_p}(\pi', \boldsymbol{\epsilon}_0)\big\|_1 \le (C_1 + C_2)\, |\pi - \pi'|/(1 - a).$$

(B4): Assumption (B4) can be shown analogously to (A4); we omit the details.

(A5): We observe that for any $r \in \mathbb{N}$

$$\begin{aligned} & \delta_{\boldsymbol{\epsilon},r;q} \\ = & \sup_{\pi \in (p_{j_p}, p_{j_p+1}]} \big\| \big\{ \overline{G}_{j_p}(\pi, \boldsymbol{\epsilon}_0) - \overline{G}_{j_p}(\pi, \boldsymbol{\epsilon}^*_{0,-r}) \big\}_{(1)} \big\|_q \\ = & \sup_{\pi \in (p_{j_p}, p_{j_p+1}]} \Big\| \Big\{ \Big[\prod_{s=0}^{r-1} A_{j_p}(\pi, \varepsilon_{-s}) \Big] \bar{b}_{j_p}(\pi, \varepsilon_{-r} - \varepsilon^*) \\ & + \sum_{t=r}^{\infty} \Big(\Big[\prod_{s=0}^{r-1} A_{j_p}(\pi, \varepsilon_{-s}) \Big] A_{j_p}(\pi, \varepsilon_{-r} - \varepsilon^*) \Big[\prod_{s=r+1}^{t} A_{j_p}(\pi, \varepsilon_{-s}) \Big] \bar{b}_{j_p}(\pi, \varepsilon_{-t-1}) \Big) \Big\}_{(1)} \Big\|_q, \end{aligned}$$

where $[\prod_{s=r+1}^{r} A_{j_p}(\pi, \varepsilon_{-s})]$ corresponds to the $\mathcal{P} \times \mathcal{P}$-dimensional identity matrix. Analogously to the proof of part (i) of Lemma 1.2.8, we can find constants C and $\widetilde{C}$ such that

$$\delta_{\boldsymbol{\epsilon},r;q} \le \Big\{ \Big[\sup_{\pi \in (p_{j_p}, p_{j_p+1}]} A_{j_p}(\pi, \|\varepsilon_0\|_q) \Big]^r \sup_{\pi \in (p_{j_p}, p_{j_p+1}]} \bar{b}_{j_p}(\pi, 2\|\varepsilon_0\|_q)$$

$$
\begin{aligned}
&+\sum_{t=r}^{\infty}\Big(\big[\sup_{\pi\in(p_{jp},p_{jp+1}]} A_{j_p}(\pi,\|\varepsilon_0\|_q)\big]^r\cdot\sup_{\pi\in(p_{jp},p_{jp+1}]} A_{j_p}(\pi,2\|\varepsilon_0\|_q)\\
&\cdot\big[\sup_{\pi\in(p_{jp},p_{jp+1}]} A_{j_p}(\pi,\|\varepsilon_0\|_q)\big]^{t-r}\,\overline{b}_{j_p}(\pi,\|\varepsilon_0\|_q)\Big)\Big\}_{(1)}\\
\leq\;& 2\Big\{\sum_{t=r-1}^{\infty}\Big(\big[\sup_{\pi\in(p_{jp},p_{jp+1}]} A_{j_p}(\pi,\|\varepsilon_0\|_q)\big]^{t+1}\,\overline{b}_{j_p}(\pi,\|\varepsilon_0\|_q)\Big)\Big\}_{(1)}\\
\leq\;& 2C\sum_{t=r-1}^{\infty}\rho_{j_p}^{t+1}\;\leq\;2C\sum_{t=r-1}^{\infty}a^{t+1}\;=\;\widetilde{C}a^r.
\end{aligned}
$$

Here, ρ_{j_p} denotes the spectral radius of $\sup_{\pi\in(p_{jp},p_{jp+1}]} A_{j_p}(\pi,\|\varepsilon_0\|_q)$, which is less than $a := (\|\varepsilon_0\|_q \max_j \sup_{\pi\in(p_{jp},p_{jp+1}]}\sum_{s=0}^{\mathcal{P}} a_{j,s}(\pi))^{1/\mathcal{P}}$ (as we have seen in the proof of part (i) of Lemma 1.2.8). Since $a<1$ by assumption (a), assertion (A5) follows.

(A7) and (A8): We shall only prove $M_{2,\alpha}(\mathbb{R})<\infty$ for $\alpha=1$ (and thus for any $\alpha\in[0,1]$) here since the other claims in (A7) and (A8) (with $\alpha=1$) follow by similar arguments. In the following, we will prove that

$$\delta^2_{\boldsymbol{\epsilon},r;2}(\mathfrak{f}_{n,i};x)\;\leq\;C\overline{a}^r/\phi_{2v}(x) \tag{1.138}$$

for some constants $\overline{a}<1$ and $v>\lambda+\alpha/2+1/2$. Then the assertion that $M_{2,\alpha}(\mathbb{R})<\infty$ immediately follows by plugging (1.138) into the definition of $M_{2,\alpha}(\mathbb{R})$.

Let $k:=\min\{1,\frac{1}{2}(\gamma-2\lambda-1),\frac{1}{2}(\frac{q}{2}-2\lambda-1)\}$. Below we will show that there exist constants $0\leq C_1,C_2<\infty$ and $0\leq\widetilde{a}<1$ such that

$$\delta^2_{\boldsymbol{\epsilon},r;2}(\mathfrak{f}_{n,i};x)\;\leq\;C_1\phi_{-4\lambda-2-2k}(x), \tag{1.139}$$
$$\delta^2_{\boldsymbol{\epsilon},r;2}(\mathfrak{f}_{n,i};x)\;\leq\;C_2\phi_{-2k}(x)\widetilde{a}^{2kr}. \tag{1.140}$$

These two inequalities imply (1.138), because

$$
\begin{aligned}
\delta^2_{\boldsymbol{\epsilon},r;2}(\mathfrak{f}_{n,i};x)\;&\leq\;C_1\phi_{-4\lambda-2-2k}(x)\,\min\big\{1,\big(C_2\phi_{-2k}(x)\widetilde{a}^{2kr}\big)/\big(C_1\phi_{-4\lambda-2-2k}(x)\big)\big\}\\
&\leq\;C_1^{1-k/2}C_2^{k/2}\,\phi_{-2\lambda(2-k)-2-k}(x)\,\widetilde{a}^{2kr}.
\end{aligned}
$$

The latter step relies on the inequality $\min\{1,z\}\leq z^{k/2}$ for $z\in\mathbb{R}_+$.

We start with the proof of (1.139). We note that $\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1}),\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})\geq\beta$ for all $i=1,\ldots,n$, where β is as in the proof of (A3). Moreover, $\|f_\varepsilon\|_{(2\lambda+1+k)}<\infty$ by assumption (b) and hence (1.136) holds (with $\gamma=2\lambda+1+k$) for some finite constant c_1 (in place of c). Consequently

$$
\begin{aligned}
&\delta^2_{\boldsymbol{\epsilon},r;2}(\mathfrak{f}_{n,i};x)\\
&\leq\;\frac{2}{\beta^2}\Big\{\big\|f_\varepsilon\big(x/\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})\big)\big\|_2^2+\big\|f_\varepsilon\big(x/\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})\big)\big\|_2^2\Big\}\\
&\leq\;\frac{2c_1^2}{\beta^2}\,\phi_{-4\lambda-2-2k}(x)\Big\{\big\|\max\big\{1,\Lambda_{n,i}^{2\lambda+1+k}(\boldsymbol{\epsilon}_{i-1})\big\}\big\|_2^2+\big\|\max\big\{1,\Lambda_{n,i}^{2\lambda+1+k}(\boldsymbol{\epsilon}^*_{i,i-r})\big\}\big\|_2^2\Big\}
\end{aligned}
$$

$$\leq \frac{2c_1^2}{\beta^2}\,\phi_{-4\lambda-2-2k}(x)\Big\{1+\big\|\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})\big\|_{4\lambda+2+2k}^{4\lambda+2+2k}+1+\big\|\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})\big\|_{4\lambda+2+2k}^{4\lambda+2+2k}\Big\}.$$

Since $4\lambda+2+2k \leq q$ and hence $\sup_{n\in\mathbb{N}}\max_{1\leq i\leq n}\|\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})\|_{4\lambda+2+2k} < \infty$ analogously to the argumentation in the proof of (A3), we may thus find a finite constant C_1 such that (1.139) holds.

For the proof of (1.140), we observe

$$\begin{aligned}\delta^2_{\boldsymbol{\epsilon},r;2}(\mathfrak{f}_{n,i};x) &\leq 2\Big\|\Lambda^{-1}_{n,i}(\boldsymbol{\epsilon}_0)\Big(f_\varepsilon\big(x/\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1})\big)-f_\varepsilon\big(x/\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})\big)\Big)\Big\|_2^2\\ &\quad +2\Big\|f_\varepsilon\big(x/\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})\big)\Big(\Lambda^{-1}_{n,i}(\boldsymbol{\epsilon}_0)-\Lambda^{-1}_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})\Big)\Big\|_2^2\\ &=: S_1(n,i;r)+S_2(n,i;r). \end{aligned} \tag{1.141}$$

Since $\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1}),\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r}) \geq \beta$ for all $i=1,\ldots,n$ and $\|f'_\varepsilon\|_{(k+1)} \leq \|f'_\varepsilon\|_{(\gamma)} < \infty$ (hence $|f'_\varepsilon(x/y)|\phi_{k+1}(x) \leq c_2\max\{1,|y|^{k+1}\}$ for any $x,y\in\mathbb{R}$, analogously to (1.136)) we obtain for the first summand

$$\begin{aligned} &S_1(n,i;r)\\ &\leq 2\beta^{-2}\Big\|\int_{\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})}^{\Lambda_{n,i}(\boldsymbol{\epsilon}_0)} -xy^{-2}f'_\varepsilon\big(\tfrac{x}{y}\big)\,dy\Big\|_2^2\\ &\leq \frac{2c_2^2}{\beta^2}\,\phi_{-2k-2}(x)\,|x|^2\Big\|\int_{\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})}^{\Lambda_{n,i}(\boldsymbol{\epsilon}_0)} y^{-2}\max\{1,y^{k+1}\}\,dy\Big\|_2^2\\ &\leq \frac{4c_2^2}{\beta^2}\,\phi_{-2k}(x)\Big\{\Big\|\int_{\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})}^{\Lambda_{n,i}(\boldsymbol{\epsilon}_0)} y^{-2}\,dy\Big\|_2^2+\Big\|\int_{\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})}^{\Lambda_{n,i}(\boldsymbol{\epsilon}_0)} y^{k-1}\,dy\Big\|_2^2\Big\}\\ &= \frac{4c_2^2}{\beta^2}\,\phi_{-2k}(x)\Big\{\big\|\Lambda^{-1}_{n,i}(\boldsymbol{\epsilon}_0)-\Lambda^{-1}_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})\big\|_2^2+\frac{1}{k^2}\big\|\Lambda_{n,i}(\boldsymbol{\epsilon}_0)-\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})\big\|_{2k}^{2k}\Big\}\\ &\leq \frac{4c_2^2}{\beta^2}\Big(k^{-2}+\beta^{-4}\Big)\,\phi_{-2k}(x)\max_{t\in\{1,k\}}\big\|\Lambda_{n,i}(\boldsymbol{\epsilon}_0)-\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})\big\|_2^{2t} \end{aligned} \tag{1.142}$$

for some finite constant c_2.

For the second summand, there exists a finite constant c_3 such that

$$\begin{aligned} S_2(n,i;r) &\leq 2c_3^2\,\phi_{-2}(x)\Big\|\max\{1,|\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})|\}\Big(\Lambda^{-1}_{n,i}(\boldsymbol{\epsilon}_0)-\Lambda^{-1}_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})\Big)\Big\|_2^2\\ &\leq 4c_3^2\,\phi_{-2}(x)\Big\{\big\|\Lambda^{-1}_{n,i}(\boldsymbol{\epsilon}_0)-\Lambda^{-1}_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})\big\|_2^2+\big\|\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})\Lambda^{-1}_{n,i}(\boldsymbol{\epsilon}_0)-1\big\|_2^2\Big\}\\ &\leq \frac{4c_3^2}{\beta^2}\,(\beta^{-2}+1)\,\phi_{-2}(x)\,\big\|\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})-\Lambda_{n,i}(\boldsymbol{\epsilon}_0)\big\|_2^2, \end{aligned}$$

where we used that $|f_\varepsilon(x/y)|\phi_1(x) \leq c_3\max\{1,|y|\}$ for any $x,y\in\mathbb{R}$ (because of the boundedness of f_ε) analogously to (1.136) in the first step and $\Lambda_{n,i}(\boldsymbol{\epsilon}_{i-1}),\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r}) \geq \beta$ for all $i=1,\ldots,n$ in the last step. In view of (1.141) and (1.142), this results in

$$\delta^2_{\boldsymbol{\epsilon},r;2}(\mathfrak{f}_{n,i};x) \leq c_4\,\phi_{-2k}(x)\max_{t\in\{1,k\}}\big\|\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})-\Lambda_{n,i}(\boldsymbol{\epsilon}_0)\big\|_2^{2t} \tag{1.143}$$

for some finite constant c_4. For the proof of (1.140) it remains to show that $\big\|\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})-\Lambda_{n,i}(\boldsymbol{\epsilon}_0)\big\|_2$ is finite. We note that the latter difference coincides with $\sum_{j=0}^{\ell}\big(G_j(\pi,\boldsymbol{\epsilon}^*_{i,i-r;1}) - G_j(\pi,\boldsymbol{\epsilon}_{i;1})\big)\mathbb{1}_{(p_j,p_{j+1}]}(i/n)$, where we use the same notation as introduced in (A3) and define analogously $\boldsymbol{\epsilon}^*_{i,i-r;1} := (1,\boldsymbol{\epsilon}^*_{i-1,i-r})$. By the same line of arguments as in the proof of (A5) we obtain

$$\big\|\Lambda_{n,i}(\boldsymbol{\epsilon}^*_{i,i-r})-\Lambda_{n,i}(\boldsymbol{\epsilon}_0)\big\|_2 \;\le\; \max_{j=0,\ldots,\ell}\sup_{\pi\in(p_j,p_{j+1}]}\big\|\big\{\overline{G}_j(\pi,\boldsymbol{\epsilon}_{i;1})-\overline{G}_j(\pi,\boldsymbol{\epsilon}^*_{i,i-r;1})\big\}_{(1)}\big\|_2 \;\le\; c_5\,\widetilde{a}^r$$

for some finite constants c_5 and $\widetilde{a}<1$. Along with (1.143) this proves (1.140).

(A9): Analogously to (1.135) we obtain that $f_p(x) := \mathbb{E}[f_\varepsilon(x/\Lambda_p(\boldsymbol{\epsilon}_0))/\Lambda_p(\boldsymbol{\epsilon}_0)]$ with $\Lambda_p(\boldsymbol{\epsilon}_0) := \sum_{j=0}^{\ell}(a_{j,0}(p)+\sum_{s=1}^{\mathcal{P}} a_{j,s}(p)G_j\big(p,\boldsymbol{\epsilon}_{-s}))\mathbb{1}_{(p_j,p_{j+1}]}(p)$ provides a Lebesgue density f_p of ξ_p. Since $\Lambda_p(\boldsymbol{\epsilon}_0)\ge\beta$, we obtain

$$\|f_p\|_\infty \le \beta^{-1}\big\|\mathbb{E}\big[f_\varepsilon\big(x/\Lambda_p(\boldsymbol{\epsilon}_0)\big)\big]\big\|_\infty \;\le\; \beta^{-1}\sup_{z\in\mathbb{R}}|f_\varepsilon(z)|,$$

so that (A9) follows in view of assumption (b). □

1.5.8 Proof of Lemma 1.3.1

As before let j_p be the unique index j with $p\in(p_j,p_{j+1})$. Then, for n sufficiently large (depending only on p_j and p_{j+1}), we have $i_{p,n}/n\in(p_j,p_{j+1})$. Without loss of generality we will only consider n being sufficiently large. Then $\xi_{p,n} := G_{j_p}(i_{p,n}/n,\boldsymbol{\epsilon}_0)$ has the same distribution as $X_{n,i_{p,n}}$. Moreover $\xi_p = G_{j_p}(p,\boldsymbol{\epsilon}_0)$. Thus for any $x\in\mathbb{R}$

$$\begin{aligned}
&|F_{p,n}(x)-F_p(x)|\\
&\quad= \big|\mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(\xi_{p,n})-\mathbb{1}_{(-\infty,x]}(\xi_p)\big]\big|\\
&\quad= \big|\mathbb{E}\big[\mathbb{1}_{(-\infty,x]}(G_{j_p}(i_{p,n}/n,\boldsymbol{\epsilon}_0))-\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_0))\big]\big|\\
&\quad\le \big|\mathbb{E}\big[\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i_{p,n}/n,\boldsymbol{\epsilon}_0))-\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_0))\big)\mathbb{1}_{\{|G_{jp}(i_{p,n}/n,\boldsymbol{\epsilon}_0)-G_{jp}(p,\boldsymbol{\epsilon}_0)|\le x\delta_n\}}\big]\big|\\
&\qquad+\big|\mathbb{E}\big[\big(\mathbb{1}_{(-\infty,x]}(G_{j_p}(i_{p,n}/n,\boldsymbol{\epsilon}_0))-\mathbb{1}_{(-\infty,x]}(G_{j_p}(p,\boldsymbol{\epsilon}_0))\big)\mathbb{1}_{\{|G_{jp}(i_{p,n}/n,\boldsymbol{\epsilon}_0)-G_{jp}(p,\boldsymbol{\epsilon}_0)|> x\delta_n\}}\big]\big|\\
&\quad=: S_1(n,x)+S_2(n,x)
\end{aligned}\tag{1.144}$$

for any $\delta_n>0$. For the first summand we have for any $x\ge1$

$$S_1(n,x)\;\le\;\mathbb{P}\big[x-x\delta_n\le G_{j_p}(i_{p,n}/n,\boldsymbol{\epsilon}_0)\le x+x\delta_n\big]\;=\;\int_{x-x\delta_n}^{x+x\delta_n} f_{n,i_{p,n}}(y)\,dy.$$

Assuming that δ_n is nonincreasing and tends to 0 as $n\to\infty$, we can choose $n_0\in\mathbb{N}$ such that $\delta_n\le1/2$ for all $n\ge n_0$. Then, for any $n\ge n_0$ and $x\ge1$

$$\phi_\lambda(x)S_1(n,x)\;\;\le\;\;\phi_\lambda(x)\int_{x(1-\delta_n)}^{x(1+\delta_n)} f_{n,i_{p,n}}(y)\,dy$$

$$
\begin{aligned}
&\leq \|f_{n,i_{p,n}}\|_{(\gamma)}\,\phi_\lambda(x)\int_{x(1-\delta_n)}^{x(1+\delta_n)}\phi_{-\gamma}(y)\,dy\\
&\leq C_1\phi_\lambda(x)\,(2x\delta_n)\sup_{y\in(x(1-\delta_n),x(1+\delta_n))}\phi_{-\gamma}(y)\\
&\leq 2C_1\delta_n\,\phi_{\lambda+1}(x)\sup_{y\in(x(1-\delta_n),x(1+\delta_n))}\phi_{-\gamma}(y)\\
&= 2C_1\delta_n\,\phi_{\lambda+1}(x)\phi_{-\gamma}(x(1-\delta_n))\ \leq\ 2C_1\delta_n\,\frac{(1+x)^{\lambda+1}}{(1+x/2)^{\gamma}}\\
&\leq 2C_1\delta_n\left(\frac{1}{1+x/2}+\frac{x}{1+x/2}\right)^{\lambda+1}\ \leq\ C_2\delta_n,
\end{aligned}
$$

where $C_2 := 2C_1 3^{\lambda+1}$ and $C_1 := \sup_{n\in\mathbb{N}}\max_{i\in I_{p,n}}\|f_{n,i}\|_{(\gamma)} < \infty$ (recall (A3)). Thus $\sup_{x\in[1,\infty)} S_1(n,x)\phi_\lambda(x) = \mathcal{O}(\delta_n)$. In the same way we obtain the analogue with "$x \in [1,\infty)$" replaced by "$x \in (-\infty,-1]$". Hence $\sup_{x\in\mathbb{R}\setminus[-1,1]} S_1(n,x)\phi_\lambda(x) = \mathcal{O}(\delta_n)$. For the second summand we have for any $x \in \mathbb{R}\setminus\{0\}$ and some constant $C_3 > 0$

$$
\begin{aligned}
S_2(n,x) &\leq \mathbb{P}\big[|G_{j_p}(i_{p,n}/n,\boldsymbol{\epsilon}_0) - G_{j_p}(p,\boldsymbol{\epsilon}_0)| > x\delta_n\big]\\
&\leq (x\delta_n)^{-q}\big\|G_{j_p}(i_{p,n}/n,\boldsymbol{\epsilon}_0) - G_{j_p}(p,\boldsymbol{\epsilon}_0)\big\|_q^q\ \leq\ C_3(x\delta_n)^{-q}n^{-q},
\end{aligned}
$$

where we used Markov's inequality and (B4). Thus $\phi_\lambda(x)S_2(n,x) \leq 2^\lambda C_3(n\delta_n)^{-q}$ for any $x \in \mathbb{R}\setminus[-1,1]$. Thus $\sup_{x\in\mathbb{R}\setminus[-1,1]}\phi_\lambda(x)S_2(n,x) = \mathcal{O}((n\delta_n)^{-q})$, and therefore $\sup_{x\in\mathbb{R}\setminus[-1,1]}\phi_\lambda(x)|F_{p,n}(x) - F_p(x)| = \mathcal{O}(\delta_n) + \mathcal{O}((n\delta_n)^{-q})$.

By the same line of arguments (but with $\leq x\delta_n$ and $> x\delta_n$ in (1.144) replaced by $\leq \delta_n$ and $> \delta_n$ respectively) we obtain $\sup_{x\in[-1,1]}|F_{p,n}(x)-F_p(x)| = \mathcal{O}(\delta_n)+\mathcal{O}((n\delta_n)^{-q})$. Altogether, $\|F_{p,n} - F_p\|_{(\lambda)} = \mathcal{O}(\delta_n) + \mathcal{O}((n\delta_n)^{-q})$. Choosing $\delta_n := n^{-q/(q+1)}$ we arrive at $\|F_{p,n} - F_p\|_{(\lambda)} = \mathcal{O}(n^{-q/(q+1)})$. Together with (C2) this gives the claim. □

Chapter 2

Integration by parts for multivariate functions

2.1 Introduction

This chapter is devoted to the development of an integration by parts formula for multivariable functions of (locally) bounded variation as defined in Section 2.4. In 1917, Young [81] elaborated such a multivariate integration by parts formula for Riemann-Stieltjes integrals, where the use of this special type of integrals forced him to assume continuity of at least one of the involved functions. For Lebesgue-Stieltjes integration, two-dimensional versions can be found for instance in Gill et al. [40], Dehling and Taqqu [25], Beutner et al. [8] and in Beutner and Zähle [11]. Recently, Berghaus et al. [7] proved a two-dimensional integration by parts formula on compact intervals by using a slightly different type of variation. The generalization thereof to multivariable functions is part of the recent work of Radulović et al. [63].

Section 2.6 below provides an integration by parts formula for multivariable functions of locally bounded variation, which is closely related to [63]. For its formulation we need some preparation. In Sections 2.2–2.3 we recall the notion of d-fold monotonically increasing functions and we discuss the connection to positive Borel measures on $\mathbb{R}^d$. In Sections 2.4–2.5 we recall the notion of functions that are locally of bounded d-fold variation and we discuss the connection to signed Borel measures on $\mathbb{R}^d$.

2.2 Multi-monotonically increasing functions

For any $\boldsymbol{a} = (a_1, \ldots, a_d)$ and $\boldsymbol{b} = (b_1, \ldots, b_d)$ in $\mathbb{R}^d$, we will write $\boldsymbol{a} \leq \boldsymbol{b}$ if $a_j \leq b_j$ for $j = 1, \ldots, d$, and $\boldsymbol{a} < \boldsymbol{b}$ if $a_j < b_j$ for $j = 1, \ldots, d$. For any $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} \leq \boldsymbol{b}$ we denote by $[\boldsymbol{a}, \boldsymbol{b}]$ the set of all $\boldsymbol{x} \in \mathbb{R}^d$ satisfying $\boldsymbol{a} \leq \boldsymbol{x} \leq \boldsymbol{b}$. For any $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$ we denote by $(\boldsymbol{a}, \boldsymbol{b}]$ the set of all $\boldsymbol{x} \in \mathbb{R}^d$ satisfying $\boldsymbol{a} < \boldsymbol{x} \leq \boldsymbol{b}$. The set

$[\boldsymbol{a},\boldsymbol{b}]$ can be seen as a (generalized) closed interval, and the set $(\boldsymbol{a},\boldsymbol{b}]$ can be seen as a (generalized) half-open interval. The cardinality of a set J will be denoted by $|J|$, using the convention $|\emptyset| := 0$.

Definition 2.2.1 *Let $F : \mathbb{R}^d \to \mathbb{R}$ be any function. For $\boldsymbol{a} = (a_1,\dots,a_d) \in \mathbb{R}^d$ and $\boldsymbol{b} = (b_1,\dots,b_d) \in \mathbb{R}^d$, we set*

$$\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F := \sum_{J\subseteq\{1,\dots,d\}} (-1)^{d-|J|}\, F(\boldsymbol{b}^{\boldsymbol{a};J}), \tag{2.1}$$

where

$$\boldsymbol{b}^{\boldsymbol{a};J} := (b_1^{\boldsymbol{a};J},\dots,b_d^{\boldsymbol{a};J}) \quad \text{with} \quad b_j^{\boldsymbol{a};J} := \begin{cases} b_j & , \; j\in J \\ a_j & , \; j\notin J \end{cases}. \tag{2.2}$$

*When $\boldsymbol{a} \le \boldsymbol{b}$, we refer to $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F$ as the d-*fold increase of F on the interval $[\boldsymbol{a},\boldsymbol{b}]$.

For illustration, let $F_i : \mathbb{R}^i \to \mathbb{R}$, $i = 1,2,3$, be any functions. Then

$$\begin{aligned}
\Delta_{a_1}^{b_1} F_1 &= F_1(b_1) - F_1(a_1), \\
\Delta_{(a_1,a_2)}^{(b_1,b_2)} F_2 &= F_2(b_1,b_2) + F_2(a_1,a_2) - F_2(a_1,b_2) - F_2(b_1,a_2), \\
\Delta_{(a_1,a_2,a_3)}^{(b_1,b_2,b_3)} F_3 &= F_3(b_1,b_2,b_3) + F_3(b_1,a_2,a_3) + F_3(a_1,b_2,a_3) + F_3(a_1,a_2,b_3) \\
&\quad - F_3(b_1,b_2,a_3) - F_3(b_1,a_2,b_3) - F_3(a_1,b_2,b_3) - F_3(a_1,a_2,a_3)
\end{aligned}$$

for all $a_i, b_i \in \mathbb{R}$, $i = 1,2,3$. The following remark justifies the name "d-fold increase" in a sense.

Remark 2.2.2 It is easy to see that

$$\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F = D_{a_d}^{b_d} \cdots D_{a_1}^{b_1} F$$

for any $\boldsymbol{a} = (a_1,\dots,a_d)$ and $\boldsymbol{b} = (b_1,\dots,b_d)$ in $\mathbb{R}^d$, where

$$D_{a_1}^{b_1} F := F_2^{\boldsymbol{a},\boldsymbol{b}}, \quad D_{a_2}^{b_2} F_2^{\boldsymbol{a},\boldsymbol{b}} := F_3^{\boldsymbol{a},\boldsymbol{b}}, \quad \dots, \quad D_{a_{d-1}}^{b_{d-1}} F_{d-1}^{\boldsymbol{a},\boldsymbol{b}} := F_d^{\boldsymbol{a},\boldsymbol{b}}, \quad D_{a_d}^{b_d} F_d^{\boldsymbol{a},\boldsymbol{b}} := \Delta_{a_d}^{b_d} F_d^{\boldsymbol{a},\boldsymbol{b}}$$

with

$$\begin{aligned}
F_2^{\boldsymbol{a},\boldsymbol{b}}(x_2,\dots,x_d) &:= F(b_1,x_2,\dots,x_d) - F(a_1,x_2,\dots,x_d) &&, \quad (x_2,\dots,x_d)\in\mathbb{R}^{d-1} \\
F_3^{\boldsymbol{a},\boldsymbol{b}}(x_3,\dots,x_d) &:= F_2^{\boldsymbol{a},\boldsymbol{b}}(b_2,x_3,\dots,x_d) - F_2^{\boldsymbol{a},\boldsymbol{b}}(a_2,x_3,\dots,x_d) &&, \quad (x_3,\dots,x_d)\in\mathbb{R}^{d-2} \\
&\;\;\vdots \\
F_d^{\boldsymbol{a},\boldsymbol{b}}(x_d) &:= F_{d-1}^{\boldsymbol{a},\boldsymbol{b}}(b_{d-1},x_d) - F_{d-1}^{\boldsymbol{a},\boldsymbol{b}}(a_{d-1},x_d) &&, \quad x_d\in\mathbb{R}^1.
\end{aligned}$$

In particular, $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F = 0$ when $a_i = b_i$ for at least one $i \in \{1,\dots,d\}$. ◇

Definition 2.2.3 *A function* $F : \mathbb{R}^d \to \mathbb{R}$ *is said to be* d-fold monotonically increasing *when* $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F \geq 0$ *for all* $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ *with* $\boldsymbol{a} < \boldsymbol{b}$. *It is said to be* d-fold constant *when* $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F = 0$ *for all* $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ *with* $\boldsymbol{a} < \boldsymbol{b}$.

Note that in dimension $d = 1$, a function is d-fold monotonically increasing (resp. d-fold constant) if and only if it is monotonically increasing (resp. constant) in the conventional sense, because the 1-fold increase $\Delta_a^b F$ coincides with the conventional increase $F(b) - F(a)$. In higher dimensions the situation is different. For $d \geq 2$ a d-fold monotonically increasing function $F : \mathbb{R}^d \to \mathbb{R}$ is not necessarily monotonically increasing in the sense that $F(\boldsymbol{a}) \leq F(\boldsymbol{b})$ for all $\boldsymbol{a} < \boldsymbol{b}$, and vice versa; see Examples 2.2.5–2.2.6. Also, for $d \geq 2$ a d-fold constant function $F : \mathbb{R}^d \to \mathbb{R}$ is not necessarily constant in the sense that $F \equiv c$ for some constant $c \in \mathbb{R}$, which can be seen from part (iii) of Proposition 2.2.7. These observations correspond to the fact that for $d \geq 2$ we do *not* have $F(\boldsymbol{b}) - F(\boldsymbol{a}) = \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F$ in general. We rather have the following representation (2.6), where for any nonempty subset $J \subseteq \{1, \ldots, d\}$ we use the notation

$$\boldsymbol{c}_J := (c_j)_{j \in J} \quad (\in \mathbb{R}^J) \qquad \text{for any } \boldsymbol{c} = (c_1, \ldots, c_d) \in \mathbb{R}^d. \tag{2.3}$$

Moreover, we define the function $F^{\boldsymbol{a};J} : \mathbb{R}^{|J|} \to \mathbb{R}$ by

$$F^{\boldsymbol{a};J}(\boldsymbol{x}) \;:=\; F(\boldsymbol{x}_J \boldsymbol{a}), \qquad \boldsymbol{x} = (x_j)_{j \in J} \in \mathbb{R}^J \tag{2.4}$$

with

$$\boldsymbol{x}_J \boldsymbol{a} := (x_J a_1, \ldots, x_J a_d) \quad \text{and} \quad x_J a_j := \begin{cases} x_j & , \ j \in J \\ a_j & , \ j \notin J \end{cases} \tag{2.5}$$

for any $\boldsymbol{a} = (a_1, \ldots, a_d) \in \mathbb{R}^d$, $\boldsymbol{x} = (x_j)_{j \in J} \in \mathbb{R}^J$, and any nonempty subset $J \subseteq \{1, \ldots, d\}$. Note that the statement of Lemma 2.2.4 can also be found as Proposition 6 in [59] (note that $\Delta_{\boldsymbol{a}_J}^{\boldsymbol{b}_J} F^{\boldsymbol{a};J} = (-1)^{|J|} \Delta_{\boldsymbol{b}_J}^{\boldsymbol{a}_J} F^{\boldsymbol{a};J}$) and as formula (8) in the proof of Theorem 2 in [51].

Lemma 2.2.4 *For any function* $F : \mathbb{R}^d \to \mathbb{R}$ *and any* $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ *we have*

$$F(\boldsymbol{b}) \;=\; F(\boldsymbol{a}) \;+ \sum_{\emptyset \neq J \subseteq \{1,\ldots,d\}} \Delta_{\boldsymbol{a}_J}^{\boldsymbol{b}_J} F^{\boldsymbol{a};J}, \tag{2.6}$$

where $F^{\boldsymbol{a};J}$ *is defined by (2.4).*

Proof We will proceed by an induction on d. For dimension $d = 1$, that is, for $F : \mathbb{R} \to \mathbb{R}$, we clearly have

$$F(b_1) \;=\; F(a_1) + (F(b_1) - F(a_1)) \;=\; F(a_1) + \Delta_{a_1}^{b_1} F.$$

Now, assume that (2.6) holds up to dimension $d - 1$ with $d - 1 \geq 1$. Let $F : \mathbb{R}^d \to \mathbb{R}$ and $\boldsymbol{a} = (a_1, \ldots, a_d), \boldsymbol{b} = (b_1, \ldots, b_d) \in \mathbb{R}^d$. Then

$\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F$

$$
\begin{aligned}
&= \Delta_{(a_1,\ldots,a_{d-1},a_d)}^{(b_1,\ldots,b_{d-1},b_d)} F \\
&= \Delta_{(a_1,\ldots,a_{d-1})}^{(b_1,\ldots,b_{d-1})} F(\,\cdot\,,\ldots,\,\cdot\,,b_d) - \Delta_{(a_1,\ldots,a_{d-1})}^{(b_1,\ldots,b_{d-1})} F(\,\cdot\,,\ldots,\,\cdot\,,a_d) \\
&= \Delta_{(a_1,\ldots,a_{d-1})}^{(b_1,\ldots,b_{d-1})} F^{(a_1,\ldots,a_{d-1},b_d);\{1,\ldots,d-1\}} - \Delta_{(a_1,\ldots,a_{d-1})}^{(b_1,\ldots,b_{d-1})} F^{\boldsymbol{a};\{1,\ldots,d-1\}} \\
&= \Big(F^{(a_1,\ldots,a_{d-1},b_d);\{1,\ldots,d-1\}}(b_1,\ldots,b_{d-1}) - F^{(a_1,\ldots,a_{d-1},b_d);\{1,\ldots,d-1\}}(a_1,\ldots,a_{d-1}) \\
&\quad - \sum_{\emptyset\neq J\subsetneq\{1,\ldots,d-1\}} \Delta_{\boldsymbol{a}_J}^{\boldsymbol{b}_J} \big(F^{(a_1,\ldots,a_{d-1},b_d);\{1,\ldots,d-1\}}\big)^{(a_1,\ldots,a_{d-1});J} \Big) - \Delta_{(a_1,\ldots,a_{d-1})}^{(b_1,\ldots,b_{d-1})} F^{\boldsymbol{a};\{1,\ldots,d-1\}} \\
&= \Big(F(b_1,\ldots,b_{d-1},b_d) - F(a_1,\ldots,a_{d-1},b_d) \\
&\quad - \sum_{\emptyset\neq J\subsetneq\{1,\ldots,d-1\}} \Delta_{\boldsymbol{a}_J}^{\boldsymbol{b}_J} F^{(a_1,\ldots,a_{d-1},b_d);J} \Big) - \Delta_{(a_1,\ldots,a_{d-1})}^{(b_1,\ldots,b_{d-1})} F^{\boldsymbol{a};\{1,\ldots,d-1\}},
\end{aligned}
$$

where we used the induction assumption for the fourth equality. Adding the telescoping sums $\sum_{\emptyset\neq J\subsetneq\{1,\ldots,d-1\}} \Delta_{\boldsymbol{a}_J}^{\boldsymbol{b}_J} F^{\boldsymbol{a};J} - \sum_{\emptyset\neq J\subsetneq\{1,\ldots,d-1\}} \Delta_{\boldsymbol{a}_J}^{\boldsymbol{b}_J} F^{\boldsymbol{a};J}$ and $F(\boldsymbol{a}) - F(\boldsymbol{a})$, we may continue with

$$
\begin{aligned}
&= F(b_1,\ldots,b_{d-1},b_d) \\
&\quad - \sum_{\emptyset\neq J\subsetneq\{1,\ldots,d-1\}} \Big(\Delta_{\boldsymbol{a}_J}^{\boldsymbol{b}_J} F^{(a_1,\ldots,a_{d-1},b_d);J} - \Delta_{\boldsymbol{a}_J}^{\boldsymbol{b}_J} F^{(a_1,\ldots,a_{d-1},a_d);J} \Big) - \sum_{\emptyset\neq J\subsetneq\{1,\ldots,d-1\}} \Delta_{\boldsymbol{a}_J}^{\boldsymbol{b}_J} F^{\boldsymbol{a};J} \\
&\quad - \big(F(a_1,\ldots,a_{d-1},b_d) - F(a_1,\ldots,a_{d-1},a_d) \big) \\
&\quad - \Delta_{(a_1,\ldots,a_{d-1})}^{(b_1,\ldots,b_{d-1})} F^{\boldsymbol{a};\{1,\ldots,d-1\}} - F(a_1,\ldots,a_{d-1},a_d) \\
&= F(b_1,\ldots,b_{d-1},b_d) \\
&\quad - \sum_{\emptyset\neq J\subsetneq\{1,\ldots,d-1\}} \Delta_{(\boldsymbol{a}_J,a_d)}^{(\boldsymbol{b}_J,b_d)} F^{\boldsymbol{a};J\cup\{d\}} - \sum_{\emptyset\neq J\subsetneq\{1,\ldots,d-1\}} \Delta_{\boldsymbol{a}_J}^{\boldsymbol{b}_J} F^{\boldsymbol{a};J} \\
&\quad - \Delta_{a_d}^{b_d} F^{\boldsymbol{a};\{d\}} \\
&\quad - \Delta_{(a_1,\ldots,a_{d-1})}^{(b_1,\ldots,b_{d-1})} F^{\boldsymbol{a};\{1,\ldots,d-1\}} - F(a_1,\ldots,a_{d-1},a_d) \\
&= F(\boldsymbol{b}) - \sum_{\emptyset\neq J\subsetneq\{1,\ldots,d\}} \Delta_{\boldsymbol{a}_J}^{\boldsymbol{b}_J} F^{\boldsymbol{a};J} - F(\boldsymbol{a}).
\end{aligned}
$$

Since $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F$ is nothing but $\Delta_{\boldsymbol{a}_J}^{\boldsymbol{b}_J} F^{\boldsymbol{a};J}$ for $J = \{1,\ldots,d\}$, the proof is complete. □

Example 2.2.5 The function $F : \mathbb{R}^d \to \mathbb{R}$ defined by

$$
F(x_1,\ldots,x_d) := \prod_{i=1}^{d} x_i, \qquad (x_1,\ldots,x_d) \in \mathbb{R}^d \tag{2.7}
$$

is d-fold monotonically increasing, because $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F = \prod_{i=1}^{d}(b_i - a_i) \geq 0$ for all $\boldsymbol{a} = (a_1,\ldots,a_d)$ and $\boldsymbol{b} = (b_1,\ldots,b_d)$ with $\boldsymbol{a} < \boldsymbol{b}$. However, for $d \geq 2$ it is not monotonically increasing in the sense that $F(\boldsymbol{a}) \leq F(\boldsymbol{b})$ for all $\boldsymbol{a} < \boldsymbol{b}$. For instance, $\boldsymbol{a} < \boldsymbol{b}$ but $F(\boldsymbol{b}) < F(\boldsymbol{a})$ when choosing $\boldsymbol{a} := (-1,-1,1,\ldots,1)$ and $\boldsymbol{b} := (0,1,1,\ldots,1)$. ◇

For the following Example 2.2.6 note that

$$\begin{aligned} F(\boldsymbol{b}) - F(\boldsymbol{a}) &= (F(b_1, b_2, \dots, b_d) - F(a_1, b_2, \dots, b_d)) \\ &\quad + (F(a_1, b_2, \dots, b_d) - F(a_1, a_2, \dots, b_d)) \\ &\quad \dots \\ &\quad + (F(a_1, a_2, \dots, b_d) - F(a_1, a_2, \dots, a_d)) \end{aligned} \tag{2.8}$$

holds for any function $F : \mathbb{R}^d \to \mathbb{R}$ and $\boldsymbol{a} = (a_1, \dots, a_d)$ and $\boldsymbol{b} = (b_1, \dots, b_d)$ in $\mathbb{R}^d$, which implies that F is monotonically increasing in the sense that $F(\boldsymbol{a}) \le F(\boldsymbol{b})$ for all $\boldsymbol{a} < \boldsymbol{b}$ if (and only if) F is monotonically increasing in each of the d coordinates (when the other respective $d-1$ coordinates are fixed).

Example 2.2.6 Generalizing Example 1.8 in [39], let for some even number $d \ge 2$ the function $F : \mathbb{R}^d \to \mathbb{R}$ be defined by

$$F(x_1, \dots, x_d) := \begin{cases} \prod_{i=1}^d (x_i + 1) & , \quad \sum_{i=1}^d x_i < 0 \\ \prod_{i=1}^d (x_i + 2) & , \quad \sum_{i=1}^d x_i \ge 0 \end{cases}$$

for $(x_1, \dots, x_d) \in [-1, 1]^d$, and by

$$F(x_1, \dots, x_d) := F(\min\{\max\{x_1; -1\}; 1\}, \dots, \min\{\max\{x_d; -1\}; 1\})$$

for $(x_1, \dots, x_d) \notin [-1, 1]^d$. On the one hand, the function F is monotonically increasing in the sense that $F(\boldsymbol{a}) \le F(\boldsymbol{b})$ for all $\boldsymbol{a} < \boldsymbol{b}$. To see this, note that for any fixed $j \in \{1, \dots, d\}$ and $(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_d) \in \mathbb{R}^{d-1}$ the mapping $x \mapsto F(x_1, \dots, x_{j-1}, x, x_{j+1}, \dots, x_d)$ from $\mathbb{R}$ to $\mathbb{R}$ is monotonically increasing, and thus, in view of (2.8), the mapping $\boldsymbol{x} \mapsto F(\boldsymbol{x})$ from $\mathbb{R}^d$ to $\mathbb{R}$ is monotonically increasing. On the other hand, F is not d-fold monotonically increasing. For instance, $\boldsymbol{a} < \boldsymbol{b}$ but $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F < 0$ when choosing $\boldsymbol{a} := (-1, 0, \dots, 0)$ and $\boldsymbol{b} := (0, 1, \dots, 1)$.

Indeed, for any $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ we observe that

$$\begin{aligned} \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F &= \sum_{J \subseteq \{1,\dots,d\}} (-1)^{d-|J|} F(\boldsymbol{b}^{\boldsymbol{a};J}) \\ &= \sum_{k=0}^{d} (-1)^{d-k} \sum_{J \subseteq \{1,\dots,d\}, |J|=k} F(\boldsymbol{b}^{\boldsymbol{a};J}) \\ &= (-1)^d F(\boldsymbol{a}) + \sum_{k=1}^{d} (-1)^{d-k} \sum_{J \subseteq \{1,\dots,d\}, |J|=k, 1 \in J} F(\boldsymbol{b}^{\boldsymbol{a};J}) \\ &\quad + \sum_{k=1}^{d-1} (-1)^{d-k} \sum_{J \subseteq \{1,\dots,d\}, |J|=k, 1 \notin J} F(\boldsymbol{b}^{\boldsymbol{a};J}). \end{aligned}$$

If $\boldsymbol{a} := (-1, 0, \ldots, 0)$ and $\boldsymbol{b} := (0, 1, \ldots, 1)$, we thus obtain

$$
\begin{aligned}
&\Delta_{(-1,0,\ldots,0)}^{(0,1,\ldots,1)} F \\
&= (-1)^d \prod_{j=1}^{d} (a_j + 1) + \sum_{k=1}^{d} (-1)^{d-k} \sum_{J \subseteq \{1,\ldots,d\}, |J|=k, 1 \in J} (b_1 + 2) \prod_{i \in J \setminus \{1\}} (b_i + 2) \prod_{j \notin J} (a_j + 2) \\
&\quad + \sum_{k=1}^{d-1} (-1)^{d-k} \sum_{J \subseteq \{1,\ldots,d\}, |J|=k, 1 \notin J} (a_1 + 2) \prod_{i \in J} (b_i + 2) \prod_{j \notin (J \cup \{1\})} (a_j + 2) \\
&= 0 + \sum_{k=1}^{d} (-1)^{d-k} \sum_{J \subseteq \{1,\ldots,d\}, |J|=k, 1 \in J} 2 \cdot 3^{k-1} \cdot 2^{d-k} \\
&\quad + \sum_{k=1}^{d-1} (-1)^{d-k} \sum_{J \subseteq \{1,\ldots,d\}, |J|=k, 1 \notin J} 1 \cdot 3^{k} \cdot 2^{d-k-1} \\
&= \sum_{k=1}^{d} (-1)^{d-k} \binom{d-1}{k-1} 3^{k-1}\, 2^{d-k+1} + \sum_{k=1}^{d-1} (-1)^{d-k} \binom{d-1}{k} 3^{k}\, 2^{d-k-1} \\
&= 2 \sum_{k=0}^{d-1} \binom{d-1}{k} 3^{k} (-2)^{d-1-k} - \sum_{k=0}^{d-1} \binom{d-1}{k} 3^{k} (-2)^{d-1-k} - (-1)^d\, 2^{d-1}.
\end{aligned}
$$

By means of the Binomial theorem we arrive at

$$\Delta_{(-1,0,\ldots,0)}^{(0,1,\ldots,1)} = 2\,(3-2)^{d-1} - (3-2)^{d-1} - (-1)^d\, 2^{d-1} = 1 - (-1)^d\, 2^{d-1},$$

which is negative for an even number $d \geq 2$. ◇

Proposition 2.2.7 *For any functions $F, G : \mathbb{R}^d \to \mathbb{R}$ the following statements hold:*

(i) *If F and G are d-fold monotonically increasing, then the same is true for $\alpha F + \beta G$ for any $\alpha, \beta \geq 0$.*

(ii) *If F and G are d-fold constant, then the same is true for $\alpha F + \beta G$ for any $\alpha, \beta \in \mathbb{R}$.*

(iii) *F is d-fold constant if it is constant in at least one component, that is, if for at least one $i \in \{1, \ldots, d\}$*

$$F(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_d) = F(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_d) \tag{2.9}$$

holds for all $x_1, \ldots, x_d \in \mathbb{R}$.

(iv) *If F is d-fold constant, then it can be represented as the sum of d functions $F_1, \ldots, F_d$ with F_i being independent of the i-th component, that is, there exist functions $F_1, \ldots, F_d : \mathbb{R}^d \to \mathbb{R}$ such that*

$$F(x_1, \ldots, x_d) = \sum_{i=1}^{d} F_i(x_1, \ldots, x_d) \qquad \text{for all } (x_1, \ldots, x_d) \in \mathbb{R}^d \tag{2.10}$$

and $F_i(x_1, \ldots, x_d)$ does not depend on x_i, $i \in \{1, \ldots, d\}$.

(v) *F is d-fold monotonically increasing if it has the representation*

$$F(x_1, \ldots, x_d) = \prod_{i=1}^{d} f_i(x_i) \qquad \text{for all } (x_1, \ldots, x_d) \in \mathbb{R}^d \tag{2.11}$$

for some monotonically increasing functions $f_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, d$.

(vi) *F is d-fold monotonically increasing if it is d times continuously differentiable with nonnegative mixed partial derivative $F^{(d)} := \frac{\partial^d F}{\partial x_d \ldots \partial x_1}$.*

Proof Assertions (i) and (ii) are obvious, and assertion (iv) is known from (5.26) in [64, p. 37].

(iii): If F is constant in the i-th component, then the difference

$$\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F = \Delta_{(a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_d)}^{(b_1,\ldots,b_{i-1},b_{i+1},\ldots,b_d)} F^{\boldsymbol{b};\{1,\ldots,i-1,i+1,\ldots,d\}} - \Delta_{(a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_d)}^{(b_1,\ldots,b_{i-1},b_{i+1},\ldots,b_d)} F^{\boldsymbol{a};\{1,\ldots,i-1,i+1,\ldots,d\}}$$

vanishes, where the functions $F^{\boldsymbol{b};\{1,\ldots,i-1,i+1,\ldots,d\}}$ and $F^{\boldsymbol{a};\{1,\ldots,i-1,i+1,\ldots,d\}}$ are defined as in (2.4). Hence, F is indeed d-fold constant.

(v): If (2.11) holds, then we have $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F = \prod_{i=1}^{d}(f_i(b_i) - f_i(a_i))$ for all $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$. This shows that F is d-fold monotonically increasing.

(vi): If F is d times continuously differentiable, then we may apply d times the fundamental theorem of calculus to obtain

$$\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F = \int_{a_d}^{b_d} \cdots \int_{a_1}^{b_1} F^{(d)}(x_1, \ldots, x_d)\, dx_1 \cdots dx_d \tag{2.12}$$

for every $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$ (cf. (15) in [59]). Since $F^{(d)}$ is nonnegative by assumption, this implies (vi). □

If we fix some arguments of a d-fold monotonically increasing function $F : \mathbb{R}^d \to \mathbb{R}$ and regard it as a new function in the remaining arguments, then the new function is not necessarily multi-monotonically increasing. The following simple Example 2.2.8 shows that if $F : \mathbb{R}^d \to \mathbb{R}$ is d-fold monotonically increasing, then the function $F^{\boldsymbol{a};J}$ defined by (2.4) is not necessarily $|J|$-fold monotonically increasing. Such a situation can also be derived from part (iii) of Proposition 2.2.7. Indeed, pick any function $G : \mathbb{R}^p \to \mathbb{R}$, with $p < d$, that is *not* p-fold monotonically increasing and regard it as a function from $\mathbb{R}^d$ to $\mathbb{R}$.

Example 2.2.8 In Example 2.2.5 we have seen that the function $F : \mathbb{R}^d \to \mathbb{R}$ defined by (2.7) is d-fold monotonically increasing. However, for $\boldsymbol{a} = (a_1, \ldots, a_d)$ the function $F^{\boldsymbol{a};J} : \mathbb{R}^{|J|} \to \mathbb{R}$ is not $|J|$-fold monotonically increasing when $|J| < d$ and $p := \prod_{j \notin J} a_j < 0$. Indeed, the mapping $(x_j)_{j \in J} \mapsto \prod_{j \in J} x_j$ is $|J|$-fold monotonically increasing by Example

2.2.5, which implies that the mapping $(x_j)_{j\in J} \mapsto F^{\boldsymbol{a};J}((x_j)_{j\in J}) = p\prod_{j\in J} x_j$ cannot have this property. $\diamond$

However, under an additional assumption on the function F we obtain that also the functions $F^{\boldsymbol{a};J}$ are multi-monotonically increasing. For instance, the distribution function of every Borel probability measure on $\mathbb{R}^d$ satisfies the assumptions of the following lemma. Here we use the notation $J^{\mathsf{c}} := \{1,\dots,d\}\setminus J$ as well as (2.2).

Lemma 2.2.9 *Let $F:\mathbb{R}^d\to\mathbb{R}$ be a d-fold monotonically increasing function, $\boldsymbol{a}\in\mathbb{R}^d$, and $J\subsetneq\{1,\dots,d\}$ be nonempty. Assume that $F(x_1,\dots,x_d)\to 0$ as $(x_j)_{j\in J^{\mathsf{c}}}\to(-\infty)_{j\in J^{\mathsf{c}}}$ for any $(x_j)_{j\in J}\in\mathbb{R}^J$. Then the function $F^{\boldsymbol{a};J}:\mathbb{R}^J\to\mathbb{R}$ defined by (2.4) is $|J|$-fold monotonically increasing.*

Proof To prove that $F^{\boldsymbol{a};J}$ is $|J|$-fold monotonically increasing we have to show that $\Delta_{\boldsymbol{u}}^{\boldsymbol{v}}F^{\boldsymbol{a};J}\geq 0$ for any $\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^J$ with $\boldsymbol{u}<\boldsymbol{v}$. For any fixed $\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^J$ with $\boldsymbol{u}<\boldsymbol{v}$, let $\widetilde{\boldsymbol{u}}=(\widetilde{u}_1,\dots,\widetilde{u}_d)$ and $\widetilde{\boldsymbol{v}}=(\widetilde{v}_1,\dots,\widetilde{v}_d)$ be defined by

$$\widetilde{u}_j := \begin{cases} u_j & ,\ j\in J\\ x_j & ,\ j\notin J\end{cases} \qquad\text{and}\qquad \widetilde{v}_j := \begin{cases} v_j & ,\ j\in J\\ a_j & ,\ j\notin J\end{cases},$$

where x_j is (arbitrarily) chosen such that $x_j<a_j$, $j\in J^{\mathsf{c}}$. In particular, $\widetilde{\boldsymbol{u}}<\widetilde{\boldsymbol{v}}$. Then we obtain

$$\begin{aligned}\Delta_{\boldsymbol{u}}^{\boldsymbol{v}}F^{\boldsymbol{a};J} &= \sum_{K\subseteq J}(-1)^{|J|-|K|}\,F^{\boldsymbol{a};J}\big(\boldsymbol{v}^{\boldsymbol{u};K}\big)\\ &= \sum_{K\subseteq J}(-1)^{|J|-|K|}\,F\big(\widetilde{\boldsymbol{v}}^{\widetilde{\boldsymbol{u}};K\cup(\{1,\dots,d\}\setminus J)}\big)\\ &= \sum_{K\subseteq J}(-1)^{d-(|K|+d-|J|)}\,F\big(\widetilde{\boldsymbol{v}}^{\widetilde{\boldsymbol{u}};K\cup(\{1,\dots,d\}\setminus J)}\big)\\ &= \Delta_{\widetilde{\boldsymbol{u}}}^{\widetilde{\boldsymbol{v}}}F-\sum_{L\in\mathbb{L}(J)}(-1)^{d-|L|}\,F\big(\widetilde{\boldsymbol{v}}^{\widetilde{\boldsymbol{u}};L}\big),\end{aligned}$$

where $\mathbb{L}(J)$ consists of all subsets $L\subseteq\{1,\dots,d\}$ that do *not* contain all of the elements of J^{c}. Our assumptions imply that $\sum_{L\in\mathbb{L}(J)}(-1)^{d-|L|}\,F(\widetilde{\boldsymbol{v}}^{\widetilde{\boldsymbol{u}};L})$ converges to 0 as $(x_j)_{j\in J^{\mathsf{c}}}\to(-\infty)_{j\in J^{\mathsf{c}}}$. Thus, since $\Delta_{\widetilde{\boldsymbol{u}}}^{\widetilde{\boldsymbol{v}}}F\geq 0$ holds for each specific choice of $(x_j)_{j\in J^{\mathsf{c}}}$ (recall that F is d-fold monotonically increasing), we indeed get $\Delta_{\boldsymbol{u}}^{\boldsymbol{v}}F^{\boldsymbol{a};J}\geq 0$. $\square$

Definition 2.2.10 *A function $F:\mathbb{R}^d\to\mathbb{R}$ is said to be* completely monotonically increasing *if all functions $F^{\boldsymbol{a};J}$, $\boldsymbol{a}\in\mathbb{R}^d$, $\emptyset\neq J\subseteq\{1,\dots,d\}$, are multi-monotonically increasing.*

Note that completely monotonically increasing functions are Borel measurable; one can argue as in Theorem 3.2 of [2] where the case of functions on compact intervals is treated. As an immediate consequence of Lemma 2.2.9 we obtain the following corollary, which shows in particular that the distribution function of every Borel probability measure on $\mathbb{R}^d$ is completely monotonically increasing as these distribution functions satisfy the assumptions of Lemma 2.2.9.

Corollary 2.2.11 *Let $F : \mathbb{R}^d \to \mathbb{R}$ be a d-fold monotonically increasing function. Assume that $F(x_1, \dots, x_d) \to 0$ as $(x_j)_{j\in J^c} \to (-\infty)_{j\in J^c}$ for any $(x_j)_{j\in J} \in \mathbb{R}^J$ and any nonempty $J \subsetneq \{1, \dots, d\}$. Then F is completely monotonically increasing.*

2.3 Measure generating functions

Let $F : \mathbb{R}^d \to \mathbb{R}$ be any d-fold monotonically increasing function. Denote by $\mathcal{I}^d$ the class of all sets $(\boldsymbol{a}, \boldsymbol{b}]$ in $\mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$, and consider the set function $\mu_{F,\mathcal{I}^d} : \mathcal{I}^d \to \mathbb{R}_+$ defined by

$$\mu_{F,\mathcal{I}^d}((\boldsymbol{a}, \boldsymbol{b}]) := \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F, \qquad \boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d \text{ with } \boldsymbol{a} < \boldsymbol{b}.$$

Theorem 2.3.2 below shows that $\mu_{F,\mathcal{I}^d}$ extends in a unique manner to a positive measure on $\mathcal{B}(\mathbb{R}^d)$ when F is in addition right continuous.

Definition 2.3.1 *A function $F : \mathbb{R}^d \to \mathbb{R}$ is said to be* right continuous *if it is coordinatewise right continuous in each coordinate, at every point $\boldsymbol{x} \in \mathbb{R}^d$.*

Theorem 2.3.2 *For any d-fold monotonically increasing and right continuous function $F : \mathbb{R}^d \to \mathbb{R}$ there exists a unique positive measure μ_F on $\mathcal{B}(\mathbb{R}^d)$ whose restriction to $\mathcal{I}^d$ coincides with $\mu_{F,\mathcal{I}^d}$.*

The preceding result can be found in Theorem I.5.27 of [64] and justifies the following definition.

Definition 2.3.3 *A function $F : \mathbb{R}^d \to \mathbb{R}$ is said to be a* measure generating function *if it is d-fold monotonically increasing and right continuous. In this case, the measure μ_F given by Theorem 2.3.2 is said to be the measure generated by F.*

Of course, the measure μ_F generated by a measure generating function F is finite when F is bounded. Conversely, for a finite measure μ on $\mathcal{B}(\mathbb{R}^d)$ we obtain by

$$F_\mu(\boldsymbol{x}) := \lim_{n\to\infty} \mu((\boldsymbol{a}_n, \boldsymbol{x}]), \qquad \boldsymbol{x} \in \mathbb{R}^d, \tag{2.13}$$

(for any $(\boldsymbol{a}_n)$ with $\lim_{n\to\infty} \|\boldsymbol{a}_n\| = \infty$ and $\boldsymbol{0} > \boldsymbol{a}_1 > \boldsymbol{a}_2 > \cdots$) a bounded measure generating function $F_\mu : \mathbb{R}^d \to \mathbb{R}_+$. That is, we have a one-to-one correspondence between a *finite* measure μ on $\mathcal{B}(\mathbb{R}^d)$ and a *bounded* measure generating function $F : \mathbb{R}^d \to \mathbb{R}_+$.

Definition 2.3.4 *For a finite measure μ on $\mathcal{B}(\mathbb{R}^d)$ the function F_μ defined in (2.13) is also referred to as corresponding* distribution function.

It is easily seen that the distribution function F_μ of a finite measure μ on $\mathcal{B}(\mathbb{R}^d)$ satisfies $F_\mu(x_1, \ldots, x_d) \to 0$ as $(x_j)_{j\in J} \to (-\infty)_{j\in J}$ for any $(x_j)_{j\in J^c} \in \mathbb{R}^{J^c}$ and any nonempty $J \subseteq \{1, \ldots, d\}$, and thus by Corollary 2.2.11 it is completely monotonically increasing (hence Borel measurable). This is not true for every measure generating function; recall Example 2.2.8.

We emphasize that Theorem 2.3.2 is somewhat different from (the respective special case of) part (a) of Theorem 3 in the recent paper [1]. Whereas the latter treats the case of a finite positive Borel measure on a compact interval and assumes that the "measure generating function" is completely monotonically increasing (in the sense of Definition 2.2.10), the former covers all σ-finite positive Borel measures on $\mathbb{R}^d$ and only requires that the measure generating function is d-fold monotonically increasing. Also, in [1] the "measure generating function" depends on the particular compact interval of interest, whereas in our context the measure generating function can be chosen "globally". For instance, a measure generating function for the Borel Lebesgue measure on $\mathbb{R}^d$ in the sense of Definition 2.3.3 is given by the function F defined in (2.7). Example 2.2.8 shows that this F is not completely monotonically increasing.

2.4 Functions of locally bounded multi-variation

In this section we will first recall the notion of d-fold variation (or Vitali variation) of multivariate functions $F : \mathbb{R}^d \to \mathbb{R}$, and we will show later on that any function $F : \mathbb{R}^d \to \mathbb{R}$ that is locally of bounded d-fold variation can be represented as difference of two d-fold monotonically increasing functions; cf. Theorem 2.4.8 and Corollary 2.4.9.

Definition 2.4.1 *For any $\boldsymbol{a} = (a_1, \ldots, a_d)$ and $\boldsymbol{b} = (b_1, \ldots, b_d)$ in $\mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$, a* grid partition of the interval $[\boldsymbol{a}, \boldsymbol{b}]$ *is a collection*

$$\{(x_{1,i_1}, \ldots, x_{d,i_d})\} \equiv \{(x_{1,i_1}, \ldots, x_{d,i_d}) : 0 \le i_1 \le n_1, \ldots, 0 \le i_d \le n_d\}$$

of points in $[\boldsymbol{a}, \boldsymbol{b}]$ with $a_j = x_{j,0} \le x_{j,1} \le \cdots \le x_{j,n_j-1} \le x_{j,n_j} = b_j$ for $1 \le j \le d$. The set of all such partitions will be denoted by $\mathcal{P}([\boldsymbol{a}, \boldsymbol{b}])$.

For any interval $[\boldsymbol{a}, \boldsymbol{b}]$ and any grid partition $\{(x_{1,i_1}, \ldots, x_{d,i_d})\} \in \mathcal{P}([\boldsymbol{a}, \boldsymbol{b}])$, we have

$$\sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \Delta_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}^{(x_{1,i_1},\ldots,x_{d,i_d})} F \;=\; \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F, \tag{2.14}$$

because the left-hand side in (2.14) is nothing but $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F$ plus some telescoping sum (see also Proposition 3 in [59]). In the following definition, and later on, we will use the notation $z^\pm := \max\{\pm z, 0\}$.

Definition 2.4.2 *Let $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ such that $\boldsymbol{a} < \boldsymbol{b}$. For a function $F : \mathbb{R}^d \to \mathbb{R}$ the* total d-fold variation*, the* positive d-fold variation*, and the* negative d-fold variation *on $[\boldsymbol{a}, \boldsymbol{b}]$ are defined by respectively*

$$
\begin{aligned}
V_F([\boldsymbol{a},\boldsymbol{b}]) &:= \sup_{\{(x_{1,i_1},\ldots,x_{d,i_d})\}\in\mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1}\cdots\sum_{i_d=1}^{n_d} \left|\Delta_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}^{(x_{1,i_1},\ldots,x_{d,i_d})} F\right|,\\
V_F^+([\boldsymbol{a},\boldsymbol{b}]) &:= \sup_{\{(x_{1,i_1},\ldots,x_{d,i_d})\}\in\mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1}\cdots\sum_{i_d=1}^{n_d} \left(\Delta_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}^{(x_{1,i_1},\ldots,x_{d,i_d})} F\right)^+,\\
V_F^-([\boldsymbol{a},\boldsymbol{b}]) &:= \sup_{\{(x_{1,i_1},\ldots,x_{d,i_d})\}\in\mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1}\cdots\sum_{i_d=1}^{n_d} \left(\Delta_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}^{(x_{1,i_1},\ldots,x_{d,i_d})} F\right)^-.
\end{aligned}
$$

It is easily seen that in Definition 2.4.2 the set $\mathcal{P}([\boldsymbol{a},\boldsymbol{b}])$ can be replaced by the set $\widetilde{\mathcal{P}}([\boldsymbol{a},\boldsymbol{b}])$ of arbitrary partitions of $[\boldsymbol{a},\boldsymbol{b}]$ into finitely many disjoint subintervals $[\boldsymbol{\alpha},\boldsymbol{\beta}]$. This was done, for instance, in [51, p. 62].

Definition 2.4.3 *A function $F : \mathbb{R}^d \to \mathbb{R}$ is said to be* locally of bounded d-fold variation *if $V_F([\boldsymbol{a},\boldsymbol{b}]) < \infty$ for all $\boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$.*

In dimension $d = 1$ the notion of locally bounded d-fold variation coincides with the conventional notion of locally bounded variation; observe that the expression $V_F([a, b])$ is nothing but the conventional variation of a function $F : \mathbb{R} \to \mathbb{R}$ on the interval $[a, b]$.

Proposition 2.4.4 *Let $F, G : \mathbb{R}^d \to \mathbb{R}$ be any functions.*

(i) *$V_{F+G}([\boldsymbol{a},\boldsymbol{b}]) \le V_F([\boldsymbol{a},\boldsymbol{b}]) + V_G([\boldsymbol{a},\boldsymbol{b}])$ holds for all $\boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$.*

(ii) *If F and G are locally of bounded d-fold variation, then the same is true for $\alpha F + \beta G$ for any $\alpha, \beta \in \mathbb{R}$.*

(iii) *If F has the representation $F = F_1 - F_2$ for two d-fold monotonically increasing functions $F_1, F_2 : \mathbb{R}^d \to \mathbb{R}$, then it is locally of bounded d-fold variation.*

(iv) *If F is d times continuously differentiable, then it is locally of bounded d-fold variation.*

Proof Assertions (i) and (ii) are obvious.

(iii): If a function $G : \mathbb{R}^d \to \mathbb{R}$ is d-fold monotonically increasing, then it is clearly locally of bounded d-fold variation with $V_G([\boldsymbol{a},\boldsymbol{b}]) = V_G^+([\boldsymbol{a},\boldsymbol{b}]) = \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} G$ for all $\boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$. For two d-fold monotonically increasing functions $F_1, F_2 : \mathbb{R}^d \to \mathbb{R}$ we thus obtain $V_{F_1-F_2}([\boldsymbol{a},\boldsymbol{b}]) \le V_{F_1}([\boldsymbol{a},\boldsymbol{b}]) + V_{F_2}([\boldsymbol{a},\boldsymbol{b}]) = \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_1 + \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_2$ for all $\boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$. Hence $F = F_1 - F_2$ is locally of bounded d-fold variation.

(iv): Let $F^{(d)}(x_1,\ldots,x_d) := \frac{\partial^{(d)} F}{\partial x_d \cdots \partial x_1}(x_1,\ldots,x_d)$, and $\boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$ arbitrary but fixed. Applying d times the fundamental theorem of calculus we obtain

$$\begin{aligned} |\Delta_{\boldsymbol{u}}^{\boldsymbol{v}} F| &= \Big| \int_{u_d}^{v_d} \cdots \int_{u_1}^{v_1} F^{(d)}(x_1,\ldots,x_d)\, dx_1 \cdots dx_d \Big| \\ &\leq \int_{u_d}^{v_d} \cdots \int_{u_1}^{v_1} \big| F^{(d)}(x_1,\ldots,x_d) \big|\, dx_1 \cdots dx_d \end{aligned}$$

for every $\boldsymbol{u},\boldsymbol{v} \in \mathbb{R}^d$ with $\boldsymbol{a} \leq \boldsymbol{u} < \boldsymbol{v} \leq \boldsymbol{b}$. It follows that $V_F([\boldsymbol{a},\boldsymbol{b}])$ is bounded above by the integral $\int_{a_d}^{b_d} \cdots \int_{a_1}^{b_1} |F^{(d)}(x_1,\ldots,x_d)|\, dx_1 \cdots dx_d$. Since the latter integral is finite by the continuity of $F^{(d)}$, we obtain $V_F([\boldsymbol{a},\boldsymbol{b}]) < \infty$. □

The following remark shows that if $F : \mathbb{R}^d \to \mathbb{R}$ is locally of bounded d-fold variation, then the function $F^{\boldsymbol{a};J} : \mathbb{R}^J \to \mathbb{R}$ defined by (2.4) is not necessarily locally of bounded $|J|$-fold variation.

Remark 2.4.5 If we fix some arguments of a function $F : \mathbb{R}^d \to \mathbb{R}$ that is locally of bounded d-fold variation and regard it as a new function, say G, in the remaining arguments, then the new function is not necessarily locally of bounded multi-variation. Indeed, pick any function $G : \mathbb{R}^p \to \mathbb{R}$, with $p < d$, that is *not* locally of bounded p-fold variation and regard it as a function F from $\mathbb{R}^d$ to $\mathbb{R}$ through

$$F(x_1,\ldots,x_p,x_{p+1},\ldots,x_d) := G(x_1,\ldots,x_p), \qquad (x_1,\ldots,x_d) \in \mathbb{R}^d.$$

Then, by part (iii) of Proposition 2.2.7 the function $F : \mathbb{R}^d \to \mathbb{R}$ is d-fold constant and thus by part (iii) of Proposition 2.4.4 also locally of bounded d-fold variation. ◇

Corollary 2.4.9 below will show that also the converse of part (iii) of Proposition 2.4.4 is true: if a function $F : \mathbb{R}^d \to \mathbb{R}$ is locally of bounded d-fold variation, then it can be represented as difference of two d-fold monotonically increasing functions.

For the proof of Theorem 2.4.8 we need the following two lemmas. The first one is a generalization of Lemma 1.16 in [39], and can also be found as Lemma 1 in [59].

Lemma 2.4.6 *Let $\boldsymbol{a} = (a_1,\ldots,a_d), \boldsymbol{y} = (y_1,\ldots,y_d), \boldsymbol{b} = (b_1,\ldots,b_d) \in \mathbb{R}^d$, and assume that $\boldsymbol{a} \leq \boldsymbol{y} \leq \boldsymbol{b}$. Let $\boldsymbol{I}_1,\ldots,\boldsymbol{I}_{2^d}$ denote the 2^d compact intervals of the shape $\times_{j=1}^d I_j$ where for $j = 1,\ldots,d$ either $I_j = [a_j,y_j]$ or $[y_j,b_j]$. Let $F : \mathbb{R}^d \to \mathbb{R}$ be any function. Then $V_F([\boldsymbol{a},\boldsymbol{b}]) < \infty$ implies $V_F(\boldsymbol{I}_i) < \infty$ for $1 \leq i \leq 2^d$ and we have*

$$V_F([\boldsymbol{a},\boldsymbol{b}]) = \sum_{i=1}^{2^d} V_F(\boldsymbol{I}_i) \qquad \textit{and} \qquad V_F^{\pm}([\boldsymbol{a},\boldsymbol{b}]) = \sum_{i=1}^{2^d} V_F^{\pm}(\boldsymbol{I}_i). \tag{2.15}$$

Proof For a grid partition $P = \{(x_{1,i_1}, \ldots, x_{d,i_d}) : 0 \leq i_1 \leq n_1, \ldots, 0 \leq i_d \leq n_d\}$ of any compact interval $\boldsymbol{I}$, we will use the notation

$$V_F(\boldsymbol{I}, P) := \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \left|\Delta_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}^{(x_{1,i_1},\ldots,x_{d,i_d})} F\right|.$$

Let $P_i \in \mathcal{P}(\boldsymbol{I}_i)$ for $1 \leq i \leq 2^d$. Of course, $P_1, \ldots, P_{2^d}$ together form a grid partition $P \in \mathcal{P}([\boldsymbol{a}, \boldsymbol{b}])$. Then

$$V_F([\boldsymbol{a}, \boldsymbol{b}], P) = \sum_{i=1}^{2^d} V_F(\boldsymbol{I}_i, P_i). \tag{2.16}$$

In particular,

$$0 \leq V_F(\boldsymbol{I}_i, P_i) \leq V_F([\boldsymbol{a}, \boldsymbol{b}], P) \leq V_F([\boldsymbol{a}, \boldsymbol{b}]) \tag{2.17}$$

for every $P_i \in \mathcal{P}(\boldsymbol{I}_i)$ and $1 \leq i \leq 2^d$. This implies $V_F(\boldsymbol{I}_i, P_i) < \infty$ for every $1 \leq i \leq 2^d$.

It remains to show (2.15). We will only show the first equation in (2.15). The analogous statements for the positive and negative variation follow by the same arguments. Let $Q \in \mathcal{P}([\boldsymbol{a}, \boldsymbol{b}])$. Then $V_F([\boldsymbol{a}, \boldsymbol{b}], Q) \leq V_F([\boldsymbol{a}, \boldsymbol{b}], P)$ when $P \in \mathcal{P}([\boldsymbol{a}, \boldsymbol{b}])$ is obtained by adjoining $\boldsymbol{y}$ to Q. Since such P can obviously be considered as a grid partition of $[\boldsymbol{a}, \boldsymbol{b}]$ obtained by collecting the grid points of certain grid partitions of $\boldsymbol{I}_1, \ldots, \boldsymbol{I}_{2^d}$, we obtain by (2.16)

$$V_F([\boldsymbol{a}, \boldsymbol{b}], Q) \leq V_F([\boldsymbol{a}, \boldsymbol{b}], P) = \sum_{i=1}^{2^d} V_F(\boldsymbol{I}_i, P_i) \leq \sum_{i=1}^{2^d} V_F(\boldsymbol{I}_i).$$

Since Q was arbitrary, this implies

$$V_F([\boldsymbol{a}, \boldsymbol{b}]) \leq \sum_{i=1}^{2^d} V_F(\boldsymbol{I}_i). \tag{2.18}$$

On the other hand, given any $\varepsilon > 0$, we can find grid partitions $P_1, \ldots, P_{2^d}$ of $\boldsymbol{I}_1, \ldots, \boldsymbol{I}_{2^d}$ respectively such that

$$V_F(\boldsymbol{I}_i) - \frac{\varepsilon}{2^d} < V_F(\boldsymbol{I}_i, P_i) \tag{2.19}$$

for every $i = 1, \ldots, 2^d$. If P denotes the grid partition of $[\boldsymbol{a}, \boldsymbol{b}]$ obtained by collecting the grid points of the grid partitions $P_1, \ldots, P_{2^d}$, then we obtain by (2.19), (2.16), and (2.17) that

$$\sum_{i=1}^{2^d} V_F(\boldsymbol{I}_i) - \varepsilon < \sum_{i=1}^{2^d} V_F(\boldsymbol{I}_i, P_i) = V_F([\boldsymbol{a}, \boldsymbol{b}], P) \leq V_F([\boldsymbol{a}, \boldsymbol{b}]).$$

Since $\varepsilon > 0$ is arbitrary, we conclude $\sum_{i=1}^{2^d} V_F(\boldsymbol{I}_i) \leq V_F([\boldsymbol{a}, \boldsymbol{b}])$. In view of (2.18), this completes the proof of the first equation in (2.15). □

Lemma 2.4.7 *Let* $F : \mathbb{R}^d \to \mathbb{R}$ *and* $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ *with* $\boldsymbol{a} \leq \boldsymbol{b}$. *Then* $V_F([\boldsymbol{a}, \boldsymbol{b}]) < \infty$ *implies*

(i) $V_F([\boldsymbol{a}, \boldsymbol{b}]) = V_F^+([\boldsymbol{a}, \boldsymbol{b}]) + V_F^-([\boldsymbol{a}, \boldsymbol{b}])$,

(ii) $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F = V_F^+([\boldsymbol{a}, \boldsymbol{b}]) - V_F^-([\boldsymbol{a}, \boldsymbol{b}])$,

(iii) $V_F^-([\boldsymbol{a}, \boldsymbol{b}]) = \frac{1}{2}(V_F([\boldsymbol{a}, \boldsymbol{b}]) - \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F)$,

(iv) $V_F^+([\boldsymbol{a}, \boldsymbol{b}]) = \frac{1}{2}(V_F([\boldsymbol{a}, \boldsymbol{b}]) + \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F)$.

Proof Part (ii) follows from

$$
\begin{aligned}
& V_F^+([\boldsymbol{a}, \boldsymbol{b}]) \\
= & \sup_{\{(x_{1,i_1},\dots,x_{d,i_d})\} \in \mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \big(\Delta_{(x_{1,i_1-1},\dots,x_{d,i_d-1})}^{(x_{1,i_1},\dots,x_{d,i_d})} F\big)^+ \\
= & \sup_{\{(x_{1,i_1},\dots,x_{d,i_d})\} \in \mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \Big(\big(\Delta_{(x_{1,i_1-1},\dots,x_{d,i_d-1})}^{(x_{1,i_1},\dots,x_{d,i_d})} F\big)^- + \Delta_{(x_{1,i_1-1},\dots,x_{d,i_d-1})}^{(x_{1,i_1},\dots,x_{d,i_d})} F\Big) \\
= & \sup_{\{(x_{1,i_1},\dots,x_{d,i_d})\} \in \mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \Big(\sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \big(\Delta_{(x_{1,i_1-1},\dots,x_{d,i_d-1})}^{(x_{1,i_1},\dots,x_{d,i_d})} F\big)^- + \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F\Big) \\
= & \sup_{\{(x_{1,i_1},\dots,x_{d,i_d})\} \in \mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \big(\Delta_{(x_{1,i_1-1},\dots,x_{d,i_d-1})}^{(x_{1,i_1},\dots,x_{d,i_d})} F\big)^- + \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F \\
= & V_F^-([\boldsymbol{a}, \boldsymbol{b}]) + \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F,
\end{aligned}
$$

where the third step is valid by (2.14). Using the same argument, we also obtain

$$
\begin{aligned}
& V_F([\boldsymbol{a}, \boldsymbol{b}]) \\
= & \sup_{\{(x_{1,i_1},\dots,x_{d,i_d})\} \in \mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \big|\Delta_{(x_{1,i_1-1},\dots,x_{d,i_d-1})}^{(x_{1,i_1},\dots,x_{d,i_d})} F\big| \\
= & \sup_{\{(x_{1,i_1},\dots,x_{d,i_d})\} \in \mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \Big(\big(\Delta_{(x_{1,i_1-1},\dots,x_{d,i_d-1})}^{(x_{1,i_1},\dots,x_{d,i_d})} F\big)^+ + \big(\Delta_{(x_{1,i_1-1},\dots,x_{d,i_d-1})}^{(x_{1,i_1},\dots,x_{d,i_d})} F\big)^-\Big) \\
= & \sup_{\{(x_{1,i_1},\dots,x_{d,i_d})\} \in \mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \Big(2\big(\Delta_{(x_{1,i_1-1},\dots,x_{d,i_d-1})}^{(x_{1,i_1},\dots,x_{d,i_d})} F\big)^- + \Delta_{(x_{1,i_1-1},\dots,x_{d,i_d-1})}^{(x_{1,i_1},\dots,x_{d,i_d})} F\Big) \\
= & \sup_{\{(x_{1,i_1},\dots,x_{d,i_d})\} \in \mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \Big(\sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} 2\big(\Delta_{(x_{1,i_1-1},\dots,x_{d,i_d-1})}^{(x_{1,i_1},\dots,x_{d,i_d})} F\big)^- + \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F\Big) \\
= & \sup_{\{(x_{1,i_1},\dots,x_{d,i_d})\} \in \mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} 2\big(\Delta_{(x_{1,i_1-1},\dots,x_{d,i_d-1})}^{(x_{1,i_1},\dots,x_{d,i_d})} F\big)^- + \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F \\
= & 2V_F^-([\boldsymbol{a}, \boldsymbol{b}]) + \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F.
\end{aligned}
$$

Together with part (ii) this implies part (i). Equation (iii) can be obtained by subtracting equation (ii) from equation (i), and equation (iv) can be obtained by plugging equation (iii) in equation (ii). $\square$

The following theorem provides a sort of Jordan decomposition for a function of locally bounded d-fold variation. The theorem complements Propositions 2.18 and 2.19 in [46] which provide a similar result in the univariate setting.

In Theorem 2 in [1], Aistleitner and Dick prove a Jordan decomposition for functions on $[0,1]^d$, which enables even the decomposition of a function F in completely monotonically increasing functions under the additional assumption that the functions $F^{(1,\ldots,1);J}$ are of bounded $|J|$-fold variation for every $J \subseteq \{1,\ldots,d\}$. However, only functions on compact intervals can be treated by Theorem 2 in [1] because the Jordan decomposition (more precisely the variation of F and $F^{(1,\ldots,1);J}$ that occurs in the definition of the functions of the Jordan decomposition) is anchored at one of the endpoints of the compact interval. The Jordan decomposition in Theorem 2.4.8 below is centered at an arbitrary point $c \in \mathbb{R}^d$, which enables to deal with functions on $\mathbb{R}^d$.

Theorem 2.4.8 (Jordan decomposition) *Let $F : \mathbb{R}^d \to \mathbb{R}$ be a function that is locally of bounded d-fold variation. For any $\boldsymbol{c} \in \mathbb{R}^d$, let the functions $F_{\boldsymbol{c},+}, F_{\boldsymbol{c},-}, F_{\boldsymbol{c},0} : \mathbb{R}^d \to \mathbb{R}$ be defined by*

$$\begin{aligned} F_{\boldsymbol{c},+}(\boldsymbol{x}) &:= (-1)^{d-|J^{\boldsymbol{c},\boldsymbol{x}}|}\, V_F^+(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{x}}), \qquad \boldsymbol{x} \in \mathbb{R}^d, & (2.20)\\ F_{\boldsymbol{c},-}(\boldsymbol{x}) &:= (-1)^{d-|J^{\boldsymbol{c},\boldsymbol{x}}|}\, V_F^-(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{x}}), \qquad \boldsymbol{x} \in \mathbb{R}^d, & (2.21)\\ F_{\boldsymbol{c},0}(\boldsymbol{x}) &:= \sum\nolimits_{J \subsetneq \{1,\ldots,d\}} (-1)^{d-|J|}\, F(\boldsymbol{x}^{\boldsymbol{c};J}), \qquad \boldsymbol{x} \in \mathbb{R}^d, & (2.22) \end{aligned}$$

where $\boldsymbol{x}^{\boldsymbol{c};J}$ is defined as in (2.2) and we set $J^{\boldsymbol{c},\boldsymbol{x}} := \{j \in \{1,\ldots,d\} : c_j < x_j\}$ as well as

$$\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{x}} := I_1^{\boldsymbol{c},\boldsymbol{x}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{x}} \quad \text{with} \quad I_j^{\boldsymbol{c},\boldsymbol{x}} := \begin{cases} [c_j, x_j] & , \quad x_j \geq c_j \\ [x_j, c_j] & , \quad x_j < c_j \end{cases}. \tag{2.23}$$

Then the following assertions hold:

(i) *The function F has the representation*

$$F = F_{\boldsymbol{c},+} - F_{\boldsymbol{c},-} - F_{\boldsymbol{c},0}. \tag{2.24}$$

(ii) *The function $F_{\boldsymbol{c},0}$ is d-fold constant. Moreover we have $F_{\boldsymbol{c},+}(\boldsymbol{x}) = F_{\boldsymbol{c},-}(\boldsymbol{x}) = 0$ for any $\boldsymbol{x} = (x_1,\ldots,x_d)$ with $x_i = c_i$ for at least one $i \in \{1,\ldots,d\}$ as well as*

$$\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_{\boldsymbol{c},+} = V_F^+([\boldsymbol{a},\boldsymbol{b}]) \quad \text{and} \quad \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_{\boldsymbol{c},-} = V_F^-([\boldsymbol{a},\boldsymbol{b}]) \tag{2.25}$$

for any $\boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$. In particular, the functions $F_{\boldsymbol{c},+}$ and $F_{\boldsymbol{c},-}$ are d-fold monotonically increasing.

(iii) *For* $F_{\boldsymbol{c},0}$ *as defined in (2.22), there do not exist any other functions* $\widetilde{F}_{\boldsymbol{c},+}$ *and* $\widetilde{F}_{\boldsymbol{c},-}$ *satisfying the properties in (i) and (ii).*

(iv) *If* F *is right continuous, then* $F_{\boldsymbol{c},+}, F_{\boldsymbol{c},-}$ *and* $F_{\boldsymbol{c},0}$ *are right continuous.*

Proof (i): By part (ii) of Lemma 2.4.7 we have

$$F_{\boldsymbol{c},+}(\boldsymbol{x}) - F_{\boldsymbol{c},-}(\boldsymbol{x}) \;=\; (-1)^{d-|J^{\boldsymbol{c},\boldsymbol{x}}|}\,\Delta_{\underline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}}}^{\overline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}}} F, \tag{2.26}$$

where $\underline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}}$ and $\overline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}}$ refer to the smallest and the largest element of $\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{x}}$, respectively. By the definition of $J^{\boldsymbol{c},\boldsymbol{x}}$ the right-hand side in (2.26) coincides with $F(\boldsymbol{x}) + F_{\boldsymbol{c},0}(\boldsymbol{x})$, which completes the proof of (2.24).

(ii): For any $\boldsymbol{x} = (x_1, \ldots, x_d)$ with $x_i = c_i$ for at least one $i \in \{1, \ldots, d\}$ the equalities $F_{\boldsymbol{c},+}(\boldsymbol{x}) = F_{\boldsymbol{c},-}(\boldsymbol{x}) = 0$ are trivial. Moreover, part (iii) of Proposition 2.2.7 implies that $F_{\boldsymbol{c},0}$ is d-fold constant. It thus remains to show (2.25). We will only consider the case where for every $i \in \{1, \ldots, d\}$ either $c_i \le a_i < b_i$ or $a_i < b_i \le c_i$ (this is not the same as assuming either $c_i \le a_i < b_i$ for all $i = 1, \ldots, d$, or $a_i < b_i \le c_i$ for all $i = 1, \ldots, d$). In the other case where $a_i < c_i < b_i$ for at least one $i \in \{1, \ldots, d\}$ the assertion can be derived therefrom by considering a grid partition of $[\boldsymbol{a}, \boldsymbol{b}]$ consisting only of intervals of the form just described. Let $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ be as just described and set

$$\boldsymbol{p} := \boldsymbol{p}_{\boldsymbol{a},\boldsymbol{b}} := (p_1, \ldots, p_d) \quad \text{with} \quad p_k := \begin{cases} a_k & , \quad c_k \le a_k < b_k \\ b_k & , \quad a_k < b_k \le c_k \end{cases}$$

as well as

$$\boldsymbol{q} := \boldsymbol{q}_{\boldsymbol{a},\boldsymbol{b}} := (q_1, \ldots, q_d) \quad \text{with} \quad q_k := \begin{cases} b_k & , \quad c_k \le a_k < b_k \\ a_k & , \quad a_k < b_k \le c_k \end{cases} .$$

Note that $\boldsymbol{p}$ (resp. $\boldsymbol{q}$) is just those edge of the rectangle $[\boldsymbol{a}, \boldsymbol{b}]$ with the smallest (resp. largest) distance to $\boldsymbol{c}$ among all edges. The d-fold increase $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_{\boldsymbol{c},\pm}$ of the function $F_{\boldsymbol{c},\pm}$ defined in (2.20)–(2.21) can be rewritten as

$$\begin{aligned} \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_{\boldsymbol{c},\pm} &= \Delta_{\boldsymbol{a}}^{\boldsymbol{b}}\big((-1)^{d-|J^{\boldsymbol{c},\bullet}|}\, V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\bullet})\big) \\ &= \Delta_{\boldsymbol{p}}^{\boldsymbol{q}}\, V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\bullet}) \\ &= \sum_{k=0}^{d} (-1)^{d-k} \sum_{Z \subseteq \{1,\ldots,d\},\, |Z|=k} V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}}), \end{aligned} \tag{2.27}$$

where $\boldsymbol{q}^{\boldsymbol{p};Z}$ is defined as in (2.2). Each of the intervals $\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}}$, $Z \subseteq \{1, \ldots, d\}$, is a finite union of some of the following subintervals of $\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}}$:

$$\begin{aligned} \boldsymbol{I}_1 &:= I_1^{\boldsymbol{c},\boldsymbol{p}} \times I_2^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{p}} \\ \boldsymbol{I}_2 &:= I_1^{\boldsymbol{p},\boldsymbol{q}} \times I_2^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{p}} \end{aligned}$$

$$
\begin{aligned}
&\vdots\\
\boldsymbol{I}_{d+1} &:= I_1^{\boldsymbol{c},\boldsymbol{p}} \times I_2^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{p},\boldsymbol{q}}\\
&\vdots\\
\boldsymbol{I}_{2^d} &:= I_1^{\boldsymbol{p},\boldsymbol{q}} \times I_2^{\boldsymbol{p},\boldsymbol{q}} \times \cdots \times I_d^{\boldsymbol{p},\boldsymbol{q}},
\end{aligned}
$$

where $I_j^{\boldsymbol{x},\boldsymbol{y}}$ is defined as in (2.23). More precisely,

$$
\begin{aligned}
I_1^{\boldsymbol{c},\boldsymbol{p}} \times I_2^{\boldsymbol{c},\boldsymbol{p}} \times I_3^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{p}} &= \boldsymbol{I}_1\\
I_1^{\boldsymbol{c},\boldsymbol{q}} \times I_2^{\boldsymbol{c},\boldsymbol{p}} \times I_3^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{p}} &= \boldsymbol{I}_1 \cup \boldsymbol{I}_2\\
I_1^{\boldsymbol{c},\boldsymbol{p}} \times I_2^{\boldsymbol{c},\boldsymbol{q}} \times I_3^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{p}} &= \boldsymbol{I}_1 \cup \boldsymbol{I}_3\\
&\vdots\\
I_1^{\boldsymbol{c},\boldsymbol{p}} \times I_2^{\boldsymbol{c},\boldsymbol{p}} \times I_3^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{q}} &= \boldsymbol{I}_1 \cup \boldsymbol{I}_{\binom{d}{1}+1}\\
I_1^{\boldsymbol{c},\boldsymbol{q}} \times I_2^{\boldsymbol{c},\boldsymbol{q}} \times I_3^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{p}} &= \boldsymbol{I}_1 \cup \boldsymbol{I}_2 \cup \boldsymbol{I}_3 \cup \boldsymbol{I}_{\binom{d}{1}+2}\\
&\vdots\\
I_1^{\boldsymbol{c},\boldsymbol{q}} \times I_2^{\boldsymbol{c},\boldsymbol{p}} \times I_3^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{q}} &= \boldsymbol{I}_1 \cup \boldsymbol{I}_2 \cup \boldsymbol{I}_{\binom{d}{1}+1} \cup \boldsymbol{I}_{\binom{d}{2}+\binom{d}{1}+1}\\
&\vdots\\
I_1^{\boldsymbol{c},\boldsymbol{q}} \times I_2^{\boldsymbol{c},\boldsymbol{q}} \times I_3^{\boldsymbol{c},\boldsymbol{q}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{q}} &= \boldsymbol{I}_1 \cup \cdots \cup \boldsymbol{I}_{2^d},
\end{aligned}
$$

where the intervals on the left-hand sides are just the intervals $\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}}$, $Z \subseteq \{1, \ldots, d\}$. Setting $W_i := V_F^{\pm}(\boldsymbol{I}_i)$ for $i = 1, \ldots, 2^d$, we obtain by Lemma 2.4.6

$$
\begin{aligned}
V_F^{\pm}(I_1^{\boldsymbol{c},\boldsymbol{p}} \times I_2^{\boldsymbol{c},\boldsymbol{p}} \times I_3^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{p}}) &= W_1\\
V_F^{\pm}(I_1^{\boldsymbol{c},\boldsymbol{q}} \times I_2^{\boldsymbol{c},\boldsymbol{p}} \times I_3^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{p}}) &= W_1 + W_2\\
V_F^{\pm}(I_1^{\boldsymbol{c},\boldsymbol{p}} \times I_2^{\boldsymbol{c},\boldsymbol{q}} \times I_3^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{p}}) &= W_1 + W_3\\
&\vdots\\
V_F^{\pm}(I_1^{\boldsymbol{c},\boldsymbol{p}} \times I_2^{\boldsymbol{c},\boldsymbol{p}} \times I_3^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{q}}) &= W_1 + W_{\binom{d}{1}+1}\\
V_F^{\pm}(I_1^{\boldsymbol{c},\boldsymbol{q}} \times I_2^{\boldsymbol{c},\boldsymbol{q}} \times I_3^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{p}}) &= W_1 + W_2 + W_3 + W_{\binom{d}{1}+2}\\
&\vdots\\
V_F^{\pm}(I_1^{\boldsymbol{c},\boldsymbol{q}} \times I_2^{\boldsymbol{c},\boldsymbol{p}} \times I_3^{\boldsymbol{c},\boldsymbol{p}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{q}}) &= W_1 + W_2 + W_{\binom{d}{1}+1} + W_{\binom{d}{2}+\binom{d}{1}+1}\\
&\vdots\\
V_F^{\pm}(I_1^{\boldsymbol{c},\boldsymbol{q}} \times I_2^{\boldsymbol{c},\boldsymbol{q}} \times I_3^{\boldsymbol{c},\boldsymbol{q}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{q}}) &= W_1 + \cdots + W_{2^d}.
\end{aligned}
$$

To compute $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_{\boldsymbol{c},\pm}$ by means of the representation (2.27), we add up the positive (resp. negative) variations above with the sign depending on the cardinality of Z, that is depending on the number of intervals $I_j^{\boldsymbol{c},\boldsymbol{q}}$ in $\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}}$. Of course, several W_i cancel out each other. To specify them, we classify the variations $V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}})$ listed above in $d+1$ blocks with $|Z|=k$, $k=0,\dots,d$, and count how many times W_i is a summand of $V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}})$ with $|Z|=k$.

We start with W_1. Of course, $Z=\emptyset$ is the unique subset of $\{1,\dots,d\}$ with $|Z|=0$. In this case we have $\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}} = \boldsymbol{I}^{\boldsymbol{c},\boldsymbol{p}}$, hence $V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}}) = V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{p}}) = W_1$. That is, there is exactly one subset Z of $\{1,\dots,d\}$ with $|Z|=0$ for which $V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}})$ contributes a summand W_1. Further, there are exactly $\binom{d}{1}$ subsets Z of $\{1,\dots,d\}$ with $|Z|=1$ for which $V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}})$ contributes a summand W_1. More generally, there are exactly $\binom{d}{k}$ subsets Z of $\{1,\dots,d\}$ with $|Z|=k$ for which $V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}})$ contributes a summand W_1.

We now turn to W_2. There is obviously no subset of $\{1,\dots,d\}$ with $|Z|=0$ for which $V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}})$ contributes a summand W_2. Further, W_2 can appear as summand of $V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}})$ only if the first component of the d-dimensional interval $\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}}$ is given by $I_1^{\boldsymbol{c},\boldsymbol{q}}$. So with $I_1^{\boldsymbol{c},\boldsymbol{q}}\times I_2^{\boldsymbol{c},\boldsymbol{p}}\times\cdots\times I_d^{\boldsymbol{c},\boldsymbol{p}}$ there is exactly one interval $\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}}$ with $|Z|=1$ for which $V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}})$ contributes a summand W_2. More generally, there are exactly $\binom{d-1}{k-1}$ subsets Z of $\{1,\dots,d\}$ with $|Z|=k$ for which $V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}})$ contributes a summand W_2. Indeed, given that the first component of the d-dimensional interval $\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}}$ is $I_1^{\boldsymbol{c},\boldsymbol{q}}$, there are exactly $\binom{d-1}{k-1}$ different set-ups where $k-1$ of the remaining $d-1$ components are $I_j^{\boldsymbol{c},\boldsymbol{q}}$ and the other $d-k$ components are $I_j^{\boldsymbol{c},\boldsymbol{p}}$.

Analogously we obtain that for every $i=3,\dots,\binom{d}{1}+1$ and $k\in\{1,\dots,d\}$ there are exactly $\binom{d-1}{k-1}$ subsets Z of $\{1,\dots,d\}$ with $|Z|=k$ for which $V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}})$ contributes a summand W_i. If we proceed with W_i for $i=\binom{d}{1}+2,\dots,2^d$ in the obvious way, we can conclude that in general there are exactly $\binom{d-i}{k-i}$ subsets Z of $\{1,\dots,d\}$ with $|Z|=k$ for which $V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}})$ contributes summands $W_{\sum_{j=1}^{i}\binom{d}{j-1}+1},\dots,W_{\sum_{j=0}^{i}\binom{d}{j}}$. Thus,

$$
\begin{aligned}
&\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_{\boldsymbol{c},\pm}\\
&= \sum_{k=0}^{d}(-1)^{d-k}\sum_{Z\subseteq\{1,\dots,d\},|Z|=k} V_F^{\pm}(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{q}^{\boldsymbol{p};Z}})\\
&= \sum_{k=0}^{d}(-1)^{d-k}\cdot\Big\{\binom{d}{k}W_1+\sum_{i=1}^{k}\binom{d-i}{k-i}\Big(W_{\sum_{j=1}^{i}\binom{d}{j-1}+1}+\cdots+W_{\sum_{j=0}^{i}\binom{d}{j}}\Big)\Big\}\\
&= \sum_{k=0}^{d}(-1)^{d-k}\binom{d}{k}W_1+\sum_{k=1}^{d}(-1)^{d-k}\binom{d-1}{k-1}W_2+\cdots\\
&\qquad \cdots+\sum_{k=d-1}^{d}(-1)^{d-k}\binom{d-(d-1)}{k-(d-1)}W_{\sum_{j=0}^{d-1}\binom{d}{j}}+\sum_{k=d}^{d}(-1)^{d-k}\binom{d-d}{k-d}W_{\sum_{j=0}^{d}\binom{d}{j}}\\
&= \sum_{k=0}^{d}(-1)^{d-k}\binom{d}{k}W_1+\sum_{k=0}^{d-1}(-1)^{(d-1)-k}\binom{d-1}{k}W_2+\cdots
\end{aligned}
$$

$$\cdots + \sum_{k=0}^{1}(-1)^{1-k}\binom{1}{k}W_{2^{d-1}} + \sum_{k=0}^{0}(-1)^{0-k}\binom{0}{k}W_{2^d}$$
$$= \quad W_{2^d},$$

where the last step is justified by the Binomial theorem. Since $\boldsymbol{I}^{\boldsymbol{p},\boldsymbol{q}} = [\boldsymbol{a}, \boldsymbol{b}]$ by the definition of $\boldsymbol{p}$ and $\boldsymbol{q}$, this implies (2.25).

(iii): Let us suppose that, for $F_{\boldsymbol{c},0}$ as defined in (2.22), the functions $F_{\boldsymbol{c},+}$ and $F_{\boldsymbol{c},-}$ in the Jordan decomposition are not uniquely determined by the properties in (i) and (ii). Then there exist functions $\widetilde{F}_{\boldsymbol{c},\pm}$ having the same properties as $F_{\boldsymbol{c},\pm}$ and a point $\boldsymbol{x} \in \mathbb{R}^d$ with $\widetilde{F}_{\boldsymbol{c},\pm}(\boldsymbol{x}) \neq F_{\boldsymbol{c},\pm}(\boldsymbol{x})$. Since $\widetilde{F}_{\boldsymbol{c},\pm}(\boldsymbol{y}) = F_{\boldsymbol{c},\pm}(\boldsymbol{y}) = 0$ as soon as $y_j = c_j$ for at least one $j \in \{1, \ldots, d\}$, this implies $\Delta_{\boldsymbol{c}}^{\boldsymbol{x}}\widetilde{F}_{\boldsymbol{c},\pm} \neq \Delta_{\boldsymbol{c}}^{\boldsymbol{x}}F_{\boldsymbol{c},\pm}$, or rather $(-1)^{d-|J^{\boldsymbol{c},\boldsymbol{x}}|}\,\Delta_{\underline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}}}^{\overline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}}}\widetilde{F}_{\boldsymbol{c},\pm} \neq (-1)^{d-|J^{\boldsymbol{c},\boldsymbol{x}}|}\,\Delta_{\underline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}}}^{\overline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}}}F_{\boldsymbol{c},\pm}$, where $\underline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}}$ and $\overline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}}$ refer to the smallest and the largest element of $\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{x}}$, respectively. Since (2.25) is satisfied for both $\widetilde{F}_{\boldsymbol{c},\pm}$ and $F_{\boldsymbol{c},\pm}$, this leads to a contradiction.

(iv): The right continuity of $F_{\boldsymbol{c},0}$ easily follows from the right continuity of F. It remains to show that $F_{\boldsymbol{c},\pm}$ is right continuous at every point $\boldsymbol{a} \in \mathbb{R}^d$. We only show right continuity in the first coordinate since the proof for the other coordinates follows by the same arguments.

Let $\boldsymbol{a} \in \mathbb{R}^d$ and $[\boldsymbol{u}, \boldsymbol{v}] := [\boldsymbol{u}_{\boldsymbol{a},\boldsymbol{c}}, \boldsymbol{v}_{\boldsymbol{a},\boldsymbol{c}}] \subsetneq \mathbb{R}^d$ with $\boldsymbol{a} \in [\boldsymbol{u}, \boldsymbol{v})$ and $\boldsymbol{c} \in [\boldsymbol{u}, \boldsymbol{v}]$. Then we define functions $G_{\boldsymbol{c},+} = G_{F,\boldsymbol{c},+}^{\boldsymbol{u},\boldsymbol{v},(1)} : [\boldsymbol{u}, \boldsymbol{v}] \to \mathbb{R}$ and $G_{\boldsymbol{c},-} = G_{F,\boldsymbol{c},-}^{\boldsymbol{u},\boldsymbol{v},(1)} : [\boldsymbol{u}, \boldsymbol{v}] \to \mathbb{R}$ by

$$G_{\boldsymbol{c},\pm}(\boldsymbol{x}) := \begin{cases} \lim_{n\to\infty} F_{\boldsymbol{c},\pm}(x_1 + \varepsilon_n, x_2, \ldots, x_d) & , \quad \boldsymbol{x} \in [\boldsymbol{u}, \boldsymbol{v}] \text{ and } x_1 \neq v_1 \\ F_{\boldsymbol{c},\pm}(x_1, x_2, \ldots, x_d) & , \quad \boldsymbol{x} \in [\boldsymbol{u}, \boldsymbol{v}] \text{ and } x_1 = v_1 \end{cases} \tag{2.28}$$

with $\varepsilon_n \searrow 0$ as $n \to \infty$. The limit in (2.28) exists since we prove in the following that there exists an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ the functions $F_{\boldsymbol{c},\pm}(x_1 + \varepsilon_n, x_2, \ldots, x_d)$ are bounded on $[\boldsymbol{u}, \boldsymbol{v}]$ and monotone in n for any fixed $\boldsymbol{x} \in [\boldsymbol{u}, \boldsymbol{v}]$. The boundedness of $F_{\boldsymbol{c},\pm}(x_1 + \varepsilon_n, x_2, \ldots, x_d)$ follows directly from the assumption that F is locally of bounded d-fold variation. To show the monotonicity, let n_0 be chosen so small that for all $n \geq n_0$ and fixed $\boldsymbol{x}$ either $x_1 + \varepsilon_n \leq c_1$ or $x_1 + \varepsilon_n > c_1$ (recall $\varepsilon_n \searrow 0$ as $n \to \infty$). If $x_1 + \varepsilon_n \leq c_1$, the positive variation $V_F^+(I_1^{\boldsymbol{c},\boldsymbol{x}+\varepsilon_n} \times I_2^{\boldsymbol{c},\boldsymbol{x}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{x}})$ and the negative variation $V_F^-(I_1^{\boldsymbol{c},\boldsymbol{x}+\varepsilon_n} \times I_2^{\boldsymbol{c},\boldsymbol{x}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{x}})$ increase in n as variations on the increasing interval $[x_1 + \varepsilon_n, c_1] \times I_2^{\boldsymbol{c},\boldsymbol{x}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{x}}$. If $c_1 < x_1 + \varepsilon_n$, they decrease as variations on the decreasing interval $[c_1, x_1 + \varepsilon_n] \times I_2^{\boldsymbol{c},\boldsymbol{x}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{x}}$. As a consequence $F_{\boldsymbol{c},\pm}(x_1 + \varepsilon_n, x_2, \ldots, x_d) = (-1)^{d-|J^{(c_1,c_2,\ldots,c_d),(x_1+\varepsilon_n,x_2,\ldots,x_d)}|}\,V_F^{\pm}(I_1^{\boldsymbol{c},\boldsymbol{x}+\varepsilon_n} \times I_2^{\boldsymbol{c},\boldsymbol{x}} \times \cdots \times I_d^{\boldsymbol{c},\boldsymbol{x}})$ is either monotonically increasing or monotonically decreasing in n depending on the prefactor $(-1)^{d-|J^{(c_1,c_2,\ldots,c_d),(x_1+\varepsilon_n,x_2,\ldots,x_d)}|}$.

By definition $G_{\boldsymbol{c},\pm}$ is right continuous in the first coordinate, at the point $\boldsymbol{a}$. For the right continuity of $F_{\boldsymbol{c},\pm}$ in the first coordinate, at the point $\boldsymbol{a}$, it suffices to prove that $G_{\boldsymbol{c},\pm}$ coincides with $F_{\boldsymbol{c},\pm}$ on the interval $[\boldsymbol{u}, \boldsymbol{v}]$. For this purpose, in view of (iii), it even suffices to show that $G_{\boldsymbol{c},+}$ and $G_{\boldsymbol{c},-}$ are d-fold monotonically increasing functions with $G_{\boldsymbol{c},\pm}(\boldsymbol{x}) = 0$ for all $\boldsymbol{x} \in [\boldsymbol{u}, \boldsymbol{v}]$ with $x_i = c_i$ for at least one $i \in \{1, \ldots, d\}$ that additionally satisfy

$$F(\boldsymbol{x}) \;=\; G_{\boldsymbol{c},+}(\boldsymbol{x}) - G_{\boldsymbol{c},-}(\boldsymbol{x}) - F_{\boldsymbol{c},0}(\boldsymbol{x}) \tag{2.29}$$

for any $\boldsymbol{x} \in [\boldsymbol{u}, \boldsymbol{v}]$ and

$$\Delta_{\boldsymbol{x}}^{\boldsymbol{y}} G_{\boldsymbol{c},\pm} = V_F^{\pm}([\boldsymbol{x}, \boldsymbol{y}]) \tag{2.30}$$

for all $\boldsymbol{x}, \boldsymbol{y} \in [\boldsymbol{u}, \boldsymbol{v}]$ with $\boldsymbol{x} < \boldsymbol{y}$.

We start with the proof of (2.29). If $\boldsymbol{x} \in [\boldsymbol{u}, \boldsymbol{v}]$ with $x_1 = v_1$, we already know by definition that $G_{\boldsymbol{c},+}(\boldsymbol{x}) - G_{\boldsymbol{c},-}(\boldsymbol{x}) = F_{\boldsymbol{c},+}(\boldsymbol{x}) - F_{\boldsymbol{c},-}(\boldsymbol{x})$. By means of decomposition (2.24), this implies (2.29). If $x_1 \neq v_1$, then we obtain by (2.28), (2.24) and the right continuity of F and $F_{\boldsymbol{c},0}$ that

$$\begin{aligned}
G_{\boldsymbol{c},+}(\boldsymbol{x}) - G_{\boldsymbol{c},-}(\boldsymbol{x}) &= \lim_{n\to\infty} \big(F_{\boldsymbol{c},+}(x_1+\varepsilon_n, x_2, \ldots, x_d) - F_{\boldsymbol{c},-}(x_1+\varepsilon_n, x_2, \ldots, x_d)\big) \\
&= \lim_{n\to\infty} \big(F(x_1+\varepsilon_n, x_2, \ldots, x_d) + F_{\boldsymbol{c},0}(x_1+\varepsilon_n, x_2, \ldots, x_d)\big) \\
&= F(x_1, x_2, \ldots, x_d) + F_{\boldsymbol{c},0}(x_1, x_2, \ldots, x_d),
\end{aligned}$$

which proves (2.29).

The d-fold monotonicity of $G_{\boldsymbol{c},\pm}$ is an immediate consequence of the definition of $G_{\boldsymbol{c},\pm}$ and the d-fold monotonicity of $F_{\boldsymbol{c},\pm}$.

For the proof of (2.30) it suffices to show

$$\Delta_{(z_1,x_2,\ldots,x_d)}^{(v_1,y_2,\ldots,y_d)} G_{\boldsymbol{c},\pm} = V_F^{\pm}([z_1, v_1] \times [x_2, y_2] \times \cdots \times [x_d, y_d]) \tag{2.31}$$

for every $\boldsymbol{x}, \boldsymbol{y} \in [\boldsymbol{u}, \boldsymbol{v}]$ with $\boldsymbol{x} < \boldsymbol{y}$ and $z_1 \in \{x_1, y_1\}$. Assertion (2.30) follows directly from (2.31) because

$$\begin{aligned}
\Delta_{\boldsymbol{x}}^{\boldsymbol{y}} G_{\boldsymbol{c},\pm} &= \Delta_{(x_1,x_2,\ldots,x_d)}^{(v_1,y_2,\ldots,y_d)} G_{\boldsymbol{c},\pm} - \Delta_{(y_1,x_2,\ldots,x_d)}^{(v_1,y_2,\ldots,y_d)} G_{\boldsymbol{c},\pm} \\
&= V_F^{\pm}([x_1, v_1] \times [x_2, y_2] \times \cdots \times [x_d, y_d]) - V_F^{\pm}([y_1, v_1] \times [x_2, y_2] \times \cdots \times [x_d, y_d]) \\
&= V_F^{\pm}([x_1, y_1] \times [x_2, y_2] \times \cdots \times [x_d, y_d])
\end{aligned}$$

by (2.14) and Lemma 2.4.6. For the proof of (2.31) we have on the one hand

$$\begin{aligned}
&V_F^{\pm}([z_1, v_1] \times [x_2, y_2] \times \cdots \times [x_d, y_d]) \\
&= \sup_{\{(x_{1,i_1},\ldots,x_{d,i_d})\} \in \mathcal{P}([z_1,v_1]\times[x_2,y_2]\times\cdots\times[x_d,y_d])} \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \big(\Delta_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}^{(x_{1,i_1},\ldots,x_{d,i_d})} F\big)^{\pm} \\
&= \sup_{\{(x_{1,i_1},\ldots,x_{d,i_d})\} \in \mathcal{P}([z_1,v_1]\times[x_2,y_2]\times\cdots\times[x_d,y_d])} \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \Big(\Delta_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}^{(x_{1,i_1},\ldots,x_{d,i_d})} G_{\boldsymbol{c},+} \\
&\qquad - \Delta_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}^{(x_{1,i_1},\ldots,x_{d,i_d})} G_{\boldsymbol{c},-}\Big)^{\pm} \\
&\leq \sup_{\{(x_{1,i_1},\ldots,x_{d,i_d})\} \in \mathcal{P}([z_1,v_1]\times[x_2,y_2]\times\cdots\times[x_d,y_d])} \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \Delta_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}^{(x_{1,i_1},\ldots,x_{d,i_d})} G_{\boldsymbol{c},\pm}
\end{aligned}$$

$$= \Delta_{(z_1,x_2,\ldots,x_d)}^{(v_1,y_2,\ldots,y_d)} G_{\boldsymbol{c},\pm}, \tag{2.32}$$

where we used in the second step that the representation (2.29) holds true and $F_{\boldsymbol{c},0}$ is d-fold constant, in the third step that $G_{\boldsymbol{c},+}$ and $G_{\boldsymbol{c},-}$ are d-fold monotonically increasing and in the last step that (2.14) is valid. On the other hand, the definition of $G_{\boldsymbol{c},\pm}$ and (2.25) yields

$$\begin{aligned}\Delta_{(z_1,x_2,\ldots,x_d)}^{(v_1,y_2,\ldots,y_d)} G_{\boldsymbol{c},\pm} &= \lim_{n\to\infty} \Delta_{(z_1+\varepsilon_n,x_2,\ldots,x_d)}^{(v_1,y_2,\ldots,y_d)} F_{\boldsymbol{c},\pm} \\ &= \lim_{n\to\infty} V_F^{\pm}([z_1+\varepsilon_n, v_1]\times[x_2,y_2]\times\cdots\times[x_d,y_d]) \\ &\leq V_F^{\pm}([z_1,v_1]\times[x_2,y_2]\times\cdots\times[x_d,y_d]),\end{aligned}$$

where the last step is valid because $V_F^{\pm}([z_1+\varepsilon_n, v_1]\times[x_2,y_2]\times\cdots\times[x_d,y_d])$ increases in n as variation on an increasing interval. Together with (2.32), this implies (2.31).

Finally, we have to show that $G_{\boldsymbol{c},\pm}(\boldsymbol{x}) = 0$ for all $\boldsymbol{x} \in [\boldsymbol{u},\boldsymbol{v}]$ with $x_i = c_i$ for at least one $i \in \{1,\ldots,d\}$. By definition (2.28) we immediately obtain

$$G_{\boldsymbol{c},\pm}(x_1,x_2,\ldots,x_d) = 0 \qquad \text{if } x_i = c_i \text{ for at least one } i \in \{2,\ldots,d\}. \tag{2.33}$$

To show $G_{\boldsymbol{c},\pm}(c_1,x_2,\ldots,x_d) = 0$ for any $\boldsymbol{x} \in [\boldsymbol{u},\boldsymbol{v}]$ with $x_i \neq c_i$ for $i \in \{2,\ldots,d\}$, we note that

$$\begin{aligned} & G_{\boldsymbol{c},\pm}(v_1,x_2,\ldots,x_d) - G_{\boldsymbol{c},\pm}(c_1,x_2,\ldots,x_d) \\ &= \Delta_{c_1}^{v_1}(G_{\boldsymbol{c},\pm})^{\boldsymbol{x};\{1\}} = \Delta_{(c_1,c_2,\ldots,c_d)}^{(v_1,x_2,\ldots,x_d)} G_{\boldsymbol{c},\pm} = (-1)^{d-1-|J^{(c_2,\ldots,c_d),(x_2,\ldots,x_d)}|}\, \Delta_{(c_1,\underline{\iota}_2^{\boldsymbol{c},\boldsymbol{x}},\ldots,\underline{\iota}_d^{\boldsymbol{c},\boldsymbol{x}})}^{(v_1,\overline{\iota}_2^{\boldsymbol{c},\boldsymbol{x}},\ldots,\overline{\iota}_d^{\boldsymbol{c},\boldsymbol{x}})} G_{\boldsymbol{c},\pm} \\ &= (-1)^{d-1-|J^{(c_2,\ldots,c_d),(x_2,\ldots,x_d)}|}\, V_F^{\pm}([c_1,v_1]\times I_2^{\boldsymbol{c},\boldsymbol{x}}\times\ldots\times I_d^{\boldsymbol{c},\boldsymbol{x}}) \\ &= (-1)^{d-1-|J^{(c_2,\ldots,c_d),(x_2,\ldots,x_d)}|}\, \Delta_{(c_1,\underline{\iota}_2^{\boldsymbol{c},\boldsymbol{x}},\ldots,\underline{\iota}_d^{\boldsymbol{c},\boldsymbol{x}})}^{(v_1,\overline{\iota}_2^{\boldsymbol{c},\boldsymbol{x}},\ldots,\overline{\iota}_d^{\boldsymbol{c},\boldsymbol{x}})} F_{\boldsymbol{c},\pm} = \Delta_{(c_1,c_2,\ldots,c_d)}^{(v_1,x_2,\ldots,x_d)} F_{\boldsymbol{c},\pm} = \Delta_{c_1}^{v_1}(F_{\boldsymbol{c},\pm})^{\boldsymbol{x};\{1\}} \\ &= F_{\boldsymbol{c},\pm}(v_1,x_2,\ldots,x_d) = G_{\boldsymbol{c},\pm}(v_1,x_2,\ldots,x_d),\end{aligned}$$

where $J^{(c_2,\ldots,c_d),(x_2,\ldots,x_d)}$ is defined analogously to $J^{\boldsymbol{c},\boldsymbol{x}}$, and $\underline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}} := (\underline{\iota}_1^{\boldsymbol{c},\boldsymbol{x}},\ldots,\underline{\iota}_d^{\boldsymbol{c},\boldsymbol{x}})$ and $\overline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}} := (\overline{\iota}_1^{\boldsymbol{c},\boldsymbol{x}},\ldots,\overline{\iota}_d^{\boldsymbol{c},\boldsymbol{x}})$ refer to the smallest and the largest element of $\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{x}} := I_1^{\boldsymbol{c},\boldsymbol{x}}\times\cdots\times I_d^{\boldsymbol{c},\boldsymbol{x}}$, respectively. Here, the second, forth and fifth step is valid by (2.33), (2.30) and (2.25), respectively. Moreover, we used that $F_{\boldsymbol{c},\pm}(y_1,y_2,\ldots,y_d) = 0$ if $y_i = c_i$ for at least one $i \in \{1,\ldots,d\}$ by (ii) for the third- and second-last step and the definition of $G_{\boldsymbol{c},\pm}$ for the last step. This implies that $G_{\boldsymbol{c},\pm}(\boldsymbol{x}) = 0$ for all $\boldsymbol{x} \in [\boldsymbol{u},\boldsymbol{v}]$ with $x_1 = c_1$.

Hence, the functions $G_{\boldsymbol{c},+}$ and $G_{\boldsymbol{c},-}$ indeed coincide with $F_{\boldsymbol{c},+}$ and $F_{\boldsymbol{c},-}$, respectively, on the interval $[\boldsymbol{u},\boldsymbol{v}]$, which completes the proof of (iv). □

In the Jordan decomposition (2.24) we can, of course, allocate the function $F_{\boldsymbol{c},0}$ to $F_{\boldsymbol{c},+}$ and $F_{\boldsymbol{c},-}$. The resulting functions are obviously still d-fold monotonically increasing. If we allocate $F_{\boldsymbol{c},0}$ in equal shares to $F_{\boldsymbol{c},+}$ and $F_{\boldsymbol{c},-}$, then we arrive at the generalization of the Jordan decomposition given in Proposition 1.17 in [39], or rather at a variant of Theorem 3 in [51] (also mentioned as Lemma 3 in [1]) for functions on $\mathbb{R}^d$ that are locally of bounded d-fold variation.

Corollary 2.4.9 *Let $F : \mathbb{R}^d \to \mathbb{R}$ be a function that is locally of bounded d-fold variation. For any $\mathbf{c} \in \mathbb{R}^d$, let the functions $F^0_{\mathbf{c},+}, F^0_{\mathbf{c},-} : \mathbb{R}^d \to \mathbb{R}$ be defined by*

$$F^0_{\mathbf{c},+}(\boldsymbol{x}) \;:=\; \frac{1}{2}\Big((-1)^{d-|J^{\mathbf{c},\boldsymbol{x}}|}\, V_F(\boldsymbol{I}^{\mathbf{c},\boldsymbol{x}}) + F(\boldsymbol{x})\Big), \qquad \boldsymbol{x} \in \mathbb{R}^d, \tag{2.34}$$

$$F^0_{\mathbf{c},-}(\boldsymbol{x}) \;:=\; \frac{1}{2}\Big((-1)^{d-|J^{\mathbf{c},\boldsymbol{x}}|}\, V_F(\boldsymbol{I}^{\mathbf{c},\boldsymbol{x}}) - F(\boldsymbol{x})\Big), \qquad \boldsymbol{x} \in \mathbb{R}^d \tag{2.35}$$

with $J^{\mathbf{c},\boldsymbol{x}} := \{j \in \{1, \ldots, d\} : c_j < x_j\}$ and $\boldsymbol{I}^{\mathbf{c},\boldsymbol{x}}$ as defined in (2.23). Then the following assertion hold:

(i) *The function F has the representation*

$$F \;=\; F^0_{\mathbf{c},+} - F^0_{\mathbf{c},-} \tag{2.36}$$

and

$$\Delta^{\boldsymbol{b}}_{\boldsymbol{a}} F^0_{\mathbf{c},+} = \Delta^{\boldsymbol{b}}_{\boldsymbol{a}} F_{\mathbf{c},+} \quad \text{and} \quad \Delta^{\boldsymbol{b}}_{\boldsymbol{a}} F^0_{\mathbf{c},-} = \Delta^{\boldsymbol{b}}_{\boldsymbol{a}} F_{\mathbf{c},-} \tag{2.37}$$

hold for all $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$, where $F_{\mathbf{c},+}$ and $F_{\mathbf{c},-}$ are defined by (2.20)–(2.21).

(ii) *For any $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$, we have*

$$\Delta^{\boldsymbol{b}}_{\boldsymbol{a}} F^0_{\mathbf{c},+} = V_F^+([\boldsymbol{a}, \boldsymbol{b}]) \quad \text{and} \quad \Delta^{\boldsymbol{b}}_{\boldsymbol{a}} F^0_{\mathbf{c},-} = V_F^-([\boldsymbol{a}, \boldsymbol{b}]).$$

In particular, the functions $F^0_{\mathbf{c},+}$ and $F^0_{\mathbf{c},-}$ are d-fold monotonically increasing.

(iii) *If F is right continuous, then the same is true for $F^0_{\mathbf{c},+}$ and $F^0_{\mathbf{c},-}$.*

Proof (i): By the definitions of $F_{c,+}$ and $F_{c,-}$ in (2.20)–(2.21), and parts (iii)–(iv) of Lemma 2.4.7, we obtain

$$F_{\mathbf{c},+}(\boldsymbol{x}) \;=\; (-1)^{d-|J^{\mathbf{c},\boldsymbol{x}}|}\, V_F^+(\boldsymbol{I}^{\mathbf{c},\boldsymbol{x}}) \;=\; (-1)^{d-|J^{\mathbf{c},\boldsymbol{x}}|}\, \frac{1}{2}\Big(V_F(\boldsymbol{I}^{\mathbf{c},\boldsymbol{x}}) + \Delta^{\overline{\boldsymbol{\iota}}^{\mathbf{c},\boldsymbol{x}}}_{\underline{\boldsymbol{\iota}}^{\mathbf{c},\boldsymbol{x}}} F\Big) \tag{2.38}$$

and

$$F_{\mathbf{c},-}(\boldsymbol{x}) \;=\; (-1)^{d-|J^{\mathbf{c},\boldsymbol{x}}|}\, V_F^-(\boldsymbol{I}^{\mathbf{c},\boldsymbol{x}}) \;=\; (-1)^{d-|J^{\mathbf{c},\boldsymbol{x}}|}\, \frac{1}{2}\Big(V_F(\boldsymbol{I}^{\mathbf{c},\boldsymbol{x}}) - \Delta^{\overline{\boldsymbol{\iota}}^{\mathbf{c},\boldsymbol{x}}}_{\underline{\boldsymbol{\iota}}^{\mathbf{c},\boldsymbol{x}}} F\Big), \tag{2.39}$$

where $\underline{\boldsymbol{\iota}}^{\mathbf{c},\boldsymbol{x}}$ and $\overline{\boldsymbol{\iota}}^{\mathbf{c},\boldsymbol{x}}$ refer to the smallest and the largest element of $\boldsymbol{I}^{\mathbf{c},\boldsymbol{x}}$, respectively. Moreover, as already noted in the proof of part (i) of Theorem 2.4.8, we have

$$(-1)^{d-|J^{\mathbf{c},\boldsymbol{x}}|}\, \Delta^{\overline{\boldsymbol{\iota}}^{\mathbf{c},\boldsymbol{x}}}_{\underline{\boldsymbol{\iota}}^{\mathbf{c},\boldsymbol{x}}} F \;=\; F(\boldsymbol{x}) + F_{\mathbf{c},0}(\boldsymbol{x}). \tag{2.40}$$

Combining (2.38)–(2.40) yields

$$\begin{aligned} F_{\mathbf{c},+}(\boldsymbol{x}) - \frac{1}{2}F_{\mathbf{c},0}(\boldsymbol{x}) \;&=\; (-1)^{d-|J^{\mathbf{c},\boldsymbol{x}}|}\, \frac{1}{2}\Big(V_F(\boldsymbol{I}^{\mathbf{c},\boldsymbol{x}}) + \Delta^{\overline{\boldsymbol{\iota}}^{\mathbf{c},\boldsymbol{x}}}_{\underline{\boldsymbol{\iota}}^{\mathbf{c},\boldsymbol{x}}} F\Big) - \frac{1}{2}F_{\mathbf{c},0}(\boldsymbol{x}) \\ &=\; (-1)^{d-|J^{\mathbf{c},\boldsymbol{x}}|}\, \frac{1}{2}V_F(\boldsymbol{I}^{\mathbf{c},\boldsymbol{x}}) + \frac{1}{2}\Big(F(\boldsymbol{x}) + F_{\mathbf{c},0}(\boldsymbol{x})\Big) - \frac{1}{2}F_{\mathbf{c},0}(\boldsymbol{x}) \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}\Big((-1)^{d-|J^{\boldsymbol{c},\boldsymbol{x}}|}\,V_F(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{x}}) + F(\boldsymbol{x})\Big) \\
&= F^0_{\boldsymbol{c},+}(\boldsymbol{x}) \qquad (2.41)
\end{aligned}
$$

and

$$
\begin{aligned}
F_{\boldsymbol{c},-}(\boldsymbol{x}) + \frac{1}{2}F_{\boldsymbol{c},0}(\boldsymbol{x}) &= (-1)^{d-|J^{\boldsymbol{c},\boldsymbol{x}}|}\,\frac{1}{2}\Big(V_F(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{x}}) - \Delta^{\overline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}}}_{\underline{\boldsymbol{\iota}}^{\boldsymbol{c},\boldsymbol{x}}}F\Big) + \frac{1}{2}F_{\boldsymbol{c},0}(\boldsymbol{x}) \\
&= (-1)^{d-|J^{\boldsymbol{c},\boldsymbol{x}}|}\,\frac{1}{2}V_F(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{x}}) - \frac{1}{2}\Big(F(\boldsymbol{x}) + F_{\boldsymbol{c},0}(\boldsymbol{x})\Big) + \frac{1}{2}F_{\boldsymbol{c},0}(\boldsymbol{x}) \\
&= \frac{1}{2}\Big((-1)^{d-|J^{\boldsymbol{c},\boldsymbol{x}}|}\,V_F(\boldsymbol{I}^{\boldsymbol{c},\boldsymbol{x}}) - F(\boldsymbol{x})\Big) \\
&= F^0_{\boldsymbol{c},-}(\boldsymbol{x}). \qquad (2.42)
\end{aligned}
$$

Now, (2.36) follows from (2.24) and (2.41)–(2.42). Moreover, (2.37) is an immediate consequence of (2.41)–(2.42) and the fact that $F_{\boldsymbol{c},0}$ is d-fold constant.

(ii): In view of (2.37), the assertion is an immediate consequence of part (ii) of Theorem 2.4.8.

(iii): The assertion follows from (2.41)–(2.42) and part (iii) of Theorem 2.4.8. □

Theorem 2.4.10 (Minimality) *Let $F : \mathbb{R}^d \to \mathbb{R}$ be a function that is locally of bounded d-fold variation, and $F^0_{\boldsymbol{c},+}$ and $F^0_{\boldsymbol{c},-}$ be defined by (2.34)–(2.35) for any $\boldsymbol{c} \in \mathbb{R}^d$. If $F_+, F_- : \mathbb{R}^d \to \mathbb{R}$ are two d-fold monotonically increasing functions such that $F = F_+ - F_-$, then $\Delta^{\boldsymbol{b}}_{\boldsymbol{a}}F_+ \geq \Delta^{\boldsymbol{b}}_{\boldsymbol{a}}F^0_{\boldsymbol{c},+}$ and $\Delta^{\boldsymbol{b}}_{\boldsymbol{a}}F_- \geq \Delta^{\boldsymbol{b}}_{\boldsymbol{a}}F^0_{\boldsymbol{c},-}$ for all $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$.*

Proof Since F_+ and F_- are d-fold monotonically increasing, we obtain

$$
\begin{aligned}
&V_F^+([\boldsymbol{a},\boldsymbol{b}]) \\
&= \sup_{\{(x_{1,i_1},\ldots,x_{d,i_d})\}\in\mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1}\cdots\sum_{i_d=1}^{n_d}\big(\Delta^{(x_{1,i_1},\ldots,x_{d,i_d})}_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}F\big)^+ \\
&= \sup_{\{(x_{1,i_1},\ldots,x_{d,i_d})\}\in\mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1}\cdots\sum_{i_d=1}^{n_d}\big(\Delta^{(x_{1,i_1},\ldots,x_{d,i_d})}_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}F_+ - \Delta^{(x_{1,i_1},\ldots,x_{d,i_d})}_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}F_-\big)^+ \\
&\leq \sup_{\{(x_{1,i_1},\ldots,x_{d,i_d})\}\in\mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1}\cdots\sum_{i_d=1}^{n_d}\Delta^{(x_{1,i_1},\ldots,x_{d,i_d})}_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}F_+ \\
&= \Delta^{\boldsymbol{b}}_{\boldsymbol{a}}F_+
\end{aligned}
$$

and

$$
\begin{aligned}
&V_F^-([\boldsymbol{a},\boldsymbol{b}]) \\
&= \sup_{\{(x_{1,i_1},\ldots,x_{d,i_d})\}\in\mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1}\cdots\sum_{i_d=1}^{n_d}\big(\Delta^{(x_{1,i_1},\ldots,x_{d,i_d})}_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}F_+ - \Delta^{(x_{1,i_1},\ldots,x_{d,i_d})}_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}F_-\big)^-
\end{aligned}
$$

$$\begin{aligned}
&\leq \sup_{\{(x_{1,i_1},\ldots,x_{d,i_d})\}\in\mathcal{P}([\boldsymbol{a},\boldsymbol{b}])} \sum_{i_1=1}^{n_1}\cdots\sum_{i_d=1}^{n_d} \Delta_{(x_{1,i_1-1},\ldots,x_{d,i_d-1})}^{(x_{1,i_1},\ldots,x_{d,i_d})} F_-\\
&= \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_-,
\end{aligned}$$

where the last "=" is justified by (2.14) in each case. Together with part (ii) of Corollary 2.4.9 this proves the claim. □

Corollary 2.4.11 *Let $F:\mathbb{R}^d\to\mathbb{R}$ be a function that is right continuous and locally of bounded d-fold variation. Then there exist unique positive measures $\mu_F^{0,+},\mu_F^{0,-}$ on $\mathcal{B}(\mathbb{R}^d)$ satisfying*

$$\mu_F^{0,+}((\boldsymbol{a},\boldsymbol{b}]) = \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_{\boldsymbol{c},+} = \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_{\boldsymbol{c},+}^0 \quad \textit{and} \quad \mu_F^{0,-}((\boldsymbol{a},\boldsymbol{b}]) = \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_{\boldsymbol{c},-} = \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_{\boldsymbol{c},-}^0$$

for all $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}\in\mathbb{R}^d$ with $\boldsymbol{a}<\boldsymbol{b}$, where $F_{\boldsymbol{c},+},F_{\boldsymbol{c},-}$ are defined by (2.20)–(2.21) and $F_{\boldsymbol{c},+}^0,F_{\boldsymbol{c},-}^0$ are defined by (2.34)–(2.35).

Proof The claim is an immediate consequence of Theorem 2.4.8, Corollary 2.4.9, Theorem 2.4.10 and Theorem 2.3.2. □

2.5 Measure generating functions and integrals with respect to signed measures

If in the setting of Corollary 2.4.11 at least one of the positive measures $\mu_F^{0,+},\mu_F^{0,-}$ is finite, then there exists a unique signed measure μ_F on $\mathcal{B}(\mathbb{R}^d)$ satisfying

$$\mu_F((\boldsymbol{a},\boldsymbol{b}]) = \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F$$

for all $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^d$ with $\boldsymbol{a}<\boldsymbol{b}$. This signed measure is given by

$$\mu_F := \mu_F^{0,+} - \mu_F^{0,-}, \tag{2.43}$$

and we will refer to it as *signed measure generated by F*. The positive measure $|\mu_F| := \mu_F^{0,+}+\mu_F^{0,-}$ will be referred to as *total variation measure* of μ_F. If both $\mu_F^{0,+}$ and $\mu_F^{0,-}$ are not finite but only finite on every compact interval $[\boldsymbol{a},\boldsymbol{b}]$ in $\mathbb{R}^d$, then μ_F is well defined at least on the ring $\mathcal{R}(\mathbb{R}^d)$ of all bounded sets from $\mathcal{B}(\mathbb{R}^d)$. In this case, we will refer to μ_F as *signed pre-measure generated by F*. Note that $\mu_F^{0,+},\mu_F^{0,-}$ are finite on compact intervals when F is bounded on compact intervals. Also note that the right-hand side of (2.43) is the Hahn–Jordan decomposition of the signed (pre-)measure μ_F. In particular, $\mu_F^{0,+}\perp\mu_F^{0,-}$ on compact intervals.

For a measurable function $G : \mathbb{R}^d \to \mathbb{R}$ the *integral of G with respect to the signed measure μ_F* is given by

$$\int_{\mathbb{R}^d} G(\boldsymbol{x})\, \mu_F(d\boldsymbol{x}) := \int_{\mathbb{R}^d} G(\boldsymbol{x})\, \mu_F^{0,+}(d\boldsymbol{x}) - \int_{\mathbb{R}^d} G(\boldsymbol{x})\, \mu_F^{0,-}(d\boldsymbol{x}). \tag{2.44}$$

We say that the integral on the left-hand side *exists*, if the integrals $\int_{\mathbb{R}^d} G^+(\boldsymbol{x})\, \mu_F^{0,+}(d\boldsymbol{x})$, $\int_{\mathbb{R}^d} G^-(\boldsymbol{x})\, \mu_F^{0,+}(d\boldsymbol{x})$, $\int_{\mathbb{R}^d} G^+(\boldsymbol{x})\, \mu_F^{0,-}(d\boldsymbol{x})$ and $\int_{\mathbb{R}^d} G^-(\boldsymbol{x})\, \mu_F^{0,-}(d\boldsymbol{x})$ are all finite, where G^+ and G^- denote the positive and the negative part of G.

We have seen that every right continuous function $F : \mathbb{R}^d \to \mathbb{R}$ that is locally of bounded d-fold variation can be written as difference of two right continuous d-fold monotonically increasing functions F_+ and F_-. In Corollary 2.4.9, for instance, we proved that $F = F^0_{\boldsymbol{c},+}(\boldsymbol{x}) - F^0_{\boldsymbol{c},-}(\boldsymbol{x})$, where $F^0_{\boldsymbol{c},+}(\boldsymbol{x})$ and $F^0_{\boldsymbol{c},-}(\boldsymbol{x})$ are defined as in (2.34) and (2.35), respectively. In the following, let $F = F_+ - F_-$ be any decomposition of F (not necessarily the Jordan decomposition from Corollary 2.4.9) into two right continuous and d-fold monotonically increasing functions F_+ and F_-. According to Theorem 2.3.2 there exist positive measures μ_F^+ and μ_F^- on $\mathcal{B}(\mathbb{R}^d)$ such that $\mu_F^\pm((\boldsymbol{a},\boldsymbol{b}]) = \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_\pm$ for all $\boldsymbol{a},\boldsymbol{b}$ with $\boldsymbol{a} < \boldsymbol{b}$. Let us stress that μ_F^+ and μ_F^- do *not necessarily* coincide with the unique positive measures $\mu_F^{0,+}$ and $\mu_F^{0,-}$ arising from the Jordan decomposition of F (cf. Corollary 2.4.11) as $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_+ \geq \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F^0_{\boldsymbol{c},+}$ and $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_- \geq \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F^0_{\boldsymbol{c},-}$ by Theorem 2.4.10. However, the signed (pre-)measure defined by

$$\mu_F := \mu_F^+ - \mu_F^-$$

on $\mathcal{B}(\mathbb{R}^d)$ satisfies

$$\mu_F((\boldsymbol{a},\boldsymbol{b}]) = \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_+ - \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F_- = \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} F \tag{2.45}$$

for all $\boldsymbol{a},\boldsymbol{b}$ with $\boldsymbol{a} < \boldsymbol{b}$ and, therefore, coincides with the unique signed measure μ_F^0 arising from the Jordan decomposition (recall (2.43)).

According to this, for some $(\mathcal{B}(\mathbb{R}^d), \mathcal{B}(\mathbb{R}))$-measurable function $G : \mathbb{R}^d \to \mathbb{R}$ the values of the integrals $\int_{\mathbb{R}^d} G(\boldsymbol{x})\, \mu_F^+(d\boldsymbol{x})$ and $\int_{\mathbb{R}^d} G(\boldsymbol{x})\, \mu_F^-(d\boldsymbol{x})$ are greater than or equal to the values of the two integrals on the right-hand side of (2.44), provided both integrals exist (meaning that $\int_{\mathbb{R}^d} G^+(\boldsymbol{x})\, \mu_F^+(d\boldsymbol{x}), \int_{\mathbb{R}^d} G^-(\boldsymbol{x})\, \mu_F^+(d\boldsymbol{x})$ and $\int_{\mathbb{R}^d} G^+(\boldsymbol{x})\, \mu_F^-(d\boldsymbol{x})$, $\int_{\mathbb{R}^d} G^-(\boldsymbol{x})\, \mu_F^-(d\boldsymbol{x})$, respectively, are finite, where G^+ and G^- denote the positive and negative part of G). Nevertheless, if we can ensure the existence of the integrals $\int_{\mathbb{R}^d} G(\boldsymbol{x})\, \mu_F^+(d\boldsymbol{x})$ and $\int_{\mathbb{R}^d} G(\boldsymbol{x})\, \mu_F^-(d\boldsymbol{x})$, the difference of both integrals corresponds to the integral of G with respect to the signed measure μ_F as defined in (2.44). This leads to the following corollary.

Corollary 2.5.1 *Let $F : \mathbb{R}^d \to \mathbb{R}$ be any right continuous function that is locally of bounded d-fold variation. Let $F = F_+ - F_-$ be any decomposition in right continuous* *and d-fold monotonically increasing functions and denote by μ_F^+ and μ_F^- the positive measures generated by F_+ and F_-, respectively.*

If the integrals $\int_{\mathbb{R}^d} G(\boldsymbol{x})\,\mu_F^+(d\boldsymbol{x})$ and $\int_{\mathbb{R}^d} G(\boldsymbol{x})\,\mu_F^-(d\boldsymbol{x})$ exist for some $(\mathcal{B}(\mathbb{R}^d),\mathcal{B}(\mathbb{R}))$-measurable function $G:\mathbb{R}^d\to\mathbb{R}$, then

$$\int_{\mathbb{R}^d} G(\boldsymbol{x})\,\mu_F(d\boldsymbol{x}) \;=\; \int_{\mathbb{R}^d} G(\boldsymbol{x})\,\mu_F^+(d\boldsymbol{x}) - \int_{\mathbb{R}^d} G(\boldsymbol{x})\,\mu_F^-(d\boldsymbol{x}).$$

2.6 Integration by parts

Let $F:\mathbb{R}^d\to\mathbb{R}$ be a function that is right continuous and locally of bounded d-fold variation. For any $\boldsymbol{c}\in\mathbb{R}^d$ and each nonempty subset $J\subseteq\{1,\dots,d\}$, let $F^{\boldsymbol{c};J}:\mathbb{R}^J\to\mathbb{R}$ be defined as in (2.4). These functions are not necessarily locally of bounded $|J|$-fold variation as we have seen in Remark 2.4.5. On the other hand, in Theorem 2.6.4 below we will need that the $F^{\boldsymbol{c};J}$ are locally of bounded $|J|$-fold variation. For this reason we will assume that the $F^{\boldsymbol{c};J}$ possess this property.

We note that right continuity of F clearly implies right continuity of $F^{\boldsymbol{c};J}$ for every nonempty subset $J\subseteq\{1,\dots,d\}$ and for every $\boldsymbol{c}\in\mathbb{R}^d$. So the $|J|$-dimensional analogues of Corollary 2.4.9 to Corollary 2.4.11 ensure the existence of the decomposition of $F^{\boldsymbol{c};J}=(F^{\boldsymbol{c};J})_{\widetilde{\boldsymbol{c}},+}-(F^{\boldsymbol{c};J})_{\widetilde{\boldsymbol{c}},-}$ into $|J|$-fold monotonically increasing functions $(F^{\boldsymbol{c};J})_{\widetilde{\boldsymbol{c}},+}$ and $(F^{\boldsymbol{c};J})_{\widetilde{\boldsymbol{c}},-}$ for some $\widetilde{\boldsymbol{c}}\in\mathbb{R}^{|J|}$. In the following we will not insist on this Jordan decomposition of $F^{\boldsymbol{c};J}$. Instead, we allow any decomposition of $F^{\boldsymbol{c};J}=F_+^{\boldsymbol{c};J}-F_-^{\boldsymbol{c};J}$ into two right continuous and $|J|$-fold monotonically increasing functions $F_+^{\boldsymbol{c};J}$ and $F_-^{\boldsymbol{c};J}$. So there exist (not necessarily unique) positive measures $\mu^+_{F^{\boldsymbol{c};J}}$ and $\mu^-_{F^{\boldsymbol{c};J}}$ on $\mathcal{B}(\mathbb{R}^J)$ satisfying $\mu^\pm_{F^{\boldsymbol{c};J}}((\boldsymbol{a},\boldsymbol{b}])=\Delta_{\boldsymbol{a}}^{\boldsymbol{b}}F_\pm^{\boldsymbol{c};J}$ for all $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^J$ with $\boldsymbol{a}<\boldsymbol{b}$. Moreover, there exist unique signed (pre-)measures

$$\mu_{F^{\boldsymbol{c};J}} \;=:\; \mu_F^{\boldsymbol{c};J} \tag{2.46}$$

on $\mathcal{B}(\mathbb{R}^J)$ with $\mu_F^{\boldsymbol{c};J}=\mu^+_{F^{\boldsymbol{c};J}}-\mu^-_{F^{\boldsymbol{c};J}}$ such that $\mu_F^{\boldsymbol{c};J}((\boldsymbol{a},\boldsymbol{b}])=\Delta_{\boldsymbol{a}}^{\boldsymbol{b}}F^{\boldsymbol{c};J}$ for all $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^J$ with $\boldsymbol{a}<\boldsymbol{b}$. In the following, we set $\boldsymbol{x}_J:=(x_j)_{j\in J}$ for any $\boldsymbol{x}=(x_1,\dots,x_d)\in\mathbb{R}^d$ and any nonempty subset $J\subseteq\{1,\dots,d\}$. Notice that $\boldsymbol{x}_J$ is an element of $\mathbb{R}^J$.

For the proof of the integration by parts formula in Theorem 2.6.4, we need the following two lemmas. The first lemma is an immediate consequence of Lemma 2.2.4 in view of (2.45) and (2.46).

Lemma 2.6.1 *Let $F:\mathbb{R}^d\to\mathbb{R}$ be any function that is right continuous and locally of bounded d-fold variation. Moreover assume that the function $F^{\boldsymbol{a};J}$ defined in (2.4) is locally of bounded $|J|$-fold variation for every nonempty subset $J\subsetneq\{1,\dots,d\}$. Then, for any $\boldsymbol{a},\boldsymbol{x}\in\mathbb{R}^d$ with $\boldsymbol{a}<\boldsymbol{x}$, we have*

$$F(\boldsymbol{x}) \;=\; F(\boldsymbol{a}) + \sum_{\emptyset\neq J\subseteq\{1,\dots,d\}} \mu_F^{\boldsymbol{a};J}((\boldsymbol{a}_J,\boldsymbol{x}_J]),$$

which can be rewritten as

$$F(\boldsymbol{x}) \;=\; F(\boldsymbol{a}) \;+\; \sum_{\emptyset\neq J\subseteq\{1,\dots,d\}} \int_{(\boldsymbol{a}_J,\boldsymbol{x}_J]} d\mu_F^{\boldsymbol{a};J}. \tag{2.47}$$

The statement of the following lemma can also be found as Proposition A.1 in [34] and as formula (42) in the proof of Theorem 15 in [63].

Lemma 2.6.2 *Let $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^d$ with $\boldsymbol{a}<\boldsymbol{b}$. Let $F:\mathbb{R}^d\to\mathbb{R}$ and $G:\mathbb{R}^d\to\mathbb{R}$ be right continuous functions that are locally of bounded d-fold variation. Further, let the function $F^{\boldsymbol{a};J}$ be locally of bounded $|J|$-fold variation for every nonempty subset $J\subseteq\{1,\dots,d\}$. Assume that the maps $(\boldsymbol{a}_J,\boldsymbol{b}_J]\ni\boldsymbol{y}\mapsto\mu^{\pm}_{F^{\boldsymbol{a};J}}((\boldsymbol{a}_J,\boldsymbol{y}])$ and $(\boldsymbol{a}_J,\boldsymbol{b}_J]\ni\boldsymbol{y}\mapsto\mu^{\pm}_{G}(I^{\boldsymbol{a},\boldsymbol{b};J}_{\boldsymbol{y}})$ are $(\mathcal{B}((\boldsymbol{a}_J,\boldsymbol{b}_J]),\mathcal{B}(\mathbb{R}))$-measurable and that the integrals $\int_{(\boldsymbol{a},\boldsymbol{b}]}\mu^{\pm}_{F^{\boldsymbol{a};J}}((\boldsymbol{a}_J,\boldsymbol{x}_J])\,\mu^{\pm}_G(d\boldsymbol{x})$ and $\int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]}\mu^{\pm}_G(I^{\boldsymbol{a},\boldsymbol{b};J}_{\boldsymbol{y}})\,\mu^{\pm}_{F^{\boldsymbol{a};J}}(d\boldsymbol{y})$ exist for every nonempty subset $J\subseteq\{1,\dots,d\}$, where $\boldsymbol{y}$ is $|J|$-dimensional (that is, $\boldsymbol{y}=(y_j)_{j\in J}$),*

$$I^{\boldsymbol{a},\boldsymbol{b};J}_{\boldsymbol{y}} \;:=\; I^{\boldsymbol{a},\boldsymbol{b};J}_{1;\boldsymbol{y}}\times\cdots\times I^{\boldsymbol{a},\boldsymbol{b};J}_{d;\boldsymbol{y}} \quad \textit{with} \quad I^{\boldsymbol{a},\boldsymbol{b};J}_{j;\boldsymbol{y}} \;:=\; \begin{cases} [y_j,b_j] & ,\ j\in J\\ (a_j,b_j] & ,\ j\notin J\end{cases}$$

and $F^{\boldsymbol{a};J}_{\pm}$ and $G_{\pm}$ are $|J|$-fold and d-fold monotonically increasing right continuous functions, respectively, satisfying $F^{\boldsymbol{a};J}=F^{\boldsymbol{a};J}_{+}-F^{\boldsymbol{a};J}_{-}$ and $G=G_+-G_-$. Then

$$\int_{(\boldsymbol{a},\boldsymbol{b}]} F(\boldsymbol{x})\,\mu_G(d\boldsymbol{x}) \;=\; F(\boldsymbol{a})\mu_G((\boldsymbol{a},\boldsymbol{b}]) + \sum_{\emptyset\neq J\subseteq\{1,\dots,d\}} \int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} \mu_G(I^{\boldsymbol{a},\boldsymbol{b};J}_{\boldsymbol{y}})\,\mu^{\boldsymbol{a};J}_F(d\boldsymbol{y}).$$

Proof We use the representation (2.47) of F to obtain

$$\begin{aligned} &\int_{(\boldsymbol{a},\boldsymbol{b}]} F(\boldsymbol{x})\,\mu_G(d\boldsymbol{x})\\ &\quad=\;\int_{(\boldsymbol{a},\boldsymbol{b}]}\Big(F(\boldsymbol{a})+\sum_{\emptyset\neq J\subseteq\{1,\dots,d\}}\int_{(\boldsymbol{a}_J,\boldsymbol{x}_J]} d\mu^{\boldsymbol{a};J}_F\Big)\,\mu_G(d\boldsymbol{x})\\ &\quad=\;F(\boldsymbol{a})\mu_G((\boldsymbol{a},\boldsymbol{b}]) \;+\; \sum_{\emptyset\neq J\subseteq\{1,\dots,d\}} \int_{(\boldsymbol{a},\boldsymbol{b}]}\int_{(\boldsymbol{a}_J,\boldsymbol{x}_J]} \mu^{\boldsymbol{a};J}_F(d\boldsymbol{y})\,\mu_G(d\boldsymbol{x}). \end{aligned} \tag{2.48}$$

By Corollary 2.5.1 we have for every nonempty subset $J\subseteq\{1,\dots,d\}$

$$\begin{aligned} &\int_{(\boldsymbol{a},\boldsymbol{b}]}\int_{(\boldsymbol{a}_J,\boldsymbol{x}_J]} \mu^{\boldsymbol{a};J}_F(d\boldsymbol{y})\,\mu_G(d\boldsymbol{x})\\ &=\;\int_{(\boldsymbol{a},\boldsymbol{b}]}\int_{(\boldsymbol{a}_J,\boldsymbol{x}_J]} \mu^{+}_{F^{\boldsymbol{a};J}}(d\boldsymbol{y})\,\mu^{+}_G(d\boldsymbol{x}) - \int_{(\boldsymbol{a},\boldsymbol{b}]}\int_{(\boldsymbol{a}_J,\boldsymbol{x}_J]} \mu^{+}_{F^{\boldsymbol{a};J}}(d\boldsymbol{y})\,\mu^{-}_G(d\boldsymbol{x})\\ &\qquad - \int_{(\boldsymbol{a},\boldsymbol{b}]}\int_{(\boldsymbol{a}_J,\boldsymbol{x}_J]} \mu^{-}_{F^{\boldsymbol{a};J}}(d\boldsymbol{y})\,\mu^{+}_G(d\boldsymbol{x}) + \int_{(\boldsymbol{a},\boldsymbol{b}]}\int_{(\boldsymbol{a}_J,\boldsymbol{x}_J]} \mu^{-}_{F^{\boldsymbol{a};J}}(d\boldsymbol{y})\,\mu^{-}_G(d\boldsymbol{x}), \end{aligned} \tag{2.49}$$

where $\mu^+_{F^{\boldsymbol{a};J}}$, $\mu^-_{F^{\boldsymbol{a};J}}$, μ^+_G and μ^-_G are positive measures that are finite on compact intervals. Applying Fubini's theorem for every nonempty subset $J \subseteq \{1, \ldots, d\}$ yields

$$\begin{aligned}
&\int_{(\boldsymbol{a},\boldsymbol{b}]} \int_{(\boldsymbol{a}_J,\boldsymbol{x}_J]} \mu^{\pm}_{F^{\boldsymbol{a};J}}(d\boldsymbol{y})\, \mu^{\pm}_G(d\boldsymbol{x}) \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^J} \mathbb{1}_{(\boldsymbol{a}_J,\boldsymbol{x}_J]}(\boldsymbol{y}) \mathbb{1}_{(\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{x})\, \mu^{\pm}_{F^{\boldsymbol{a};J}}(d\boldsymbol{y}) \mu^{\pm}_G(d\boldsymbol{x}) \\
&= \int_{\mathbb{R}^J} \int_{\mathbb{R}^d} \mathbb{1}_{(\boldsymbol{a}_J,\boldsymbol{b}_J]}(\boldsymbol{y}) \mathbb{1}_{[\boldsymbol{y},\boldsymbol{b}_J]}(\boldsymbol{x}_J) \mathbb{1}_{(\boldsymbol{a}_{J^c},\boldsymbol{b}_{J^c}]}(\boldsymbol{x}_{J^c})\, \mu^{\pm}_G(d\boldsymbol{x})\, \mu^{\pm}_{F^{\boldsymbol{a};J}}(d\boldsymbol{y}) \\
&= \int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} \mu^{\pm}_G(I^{\boldsymbol{a},\boldsymbol{b};J}_{\boldsymbol{y}})\, \mu^{\pm}_{F^{\boldsymbol{a};J}}(d\boldsymbol{y})
\end{aligned}$$

with $J^c := \{1, \ldots, d\} \setminus J$. Along with (2.48) and (2.49) this finishes the proof. □

For any $\boldsymbol{x} \in \mathbb{R}^d$ and any right continuous function $F : \mathbb{R}^d \to \mathbb{R}$ with $F^{\boldsymbol{c};J}$ being locally of bounded $|J|$-fold variation for each nonempty subset $J \subseteq \{1, \ldots, d\}$ and every $\boldsymbol{c} \in \mathbb{R}^d$, we define

$$F(\boldsymbol{x}-) := \lim_{\boldsymbol{y} \nearrow \boldsymbol{x}} F(\boldsymbol{y}). \tag{2.50}$$

The existence of the left-hand limit in (2.50) is ensured by the following remark.

Remark 2.6.3 If $F : \mathbb{R}^d \to \mathbb{R}$ is a right continuous function that is locally of bounded d-fold variation and if additionally the functions $F^{\boldsymbol{c};J}$ are locally of bounded $|J|$-fold variation for each nonempty subset $J \subseteq \{1, \ldots, d\}$ and every $\boldsymbol{c} \in \mathbb{R}^d$, then the left-hand limit of F exists at every point $\boldsymbol{x} \in \mathbb{R}^d$. Indeed, by Lemma 2.2.4 we obtain for any $\boldsymbol{x} \in \mathbb{R}^d$ that $\lim_{\boldsymbol{y} \nearrow \boldsymbol{x}} F(\boldsymbol{y}) = F(\boldsymbol{c}) + \lim_{\boldsymbol{y} \nearrow \boldsymbol{x}} \sum_{\emptyset \neq J \subseteq \{1,\ldots,d\}} \Delta^{\boldsymbol{y}_J}_{\boldsymbol{c}_J} F^{\boldsymbol{c};J} = F(\boldsymbol{c}) + \lim_{\boldsymbol{y} \nearrow \boldsymbol{x}} \sum_{\emptyset \neq J \subseteq \{1,\ldots,d\}} \mu^{\boldsymbol{c};J}_F((\boldsymbol{c}_J, \boldsymbol{y}_J])$ for some $\boldsymbol{c} \in \mathbb{R}^d$ with $\boldsymbol{c} < \boldsymbol{x}$. The existence of the left-hand limits thus follows from the continuity from below of the (pre-)measures μ_F and $\mu^{\boldsymbol{c};J}_F$. ◇

The integration by parts formula (2.51) in Theorem 2.6.4 below is already known from Theorem 15 in [63] where Radulović et al. impose assumptions on the involved functions that differ from ours. We briefly discuss these differences subsequent to Corollary 2.6.5 dealing with the extension of the integration by parts formula to integrals over $\mathbb{R}^d$.

Theorem 2.6.4 (integration by parts formula) *Let $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} < \boldsymbol{b}$. Let $F : \mathbb{R}^d \to \mathbb{R}$ and $G : \mathbb{R}^d \to \mathbb{R}$ be right continuous functions and assume that the functions $F^{\boldsymbol{c};J}$ and $G^{\boldsymbol{c};J}$ are locally of bounded $|J|$-fold variation for every nonempty subset $J \subseteq \{1, \ldots, d\}$ and for every $\boldsymbol{c} = (c_1, \ldots, c_d) \in \mathbb{R}^d$ with $c_i \in \{a_i, b_i\}$ for $i = 1, \ldots, d$. Further assume that for every such J and $\boldsymbol{c}$ the maps $(\boldsymbol{a}_J, \boldsymbol{b}_J] \ni \boldsymbol{x}_J \mapsto F^{\boldsymbol{a};J}_{\pm}(\boldsymbol{x}_J)$ and $(\boldsymbol{a}_J, \boldsymbol{b}_J] \ni \boldsymbol{x}_J \mapsto G^{\boldsymbol{c};J}_{\pm}(\boldsymbol{x}_J-)$ are $(\mathcal{B}((\boldsymbol{a}_J, \boldsymbol{b}_J]), \mathcal{B}(\mathbb{R}))$-measurable and the integrals*

$\int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} F_\pm^{\boldsymbol{a};J}(\boldsymbol{x}_J)\,\mu^\pm_{G^{\boldsymbol{c};J}}(d\boldsymbol{x}_J)$ *and* $\int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} G_\pm^{\boldsymbol{c};J}(\boldsymbol{x}_J-)\,\mu^\pm_{F^{\boldsymbol{c};J}}(d\boldsymbol{x}_J)$ *exist, where* $F_\pm^{\boldsymbol{c};J}$ *and* $G_\pm^{\boldsymbol{c};J}$ *are* $|J|$*-fold monotonically increasing and right continuous functions satisfying* $F^{\boldsymbol{c};J} = F_+^{\boldsymbol{c};J} - F_-^{\boldsymbol{c};J}$ *and* $G^{\boldsymbol{c};J} = G_+^{\boldsymbol{c};J} - G_-^{\boldsymbol{c};J}$*. Then*

$$\begin{aligned}\int_{(\boldsymbol{a},\boldsymbol{b}]} F(\boldsymbol{x})\,\mu_G(d\boldsymbol{x}) &= \sum_{K\subseteq\{1,\ldots,d\}} (-1)^{d-|K|}\,G(\boldsymbol{b}^{\boldsymbol{a};K})F(\boldsymbol{b}^{\boldsymbol{a};K}) \\ &\quad + \sum_{J,K\subseteq\{1,\ldots,d\}\text{ disjoint},J\neq\emptyset} (-1)^{d-|K|}\int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} G(\boldsymbol{y}_K^{\boldsymbol{a},\boldsymbol{b};J}-)\,\mu_F^{(\boldsymbol{b}^{\boldsymbol{a};K});J}(d\boldsymbol{y}),\end{aligned} \tag{2.51}$$

where $\boldsymbol{b}^{\boldsymbol{a};K}$ *is defined as in (2.2) and the integration variable* $\boldsymbol{y}$ *is* $|J|$*-dimensional in the summand corresponding to* J *with* $G(\boldsymbol{y}_K^{\boldsymbol{a},\boldsymbol{b};J}-) := \lim_{\boldsymbol{u}\nearrow\boldsymbol{y}} G(\boldsymbol{u}_K^{\boldsymbol{a},\boldsymbol{b};J})$ *for*

$$\boldsymbol{u}_K^{\boldsymbol{a},\boldsymbol{b};J} := (u_{1;K}^{\boldsymbol{a},\boldsymbol{b};J},\ldots,u_{d;K}^{\boldsymbol{a},\boldsymbol{b};J}) \quad \textit{with} \quad u_{j;K}^{\boldsymbol{a},\boldsymbol{b};J} := \begin{cases} u_j &, \; j\in J \\ b_j &, \; j\in K \\ a_j &, \; j\notin J\cup K \end{cases}.$$

We note that for every nonempty subset $J\subseteq\{1,\ldots,d\}$ and every $\boldsymbol{c}=(c_1,\ldots,c_d)\in\mathbb{R}^d$ with $c_i\in\{a_i,b_i\}$ for $i=1,\ldots,d$, it is sufficient to replace the functions $F_\pm^{\boldsymbol{c};J}$ and $G_\pm^{\boldsymbol{c};J}$ in Theorem 2.6.4 by the Jordan functions $(F^{\boldsymbol{c};J})^0_{\widetilde{\mathbf{c}},\pm}$ and $(G^{\boldsymbol{c};J})^0_{\widetilde{\mathbf{c}},\pm}$ defined in (2.34) and (2.35), respectively, for some $\widetilde{c}\in\mathbb{R}^{|J|}$. However, in applications it might be difficult to verify that $(F^{\boldsymbol{c};J})^0_{\widetilde{\mathbf{c}},\pm}$ and $(G^{\boldsymbol{c};J})^0_{\widetilde{\mathbf{c}},\pm}$ are measurable. That's why we allow any other decomposition of $F_\pm^{\boldsymbol{c};J}$ and $G_\pm^{\boldsymbol{c};J}$ in $|J|$-fold monotonically increasing and right continuous functions at the expense of slightly stronger conditions on the existence of the integrals, see Corollary 2.5.1 and the discussion beforehand.

Proof of Theorem 2.6.4 To prove the assertion, we proceed as in the proof of Theorem 15 in [63]. In view of Lemma 2.6.2 it suffices to show

$$\begin{aligned}&\sum_{J\subseteq\{1,\ldots,d\}}\int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} \mu_G(I_{\boldsymbol{y}}^{\boldsymbol{a},\boldsymbol{b};J})\,\mu_F^{\boldsymbol{a};J}(d\boldsymbol{y}) \\ &\quad= \sum_{J,K\subseteq\{1,\ldots,d\}\text{ disjoint}} (-1)^{d-|K|}\int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} G(\boldsymbol{y}_K^{\boldsymbol{a},\boldsymbol{b};J}-)\,\mu_F^{(\boldsymbol{b}^{\boldsymbol{a};K});J}(d\boldsymbol{y}),\end{aligned} \tag{2.52}$$

where we use the notation that for $J=\emptyset$ the integral on the left-hand side corresponds to $F(\boldsymbol{a})\mu_G((\boldsymbol{a},\boldsymbol{b}])$ and that for $J=\emptyset$ the sum on the right-hand side is given by $\sum_{K\subseteq\{1,\ldots,d\}}(-1)^{d-|K|}\,G(\boldsymbol{b}^{\boldsymbol{a};K})F(\boldsymbol{b}^{\boldsymbol{a};K})$.

For the proof of (2.52) we obtain by the continuity from below of the signed (pre-) measure μ_G that

$$\sum_{J\subseteq\{1,\ldots,d\}}\int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} \mu_G(I_{\boldsymbol{y}}^{\boldsymbol{a},\boldsymbol{b};J})\,\mu_F^{\boldsymbol{a};J}(d\boldsymbol{y})$$

$$= \sum_{J\subseteq\{1,\dots,d\}} \int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} \lim_{\boldsymbol{u}\nearrow\boldsymbol{y}} \mu_G\big(I_1\times\cdots\times I_d\big)\, \mu_F^{\boldsymbol{a};J}(d\boldsymbol{y})$$

holds, where $I_j := (u_j, b_j]$ if $j \in J$, and $I_j := (a_j, b_j]$ if $j \in J^{\mathsf{c}}$. Due to (2.45) and (2.1) (where J in (2.1) corresponds to $(J\setminus T_1)\cup T_2$) we observe

$$\begin{aligned}
&\sum_{J\subseteq\{1,\dots,d\}} \int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} \mu_G(I_{\boldsymbol{y}}^{\boldsymbol{a},\boldsymbol{b};J})\, \mu_F^{\boldsymbol{a};J}(d\boldsymbol{y}) \\
&\quad= \sum_{J\subseteq\{1,\dots,d\}} \int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} \lim_{\boldsymbol{u}_J\nearrow\boldsymbol{y}} \Delta_{\boldsymbol{u}_J\cup\boldsymbol{a}_{J^{\mathsf{c}}}}^{\boldsymbol{b}_J\cup\boldsymbol{b}_{J^{\mathsf{c}}}} G\, \mu_F^{\boldsymbol{a};J}(d\boldsymbol{y}) \\
&\quad= \sum_{J\subseteq\{1,\dots,d\}} \\
&\qquad \int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} \lim_{\boldsymbol{u}_J\nearrow\boldsymbol{y}} \sum_{T_1\subseteq J}\sum_{T_2\subseteq J^{\mathsf{c}}} (-1)^{d-|J\setminus T_1|-|T_2|}\, G\big(\boldsymbol{u}_{T_1}\cup\boldsymbol{b}_{J\setminus T_1}\cup\boldsymbol{a}_{J^{\mathsf{c}}\setminus T_2}\cup\boldsymbol{b}_{T_2}\big)\, \mu_F^{\boldsymbol{a};J}(d\boldsymbol{y}) \\
&\quad= \sum_{J\subseteq\{1,\dots,d\}} \qquad\qquad (2.53)\\
&\qquad \sum_{T_1\subseteq J}\sum_{T_2\subseteq J^{\mathsf{c}}} (-1)^{d-|J\setminus T_1|-|T_2|} \int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} G\big((\boldsymbol{y}_{T_1}-)\cup\boldsymbol{b}_{J\setminus T_1}\cup\boldsymbol{a}_{J^{\mathsf{c}}\setminus T_2}\cup\boldsymbol{b}_{T_2}\big)\, \mu_F^{\boldsymbol{a};J}(d\boldsymbol{y})
\end{aligned}$$

with $G\big((\boldsymbol{y}_{T_1}-)\cup\boldsymbol{b}_{J\setminus T_1}\cup\boldsymbol{a}_{J^{\mathsf{c}}\setminus T_2}\cup\boldsymbol{b}_{T_2}\big) := \lim_{\boldsymbol{u}_{T_1}\nearrow\boldsymbol{y}} G\big(\boldsymbol{u}_{T_1}\cup\boldsymbol{b}_{J\setminus T_1}\cup\boldsymbol{a}_{J^{\mathsf{c}}\setminus T_2}\cup\boldsymbol{b}_{T_2}\big)$, where

$$\boldsymbol{x}_J\cup\boldsymbol{y}_K\cup\boldsymbol{z}_{(J\cup K)^{\mathsf{c}}} := (\alpha_1,\dots,\alpha_d) \quad \text{with} \quad \alpha_i := \begin{cases} x_j &, \; j\in J \\ y_j &, \; j\in K \\ z_j &, \; j\in(J\cup K)^{\mathsf{c}} \end{cases}$$

for any subsets $J,K\subseteq\{1,\dots,d\}$ and $\boldsymbol{x}_J\in\mathbb{R}^{|J|}$, $\boldsymbol{y}_K\in\mathbb{R}^{|K|}$ and $\boldsymbol{z}_{(J\cup K)^{\mathsf{c}}}\in\mathbb{R}^{|(J\cup K)^{\mathsf{c}}|}$.

In the next step we use that the integrand of the latter integral in (2.53) does not depend on y_j for $j\in J\setminus T_1$. Because of the special shape of the measure (as indicated in Remark 2.2.2) an evaluation of the integral on $(\boldsymbol{a}_{J\setminus T_1},\boldsymbol{b}_{J\setminus T_1}]$ results in

$$\begin{aligned}
&\sum_{J\subseteq\{1,\dots,d\}} \int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} \mu_G(I_{\boldsymbol{y}}^{\boldsymbol{a},\boldsymbol{b};J})\, \mu_F^{\boldsymbol{a};J}(d\boldsymbol{y}) \\
&\quad= \sum_{J\subseteq\{1,\dots,d\}}\sum_{T_1\subseteq J}\sum_{T_2\subseteq J^{\mathsf{c}}} (-1)^{d-|J\setminus T_1|-|T_2|} \sum_{K\subseteq J\setminus T_1} (-1)^{|J\setminus T_1|-|K|} \\
&\qquad\qquad \int_{(\boldsymbol{a}_{T_1},\boldsymbol{b}_{T_1}]} G\big((\boldsymbol{y}_{T_1}-)\cup\boldsymbol{b}_{J\setminus T_1}\cup\boldsymbol{a}_{J^{\mathsf{c}}\setminus T_2}\cup\boldsymbol{b}_{T_2}\big)\, \mu_F^{(\boldsymbol{b}^{\boldsymbol{a};K});T_1}(d\boldsymbol{y}_{T_1}) \\
&\quad= \sum_{J\subseteq\{1,\dots,d\}}\sum_{T_1\subseteq J}\sum_{K\subseteq J\setminus T_1} (-1)^{d-|K|} \sum_{T_2\subseteq J^{\mathsf{c}}} (-1)^{|T_2|} \qquad\qquad (2.54)\\
&\qquad\qquad \int_{(\boldsymbol{a}_{T_1},\boldsymbol{b}_{T_1}]} G\big((\boldsymbol{y}_{T_1}-)\cup\boldsymbol{a}_{J^{\mathsf{c}}\setminus T_2}\cup\boldsymbol{b}_{J\setminus(T_1\cup K)}\cup\boldsymbol{b}_K\cup\boldsymbol{b}_{T_2}\big)\, \mu_F^{(\boldsymbol{b}^{\boldsymbol{a};K});T_1}(d\boldsymbol{y}_{T_1}),
\end{aligned}$$

where $\mu_F^{(\boldsymbol{b}^{\boldsymbol{a};K});T_1}$ is the measure generated by the function $F^{(\boldsymbol{b}^{\boldsymbol{a};K});T_1}$ with

$$F^{(\boldsymbol{b}^{\boldsymbol{a};K});T_1}(\boldsymbol{x}_{T_1}) := F(\boldsymbol{b}_K \cup \boldsymbol{a}_{J\setminus(T_1\cup K)} \cup \boldsymbol{a}_{J^c\setminus T_2} \cup \boldsymbol{a}_{T_2} \cup \boldsymbol{x}_{T_1}), \qquad \boldsymbol{x}_{T_1} \in \mathbb{R}^{|T_1|},$$

for fixed $\boldsymbol{b}_K \cup \boldsymbol{a}_{J\setminus(T_1\cup K)} \cup \boldsymbol{a}_{J^c\setminus T_2} \cup \boldsymbol{a}_{T_2} \in \mathbb{R}^{d-|T_1|}$. Apparently, some integrals in (2.54) appear several times. Regardless of whether elements belong to T_2 or to $J \setminus (T_1 \cup K)$, if the other subsets T_1, K and $J^c \setminus T_2$ stay the same, we obtain the same integral. However, we will see by case differentiation that the sign changes depending on the cardinality of T_2 and $J \setminus (T_1 \cup K)$ so that multiple summands cancel out each other. In the following we examine the two types of scenarios that $|T_2 \cup J \setminus (T_1 \cup K)|$ is odd and that $|T_2 \cup J \setminus (T_1 \cup K)|$ is even.

If the number of elements in $T_2 \cup J \setminus (T_1 \cup K)$ is odd, then the cardinality $|T_2|$ is even (and thus $(-1)^{|T_2|} = 1$) whereas $|J \setminus (T_1 \cup K)|$ is odd or vice versa. That means, if T_2 plays the role of $J \setminus (T_1 \cup K)$ and $J \setminus (T_1 \cup K)$ plays the role of T_2, then the sign changes. As a consequence the two summands, with $J \setminus (T_1 \cup K)$ and T_2 changing places, add up to zero.

If the number of elements in $T_2 \cup J \setminus (T_1 \cup K)$ is even but not equal to zero, we fix one element $j_0 \in T_2 \cup J \setminus (T_1 \cup K)$. Without loss of generality assume that $j_0 \in T_2$ (the case that $j_0 \in J \setminus (T_1 \cup K)$ can be proven analogously). Then the cardinality $|(T_2 \cup J \setminus (T_1 \cup K)) \setminus \{j_0\}|$ is odd and, by the same argumentation as above, those summands with $T_2 \setminus \{j_0\}$ and $J \setminus (T_1 \cup K)$ reversing roles cancel out each other by the same argumentation as above.

Since we sum over all subsets in $\{1, \dots, d\}$, all summands with $T_2 \neq \emptyset$ and $J \setminus (T_1 \cup K) \neq \emptyset$ for any fixed K, T_1 and $J^c \setminus T_2$ vanish. Hence $K = J \setminus T_1$ and (2.54) reduces to

$$\begin{aligned}
&\sum_{J\subseteq\{1,\dots,d\}} \int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]} \mu_G(I_{\boldsymbol{y}}^{\boldsymbol{a},\boldsymbol{b};J})\, \mu_F^{\boldsymbol{a};J}(d\boldsymbol{y}) \\
&= \sum_{J\subseteq\{1,\dots,d\}} \sum_{T_1\subseteq J} (-1)^{d-|J\setminus T_1|} \int_{(\boldsymbol{a}_{T_1},\boldsymbol{b}_{T_1}]} G\big((\boldsymbol{y}_{T_1}-) \cup \boldsymbol{a}_{J^c} \cup \boldsymbol{b}_{J\setminus T_1}\big)\, \mu_F^{(\boldsymbol{b}^{\boldsymbol{a};J\setminus T_1});T_1}(d\boldsymbol{y}_{T_1}) \\
&= \sum_{T_1,K\subseteq\{1,\dots,d\} \text{ disjoint}} (-1)^{d-|K|} \int_{(\boldsymbol{a}_{T_1},\boldsymbol{b}_{T_1}]} G\big((\boldsymbol{y}_{T_1}-) \cup \boldsymbol{a}_{(T_1\cup K)^c} \cup \boldsymbol{b}_K\big)\, \mu_F^{(\boldsymbol{b}^{\boldsymbol{a};K});T_1}(d\boldsymbol{y}_{T_1}),
\end{aligned}$$

which implies (2.52). □

In dimension $d = 2$, for instance, the integration by parts formula (2.51) is given by

$$\begin{aligned}
&\int_{(a_1,b_1]\times(a_2,b_2]} F(x_1,x_2)\, \mu_G(d(x_1,x_2)) \\
&= \int_{(a_1,b_1]\times(a_2,b_2]} G((x_1,x_2)-)\, \mu_F(d(x_1,x_2)) \\
&\quad - \int_{(a_1,b_1]} G(x_1-,b_2)\, \mu_F^{\boldsymbol{b};\{1\}}(dx_1) + \int_{(a_1,b_1]} G(x_1-,a_2)\, \mu_F^{\boldsymbol{a};\{1\}}(dx_1)
\end{aligned}$$

$$-\int_{(a_2,b_2]} G(b_1,x_2-)\,\mu_F^{\boldsymbol{b};\{2\}}(dx_2)+\int_{(a_2,b_2]} G(a_1,x_2-)\,\mu_F^{\boldsymbol{a};\{2\}}(dx_2)$$
$$+F(b_1,b_2)G(b_1,b_2)-F(b_1,a_2)G(b_1,a_2)-F(a_1,b_2)G(a_1,b_2)+F(a_1,a_2)G(a_1,a_2)$$

for any $\boldsymbol{a}=(a_1,a_2)$ and $\boldsymbol{b}=(b_1,b_2)$ in $\mathbb{R}^2$, provided the assumptions of Theorem 2.6.4 are fulfilled.

In the following corollary we use the same notation as introduced in Theorem 2.6.4.

Corollary 2.6.5 *Let $F:\mathbb{R}^d\to\mathbb{R}$ and $G:\mathbb{R}^d\to\mathbb{R}$ be right continuous functions and assume that the functions $F^{\mathbf{c};J}$ and $G^{\mathbf{c};J}$ are locally of bounded $|J|$-fold variation for every $\boldsymbol{c}\in\mathbb{R}^d$ and every nonempty subset $J\subseteq\{1,\ldots,d\}$. Further assume that for every $\boldsymbol{c}\in\mathbb{R}^d$ and every nonempty subset $J\subseteq\{1,\ldots,d\}$ the maps $\boldsymbol{x}_J\mapsto F_\pm^{\mathbf{c};J}(\boldsymbol{x}_J)$ and $\boldsymbol{x}_J\mapsto G_\pm^{\mathbf{c};J}(\boldsymbol{x}_J-)$ are $(\mathcal{B}(\mathbb{R}^{|J|}),\mathcal{B}(\mathbb{R}))$-measurable and the integrals $\int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]}F_\pm^{\mathbf{c};J}(\boldsymbol{x}_J)\,\mu^\pm_{G^{\mathbf{c};J}}(d\boldsymbol{x}_J)$ and $\int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]}G_\pm^{\mathbf{c};J}(\boldsymbol{x}_J-)\,\mu^\pm_{F^{\mathbf{c};J}}(d\boldsymbol{x}_J)$ exist for every finite interval $(\boldsymbol{a}_J,\boldsymbol{b}_J]\subsetneq\mathbb{R}^{|J|}$, where $F_\pm^{\mathbf{c};J}$ and $G_\pm^{\mathbf{c};J}$ are $|J|$-fold monotonically increasing and right continuous functions satisfying $F^{\mathbf{c};J}=F_+^{\mathbf{c};J}-F_-^{\mathbf{c};J}$ and $G^{\mathbf{c};J}=G_+^{\mathbf{c};J}-G_-^{\mathbf{c};J}$. If the integral $\int_{\mathbb{R}^d}F(\boldsymbol{x})\,\mu_G(d\boldsymbol{x})$ exists and the limits*

$$\lim_{a_1,\ldots,a_d\to-\infty,b_1,\ldots,b_d\to+\infty}\sum_{K\subseteq\{1,\ldots,d\}}(-1)^{d-|K|}G(\boldsymbol{b}^{\boldsymbol{a};K})F(\boldsymbol{b}^{\boldsymbol{a};K}),$$
$$\lim_{a_1,\ldots,a_d\to-\infty,b_1,\ldots,b_d\to+\infty}\sum_{J,K\subsetneq\{1,\ldots,d\}\text{ disjoint},J\neq\emptyset}(-1)^{d-|K|}\int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]}G(\boldsymbol{y}_K^{\boldsymbol{a},\boldsymbol{b};J}-)\,\mu_F^{(\boldsymbol{b}^{\boldsymbol{a};K});J}(d\boldsymbol{y})$$

exist and equal zero, then

$$\int_{\mathbb{R}^d}F(\boldsymbol{x})\,\mu_G(d\boldsymbol{x}) \;=\; (-1)^d\int_{\mathbb{R}^d}G(\boldsymbol{x}-)\,\mu_F(d\boldsymbol{x}). \tag{2.55}$$

Here, we think of the expression "$\lim_{a_1,\ldots,a_d\to-\infty,b_1,\ldots,b_d\to+\infty}(\ldots)$" as convergence of the net $(\ldots)_{(n_1,\ldots,n_{2d})\in\mathbb{N}^{2d}}$, with $(-a_1,\ldots,-a_d,b_1,\ldots,b_d)$ playing the role of $(n_1,\ldots,n_{2d})$.

Proof For any $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^d$ with $\boldsymbol{a}<\boldsymbol{b}$ we have

$$\begin{aligned}\int_{(\boldsymbol{a},\boldsymbol{b}]}F(\boldsymbol{x})\,\mu_G(d\boldsymbol{x}) \;=\;& \sum_{K\subseteq\{1,\ldots,d\}}(-1)^{d-|K|}\,G(\boldsymbol{b}^{\boldsymbol{a};K})F(\boldsymbol{b}^{\boldsymbol{a};K})\\ &+\sum_{J,K\subsetneq\{1,\ldots,d\}\text{ disjoint},J\neq\emptyset}(-1)^{d-|K|}\int_{(\boldsymbol{a}_J,\boldsymbol{b}_J]}G(\boldsymbol{y}_K^{\boldsymbol{a},\boldsymbol{b};J}-)\,\mu_F^{(\boldsymbol{b}^{\boldsymbol{a};K});J}(d\boldsymbol{y})\\ &+(-1)^d\int_{(\boldsymbol{a},\boldsymbol{b}]}G(\boldsymbol{y}-)\,\mu_F(d\boldsymbol{y})\end{aligned} \tag{2.56}$$

by Theorem 2.6.4. We note that the limit $\lim_{a_1,\ldots,a_d\to-\infty,b_1,\ldots,b_d\to+\infty}\int_{(\boldsymbol{a},\boldsymbol{b}]}G(\boldsymbol{y}-)\,\mu_F(d\boldsymbol{y})$ exists because the limit of the integral on the left-hand side of (2.56) exists and the limits of the other integrals on the right-hand side of (2.56) exist (and equal zero) by our assumptions. The assertion follows by letting $a_1, \dots, a_d \to -\infty, b_1, \dots, b_d \to +\infty$. $\square$

In the literature, functions $F : [\boldsymbol{a}, \boldsymbol{b}] \to \mathbb{R}$ with finite total d-fold variation on interval $[\boldsymbol{a}, \boldsymbol{b}] \subsetneq \mathbb{R}^d$ are sometimes called *of bounded Hardy-Krause variation on* $[\boldsymbol{a}, \boldsymbol{b}]$ *anchored at* $\boldsymbol{a}$ *(resp. at* $\boldsymbol{b}$*)* if the total $|J|$-fold variation of the functions $F^{\boldsymbol{a};J}$ (resp. $F^{\boldsymbol{b};J}$) on $[\boldsymbol{a}_J, \boldsymbol{b}_J]$ is also finite for each nonempty subset $J \subsetneq \{1, \dots, d\}$. We may thus define functions $F : \mathbb{R}^d \to \mathbb{R}$ to be *locally of bounded Hardy-Krause variation* if F is of bounded Hardy-Krause variation on $\boldsymbol{I_c}$ anchored at $\boldsymbol{c}$ for every $\boldsymbol{c} \in \mathbb{R}^d$ and $\boldsymbol{I_c} \in \mathcal{I}_{\boldsymbol{c}}$, where $\mathcal{I}_{\boldsymbol{c}}$ denotes the set of all compact intervals $I_{\boldsymbol{c};1} \times \cdots \times I_{\boldsymbol{c};d}$ having c_j as one of the endpoints of $I_{\boldsymbol{c};j}$ for $j = 1, \dots, d$. We note that functions $F : [\boldsymbol{a}, \boldsymbol{b}] \to \mathbb{R}$ of bounded Hardy-Krause variation (anchored at one of the endpoints) are measurable as they are decomposable in completely monotonically increasing and thus measurable functions by Theorem 2 in [1] and Theorem 3.2 in [2] (cf. Theorem 3.1 in [2]).

In [63, Theorem 15], Radulović et al. present a multivariate integration by parts formula on compact intervals by supposing the involved functions to be of bounded Hardy-Krause variation anchored at one of the endpoints of this compact interval. Extended to integrals on $\mathbb{R}^d$, we would probably need to assume that the involved functions are locally of bounded Hardy-Krause variation. The latter assumption coincides with our assumption that $F^{\mathbf{c};J}$ and $G^{\mathbf{c};J}$ are locally of bounded J-fold variation for every $\boldsymbol{c} \in \mathbb{R}^d$ and every nonempty subset $J \subseteq \{1, \dots, d\}$. But in contrast to Theorem 15 in [63], we do not automatically obtain the measurability of the functions of the Jordan decomposition. However, we allow any decomposition of $F^{\mathbf{c};J}$ and $G^{\mathbf{c};J}$ into two $|J|$-fold monotonically increasing right continuous functions $F_+^{\mathbf{c};J}$, $F_-^{\mathbf{c};J}$ and $G_+^{\mathbf{c};J}$, $G_-^{\mathbf{c};J}$, respectively, at the expense of slightly stronger conditions on the existence of the integrals. Then, in some applications as for instance in Chapter 3, the measurability assumptions of Theorem 2.6.4 are not difficult to check.

Chapter 3

Extended continuous mapping approach to the asymptotics of von Mises-statistics

3.1 Introduction

The asymptotics of von Mises-statistics, or V-statistics for short, and the closely related U-statistics were first studied by Halmos [42], Hoeffding [44] and von Mises [74]. The most common approach to study the limit distribution of these statistics is based on the Hoeffding decomposition [44]. Using this decomposition many central limit theorems have been established for non-degenerate and degenerate U- and V-statistics. For independent identically distributed sequences of random variables we refer for instance to the standard textbooks Denker [27], Lee [50], Sen [66, 67] and Serfling [68]. For dependent sequences of random variables, the asymptotics of non-degenerate and degenerate U- and V-statistics have been studied by means of Hoeffding's decomposition among others in Dehling [24], Dehling and Wendler [26], Denker and Keller [28], Sen [65] and Yoshihara [80] for weakly dependent data under various mixing conditions, in Dewan and Prakasa Rao [30, 31] and Garg and Dewan [37, 38] for associated random variables, and in Dehling and Taqqu [25] for strongly dependent data (data with long-memory). Other approaches to obtain the limit distribution of U- and V-statistics under weak dependence are for instance based on a spectral (resp. wavelet) decomposition of the kernel function, see Dewan and Prakasa Rao [29] for non-degenerate and degenerate U-statistics and Leucht and Neumann [53, 54] (resp. Leucht [52]) for degenerate U- and V-statistics. Further, Zhou [83] assumes the kernel function to admit a Fourier representation and investigates non-degenerate and degenerate V-statistics from a Fourier analysis point of view. In [10], Beutner and Zähle use quasi-Hadamard differentiability (introduced in [9]) of the V-functionals to derive the asymptotic behavior of U- and V-statistics of weakly dependent data, which is also a suitable method to derive the limit distribution for a certain class of U- and V-statistics based on long-memory sequences as shown in Beutner et al. [8].

In [11], Beutner and Zähle present a new representation for U- and V-statistics so that the asymptotics of non-degenerate and degenerate U- and V-statistics can be derived therefrom by a direct application of the continuous mapping theorem. The objective of this chapter is to put forward this continuous mapping approach. In [11] the authors restricted themselves to two-sample statistics of degree $d = 2$ and kernel functions $h_n = h$ in the setting below. In what follows we will study multi-sample statistics of degree d with kernels h_n depending on n. However, we will concentrate on V-statistics only. The corresponding results for U-statistics can be inferred from those for V-statistics because their asymptotic distributions coincide under suitable assumptions. For instance, one-sample U- and V-statistics of degree $d = 2$ have the same asymptotic distribution, if the kernel function $h_n = h$ satisfies $\mathbb{E}[|h(X,X)|] < \infty$ for some random variable X with distribution function F in the setting below, see Remark 2.5 of [10]. In the following, we will retain the notation introduced in Chapter 2.

Let $d \in \mathbb{N}$. For every $n \in \mathbb{N}_0 := \{0, 1, \ldots\}$, let $h_n : \mathbb{R}^d \to \mathbb{R}$ be a Borel measurable function and consider the functional $\mathcal{V}_{h_n} : \mathbf{F}_{h_n} \to \mathbb{R}$ defined by

$$\mathcal{V}_{h_n}(F^{(1)}, \ldots, F^{(d)}) := \int_{\mathbb{R}^d} h_n(x_1, \ldots, x_d)\,(\mu_{F^{(1)}} \otimes \cdots \otimes \mu_{F^{(d)}})(d(x_1, \ldots, x_d)), \tag{3.1}$$

where $\mathbf{F}_{h_n}$ is the set of all d-tuples $(F^{(1)}, \ldots, F^{(d)})$ of distribution functions on $\mathbb{R}$ for which the integral in (3.1) exists and μ_F refers to the Borel probability measure generated by F. Let $(F_0^{(1)}, \ldots, F_0^{(d)}) \in \mathbf{F}_{h_0}$ be fixed, and let $\widehat{F}_n^{(j)}$ be an estimator of $F_0^{(j)}$ for every $j = 1, \ldots, d$ and $n \in \mathbb{N}$ such that $(\widehat{F}_n^{(1)}, \ldots, \widehat{F}_n^{(d)}) \in \mathbf{F}_{h_n}$ (ω-wise) for every $n \in \mathbb{N}$. Then

$$\mathcal{V}_{h_n}(\widehat{F}_n^{(1)}, \ldots, \widehat{F}_n^{(d)}) = \int_{\mathbb{R}^d} h_n(x_1, \ldots, x_d)\,(\mu_{\widehat{F}_n^{(1)}} \otimes \cdots \otimes \mu_{\widehat{F}_n^{(d)}})(d(x_1, \ldots, x_d)) \tag{3.2}$$

can provide a reasonable estimator for the expression in (3.1). In the special case where $\widehat{F}_n^{(j)}$ is the empirical distribution function $\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{[X_i^{(j)},\infty)}$ of random variables $X_1^{(j)}, \ldots, X_n^{(j)}$, $j = 1, \ldots, d$, the estimator in (3.2) is a *d-sample V-statistic* of degree d with kernel h_n, i.e.

$$\mathcal{V}_{h_n}(\widehat{F}_n^{(1)}, \ldots, \widehat{F}_n^{(d)}) = \frac{1}{n^d} \sum_{i_1=1}^n \cdots \sum_{i_d=1}^n h_n\big(X_{i_1}^{(1)}, \ldots, X_{i_d}^{(d)}\big).$$

This expression is a suitable estimator of $\mathcal{V}_{h_n}(F_0^{(1)}, \ldots, F_0^{(d)})$ when $X_1^{(j)}, \ldots, X_n^{(j)}$ are identically distributed according to $F_0^{(j)}$, $j = 1, \ldots, d$, (and "sufficiently independent"), and $\{X_i^{(1)}\}_{i=1}^n, \ldots, \{X_i^{(d)}\}_{i=1}^n$ are independent. However we will not insist on this particular setting. In the general case the empirical error has the decomposition

$$\mathcal{V}_{h_n}(\widehat{F}_n^{(1)}, \ldots, \widehat{F}_n^{(d)}) - \mathcal{V}_{h_n}(F_0^{(1)}, \ldots, F_0^{(d)})$$

$$= \sum_{\emptyset \neq J \subseteq \{1,\dots,d\}} \int_{\mathbb{R}^J} h_{n;J,F_0}\big((x_j)_{j\in J}\big) \Big(\bigotimes_{j\in J} \mu_{(\widehat{F}_n^{(j)} - F_0^{(j)})}\Big)(d((x_j)_{j\in J})), \tag{3.3}$$

provided the involved integrals all exist, where $h_{n;\{1,\dots,d\},F_0}(x_1,\dots,x_d) := h_n(x_1,\dots,x_d)$ and $h_{n;J,F_0} : \mathbb{R}^J \to \mathbb{R}$ for $\emptyset \neq J \subsetneq \{1,\dots,d\}$ is given by

$$h_{n;J,F_0}(\boldsymbol{y}) := \int_{\mathbb{R}^{J^c}} h_n(\boldsymbol{y}_J\boldsymbol{x}) \Big(\bigotimes_{j\in J^c} \mu_{F_0^{(j)}}\Big)(d((x_j)_{j\in J^c}))$$

for $\boldsymbol{y} := (y_j)_{j\in J} \in \mathbb{R}^J$, $\boldsymbol{x} := (x_1,\dots,x_d) \in \mathbb{R}^d$ and $\boldsymbol{y}_J\boldsymbol{x}$ as defined in (2.5), see Lemma 3.2.2. Recall that $J^c := \{1,\dots,d\} \setminus J$. The analogue of this decomposition for one-sample V-statistics with symmetric kernel functions is sometimes called *von Mises decomposition* of $\mathcal{V}_{h_n}(\widehat{F}_n,\dots,\widehat{F}_n) - \mathcal{V}_{h_n}(F_0,\dots,F_0)$; see [48, p. 40].

In Chapter 2 we derived a multivariate integration by parts formula. Applying this formula to the integrals in (3.3), we obtain under suitable assumptions (see Lemma 3.2.3 below) that

$$\begin{aligned}
&\mathcal{V}_{h_n}(\widehat{F}_n^{(1)},\dots,\widehat{F}_n^{(d)}) - \mathcal{V}_{h_n}(F_0^{(1)},\dots,F_0^{(d)}) \\
&= \sum_{\emptyset \neq J \subseteq \{1,\dots,d\}} (-1)^{|J|} \int_{\mathbb{R}^J} \prod_{j\in J} \big(\widehat{F}_n^{(j)}(x_j-) - F_0^{(j)}(x_j-)\big)\, \mu_{h_{n;J,F_0}}(d((x_j)_{j\in J})).
\end{aligned} \tag{3.4}$$

Note at this point that $\mu_{h_{n;J,F_0}}$ for $J \subseteq \{1,\dots,d\}$ can be signed measures. In the following, we will refer to representation (3.4) as *generalized von Mises decomposition.* Based on this representation the asymptotic distribution of V-statistics may be derived by a direct application of the extended continuous mapping theorem. Except for some minor assumptions, we mainly need the limit distribution of $a_n(\widehat{F}_n^{(j)} - F_0^{(j)})$, $j = 1,\dots,d$, to determine the limit distribution of $a_n\big(\mathcal{V}_{h_n}(\widehat{F}_n^{(1)},\dots,\widehat{F}_n^{(d)}) - \mathcal{V}_{h_n}(F_0^{(1)},\dots,F_0^{(d)})\big)$ for a suitable sequence (a_n) in $(0,\infty)$ with $a_n \to \infty$ via this extended continuous mapping approach, provided the assumptions of the generalized von Mises decomposition are fulfilled.

If $\widehat{F}_n$ is the empirical distribution function, weak convergence theorems (with respect to weighted sup-metrics) for the empirical process $\sqrt{n}(\widehat{F}_n - F)$ have been established under various conditions. For instance, we refer to Shorack and Wellner [70] for independent identically distributed data, to Arcones and Yu [4], Shao and Yu [69] and Wu [78] for stationary sequences of weakly dependent random variables, and to Beutner et al. [8] for stationary sequences of strongly dependent random variables. More details can be found subsequent to Remark 3.3.2.

In Chapter 1, we proved such a weak convergence theorem for the local empirical process of piece-wise locally stationary time series. Combined with the extended continuous mapping approach, this enables to investigate the asymptotic distribution of (weighted) V-statistics for non-stationary time series. In the literature, U- and V-statistics and their asymptotics have mainly been studied for stationary sequences of random variables, see

Dehling and Taqqu [25], Dehling and Wendler [26], Denker and Keller [28], Dewan and Prakasa Rao [29, 30, 31], Garg and Dewan [37, 38], Leucht [52], Sen [65] and Yoshihara [80]. In [83], Zhou already treated weighted V-statistics of degree 2 for non-stationary time series from the perspective of Fourier analysis. With our approach, this result can be reproduced under similar assumptions and generalized to weighted V-statistics of higher degree. Moreover, we regain many results existing in the literature concerning one-sample V-statistics of degree greater than or equal to 2 for stationary sequences of random variables, which are suitable for kernel functions of bounded variation. In addition, multi-sample V-statistics of degree d can be dealt with.

The rest of the chapter is organized as follows. In Section 3.2 we clarify the assumptions under which the generalized von Mises-representation holds and give some illustrating examples. In Section 3.3 we present our main result: the weak limit theorem for one-sample and multi-sample V-statistics. More precisely, we determine the limit distribution of the vector-valued random variable

$$a_n \left(\begin{bmatrix} \mathcal{V}_{h_{n,1}}(\widehat{F}_n^{(11)}, \ldots, \widehat{F}_n^{(1d_1)}) \\ \vdots \\ \mathcal{V}_{h_{n,k}}(\widehat{F}_n^{(k1)}, \ldots, \widehat{F}_n^{(kd_k)}) \end{bmatrix} - \begin{bmatrix} \mathcal{V}_{h_{n,1}}(F_0^{(11)}, \ldots, F_0^{(1d_1)}) \\ \vdots \\ \mathcal{V}_{h_{n,k}}(F_0^{(k1)}, \ldots, F_0^{(kd_k)}) \end{bmatrix} \right)$$

for different kernel functions $h_{n,1} : \mathbb{R}^{d_1} \to \mathbb{R}, \ldots, h_{n,k} : \mathbb{R}^{d_k} \to \mathbb{R}$ and distribution functions $F_0^{(11)}, \ldots, F_0^{(kd_k)}$ with estimators $\widehat{F}_n^{(11)}, \ldots, \widehat{F}_n^{(kd_k)}$, where in the case of one-sample V-statistics $\widehat{F}_n^{(ij)} = \widehat{F}_n$ and $F_0^{(ij)} = F_0$ for all $i = 1, \ldots, k$ and $j = 1, \ldots, d_i$. Due to this vector-valued result, also the asymptotic distribution of suitable compositions of different V-statistics such as the skewness or kurtosis of probability distributions can be studied. The example of the skewness is carried out in detail in Subsection 3.3.3. In Section 3.4, we finally investigate the asymptotics of weighted V-statistics of degree d for non-stationary time series by means of the extended continuous mapping theorem and the weak convergence theorem of the local empirical process from Chapter 1.

3.2 Generalized von Mises decomposition

3.2.1 Assumptions and proof of representation (3.4)

First of all, we state an assumption on h_n, $(F_0^{(1)}, \ldots, F_0^{(d)})$ and $(\widehat{F}_n^{(1)}, \ldots, \widehat{F}_n^{(d)})$ that ensures the well-definedness of the representation (3.3).

Assumption 3.2.1 *For all $n \in \mathbb{N}$ we have $\mathbb{P}$-a.s. for every subset $J \subseteq \{1, \ldots, d\}$ that* $\int_{\mathbb{R}^d} \left| h_n(x_1, \ldots, x_d) \right| \left(\bigotimes_{j \in J^c} \mu_{F_0^{(j)}} \otimes \bigotimes_{j \in J} \mu_{\widehat{F}_n^{(j)}} \right) (d((x_j)_{j \in J^c}, (x_j)_{j \in J})) < \infty.$

We note that we already assumed in the introduction that $(\widehat{F}_n^{(1)}, \ldots, \widehat{F}_n^{(d)}) \in \mathbf{F}_{h_n}$ (ω-wise) for every $n \in \mathbb{N}$, which corresponds to Assumption 3.2.1 for $J = \{1, \ldots, d\}$.

If $\widehat{F}_n^{(j)}$ is the empirical distribution function of the random variables $X_1^{(j)}, \ldots, X_n^{(j)}$, $j = 1, \ldots, d$, then Assumption 3.2.1 is fulfilled for every nonempty subset $J \subseteq \{1, \ldots, d\}$ and thus boils down to the assumption that $(F_0^{(1)}, \ldots, F_0^{(d)}) \in \mathbf{F}_{h_n}$ for all $n \in \mathbb{N}$.

Lemma 3.2.2 *Under Assumption 3.2.1, representation (3.3) holds $\mathbb{P}$-a.s. for all $n \in \mathbb{N}$.*

Proof By Fubini's theorem, we have

$$
\begin{aligned}
&\mathcal{V}_{h_n}(\widehat{F}_n^{(1)}, \ldots, \widehat{F}_n^{(d)}) - \mathcal{V}_{h_n}(F_0^{(1)}, \ldots, F_0^{(d)}) \\
&= \int_{\mathbb{R}^d} h_n(x_1, \ldots, x_d) \Big(\bigotimes_{j=1}^{d} \big(\mu_{(\widehat{F}_n^{(j)} - F_0^{(j)})} + \mu_{F_0^{(j)}}\big)\Big)(d(x_1, \ldots, x_d)) - \mathcal{V}_{h_n}(F_0^{(1)}, \ldots, F_0^{(d)}) \\
&= \sum_{J \subseteq \{1,\ldots,d\}} \int_{\mathbb{R}^d} h_n(x_1, \ldots, x_d) \Big(\bigotimes_{j \in J^{\mathrm{c}}} \mu_{F_0^{(j)}} \otimes \bigotimes_{j \in J} \mu_{(\widehat{F}_n^{(j)} - F_0^{(j)})}\Big)(d((x_j)_{j \in J^{\mathrm{c}}}, (x_j)_{j \in J})) \\
&\quad - \int_{\mathbb{R}^d} h_n(x_1, \ldots, x_d) \Big(\bigotimes_{j=1}^{d} \mu_{F_0^{(j)}}\Big)(d(x_1, \ldots, x_d)) \\
&= \sum_{\emptyset \neq J \subseteq \{1,\ldots,d\}} \int_{\mathbb{R}^J} h_{n;J,F_0}\big((x_j)_{j \in J}\big) \Big(\bigotimes_{j \in J} \mu_{(\widehat{F}_n^{(j)} - F_0^{(j)})}\Big)(d((x_j)_{j \in J}))
\end{aligned}
$$

$\mathbb{P}$-a.s. for all $n \in \mathbb{N}$. □

In line with the notation in Section 2.6, let $\boldsymbol{x}_K := (x_k)_{k \in K}$ and let the function $(h_{n;J,F_0})^{\boldsymbol{c}_J;K} : \mathbb{R}^K \to \mathbb{R}$ be defined as in (2.4) for any nonempty subsets $J, K \subseteq \{1, \ldots, d\}$ with $K \subseteq J$, $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$ and $n \in \mathbb{N}$. If we apply for $\mathbb{P}$-almost every ω the integration by parts formula in the form of Corollary 2.6.5 to each summand on the right-hand side of (3.3), we immediately obtain the following result.

Lemma 3.2.3 *Suppose that the following conditions are fulfilled.*

(a) *Assumption 3.2.1 holds.*

(b) *For every nonempty subset $J \subseteq \{1, \ldots, d\}$ and every $n \in \mathbb{N}$, the function $h_{n;J,F_0}$ is right continuous and $(h_{n;J,F_0})^{\boldsymbol{c}_J;K}$ is locally of bounded $|K|$-fold variation for every nonempty subset $K \subseteq J$ and every fixed $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$.*

(c) *For every nonempty subset $J \subseteq \{1, \ldots, d\}$ and every $n \in \mathbb{N}$, the function $(h_{n;J,F_0})_{\pm}^{\boldsymbol{c}_J;K}$ is $(\mathcal{B}(\mathbb{R}^{|K|}), \mathcal{B}(\mathbb{R}))$-measurable for every nonempty subset $K \subseteq J$ and every fixed $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$, and the integral*

$$
\int_{\boldsymbol{I}_K^{\boldsymbol{a},\boldsymbol{b}}} (h_{n;J,F_0})_{\pm}^{\boldsymbol{c}_J;K}(\boldsymbol{x}_K) \Big(\bigotimes_{l \in L \cap K} \mu_{\widehat{F}_n^{(l)}} \otimes \bigotimes_{j \in (J \setminus L) \cap K} \mu_{F_0^{(j)}}\Big)(d((x_l)_{l \in L \cap K}, (x_j)_{j \in (J \setminus L) \cap K}))
$$

exists for all subsets $K, L \subseteq J$ with $K \neq \emptyset$, $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$ and for every finite interval $\boldsymbol{I}_K^{\boldsymbol{a},\boldsymbol{b}} := (\boldsymbol{a}_{L \cap K}, \boldsymbol{b}_{L \cap K}] \times (\boldsymbol{a}_{(J \setminus L) \cap K}, \boldsymbol{b}_{(J \setminus L) \cap K}] \subsetneq \mathbb{R}^{|K|}$, $\mathbb{P}$-a.s., where $(h_{n;J,F_0})_{+}^{\boldsymbol{c}_J;K}$ and

$(h_{n;J,F_0})_-^{\boldsymbol{c}_J;K}$ *are* $|K|$*-fold monotonically increasing and right continuous functions satisfying* $(h_{n;J,F_0})^{\boldsymbol{c}_J;K} = (h_{n;J,F_0})_+^{\boldsymbol{c}_J;K} - (h_{n;J,F_0})_-^{\boldsymbol{c}_J;K}$.

(d) *The following limits exist and equal zero* $\mathbb{P}$*-a.s. for every nonempty subset* $J \subseteq \{1,\dots,d\}$ *and for every* $n \in \mathbb{N}$:

$$\lim_{\{-a_j,\, b_j\}_{j\in J}\to+\infty} \Big[\sum_{K\subseteq J}(-1)^{|J|-|K|}\prod_{j\in J}\Big(\widehat{F}_n^{(j)}\big(b_j^{\boldsymbol{a}_J;K}\big) - F_0^{(j)}\big(b_j^{\boldsymbol{a}_J;K}\big)\Big)h_{n;J,F_0}\big(\boldsymbol{b}_J^{\boldsymbol{a}_J;K}\big)\Big],$$

$$\lim_{\{-a_j,\, b_j\}_{j\in J}\to+\infty} \Big[\sum_{L,K\subsetneq J \text{ disjoint}, L\neq\emptyset}(-1)^{|J|-|K|} \cdot \prod_{k\in K}\big(\widehat{F}_n^{(k)}(b_k) - F_0^{(k)}(b_k)\big)\prod_{j\in J\setminus(L\cup K)}\big(\widehat{F}_n^{(j)}(a_j) - F_0^{(j)}(a_j)\big) \cdot \int_{(\boldsymbol{a}_L,\boldsymbol{b}_L]}\prod_{j\in L}\big(\widehat{F}_n^{(j)}(x_j-) - F_0^{(j)}(x_j-)\big)\mu_{h_{n;J,F_0}}^{(\boldsymbol{b}_J^{\boldsymbol{a}_J;K});L}(d((x_j)_{j\in L}))\Big],$$

where we use the convention that $\prod_{j\in\emptyset}(\dots) := 1$ *and* $\boldsymbol{b}_J^{\boldsymbol{a}_J;K} := (b_j^{\boldsymbol{a}_J;K})_{j\in J}$ *is defined as in (2.2).*

Then the representation (3.4) holds true $\mathbb{P}$*-a.s. for all* $n \in \mathbb{N}$.

Analogously to the notation in Corollary 2.6.5, the expression "$\lim_{\{-a_j,b_j\}_{j\in J}\to\infty}(\dots)$" in Lemma 3.2.3 is understood as convergence of the net $(\dots)_{(n_1,\dots,n_{2|J|})\in\mathbb{N}^{2|J|}}$, where $n_1,\dots,n_{2|J|}$ corresponds to $(-a_j)_{j\in J} \cup (b_j)_{j\in J}$.

Proof According to condition (a) the representation in (3.3) is valid, see Lemma 3.2.2. To show that the representation (3.4) holds true $\mathbb{P}$-a.s. for all $n \in \mathbb{N}$, it suffices to prove that for every nonempty subset $J \subseteq \{1,\dots,d\}$, $n \in \mathbb{N}$ and for $\mathbb{P}$-almost every ω the conditions of the integration by parts formula (in form of Corollary 2.6.5) on the functions $\prod_{j\in J}(\widehat{F}_n^{(j)} - F_0^{(j)})$ and $h_{n;J,F_0}$ are satisfied. The claim then follows immediately from Corollary 2.6.5 applied to each summand on the right-hand side in (3.3) for every $n \in \mathbb{N}$ and $\mathbb{P}$-almost every ω.

Let $J \subseteq \{1,\dots,d\}$ be any nonempty subset and $n \in \mathbb{N}$. We first note that for $\mathbb{P}$-almost every ω the right continuity of $\prod_{j\in J}(\widehat{F}_n^{(j)} - F_0^{(j)})$ follows directly from the right continuity of the distribution functions $\widehat{F}_n^{(j)}$ and $F_0^{(j)}$.

As a second step, we show that $\mathbb{P}$-a.s. the product $\big(\prod_{j\in J}(\widehat{F}_n^{(j)} - F_0^{(j)})\big)^{\boldsymbol{c}_J;K}$ is locally of bounded $|K|$-fold variation for every nonempty subset $K \subseteq J$ and every $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$. For $\mathbb{P}$-almost every ω we observe that

$$\Big(\prod_{j\in J}(\widehat{F}_n^{(j)} - F_0^{(j)})\Big)^{\boldsymbol{c}_J;K} = \Big(\sum_{L\subseteq J}(-1)^{|J\setminus L|}\prod_{l\in L}\widehat{F}_n^{(l)}\prod_{j\in J\setminus L}F_0^{(j)}\Big)^{\boldsymbol{c}_J;K} \tag{3.5}$$
$$= \Big(\prod_{j\in J}(\widehat{F}_n^{(j)} - F_0^{(j)})\Big)_+^{\boldsymbol{c}_J;K} - \Big(\prod_{j\in J}(\widehat{F}_n^{(j)} - F_0^{(j)})\Big)_-^{\boldsymbol{c}_J;K}$$

for all nonempty subsets $K \subseteq J$ and all $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$ with

$$\Big(\prod_{j\in J}(\widehat{F}_n^{(j)} - F_0^{(j)})\Big)_+^{\boldsymbol{c}_J;K} := \Big(\sum_{\substack{L\subseteq J,\\ |J\setminus L| \text{ even}}} \prod_{l\in L} \widehat{F}_n^{(l)} \prod_{j\in J\setminus L} F_0^{(j)}\Big)^{\boldsymbol{c}_J;K} \tag{3.6}$$

and

$$\Big(\prod_{j\in J}(\widehat{F}_n^{(j)} - F_0^{(j)})\Big)_-^{\boldsymbol{c}_J;K} := \Big(\sum_{\substack{L\subseteq J,\\ |J\setminus L| \text{ odd}}} \prod_{l\in L} \widehat{F}_n^{(l)} \prod_{j\in J\setminus L} F_0^{(j)}\Big)^{\boldsymbol{c}_J;K}. \tag{3.7}$$

If we can show that $\big(\prod_{j\in J}(\widehat{F}_n^{(j)} - F_0^{(j)})\big)_+^{\boldsymbol{c}_J;K}$ and $\big(\prod_{j\in J}(\widehat{F}_n^{(j)} - F_0^{(j)})\big)_-^{\boldsymbol{c}_J;K}$ are $\mathbb{P}$-a.s. $|K|$-fold monotonically increasing for every nonempty subset $K \subseteq J$ and every $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$, then we can conclude from Proposition 2.4.4(iii) that the expression on the left-hand side of (3.5) is $\mathbb{P}$-a.s. locally of bounded $|K|$-fold variation for every nonempty subset $K \subseteq J$ and every $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$. For the proof that the functions in (3.6) and (3.7) are $\mathbb{P}$-a.s. $|K|$-fold monotonically increasing for every nonempty subset $K \subseteq J$ and every $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$, we note that the product $\prod_{l\in L} \widehat{F}_n^{(l)} \prod_{j\in J\setminus L} F_0^{(j)}$ is $\mathbb{P}$-a.s. $|J|$-fold monotonically increasing for every subset $L \subseteq J$ by part (v) of Proposition 2.2.7. Hence, in view of Proposition 2.2.7 (i), the sums $\sum_{L\subseteq J,\ |J\setminus L| \text{ odd}}(\prod_{l\in L} \widehat{F}_n^{(l)} \prod_{j\in J\setminus L} F_0^{(j)})$ and $\sum_{L\subseteq J,\ |J\setminus L| \text{ even}}(\prod_{l\in L} \widehat{F}_n^{(l)} \prod_{j\in J\setminus L} F_0^{(j)})$ are $\mathbb{P}$-a.s. $|J|$-fold monotonically increasing. Particularly since we consider for $\mathbb{P}$-almost every ω in each summand products of distribution functions that converge to zero if at least one function argument x_j, $j \in J$, tends to $-\infty$, Lemma 2.2.9 yields that the functions in (3.6) and (3.7) are indeed $\mathbb{P}$-a.s. $|K|$-fold monotonically increasing for every nonempty subset $K \subseteq J$ and every $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$.

Third, the functions in (3.6) and (3.7) are not only $\mathbb{P}$-a.s. $|K|$-fold monotonically increasing for every nonempty subset $K \subseteq J$ and every $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$ but also $\mathbb{P}$-a.s. right continuous as composition of right continuous functions.

In a fourth step, we show that $\mathbb{P}$-a.s. for every nonempty subset $K \subseteq J$ and every fixed $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$ the map $\boldsymbol{x}_K \mapsto \lim_{\boldsymbol{u}_K \nearrow \boldsymbol{x}_K}(\prod_{j\in J}(\widehat{F}_n^{(j)} - F_0^{(j)}))_\pm^{\boldsymbol{c}_J;K}(\boldsymbol{u}_K)$ is $(\mathcal{B}(\mathbb{R}^{|K|}), \mathcal{B}(\mathbb{R}))$-measurable. In view of (3.6) and (3.7) we obtain that $\mathbb{P}$-a.s.

$$\begin{aligned}
&\lim_{\boldsymbol{u}_K \nearrow \boldsymbol{x}_K} \Big(\prod_{j\in J}\big(\widehat{F}_n^{(j)} - F_0^{(j)}\big)\Big)_\pm^{\boldsymbol{c}_J;K}(\boldsymbol{u}_K)\\
&= \sum_{\substack{L\subseteq J,\\ |J\setminus L| \text{ even/odd}}} \prod_{k\in L\cap(J\setminus K)} \widehat{F}_n^{(k)}(c_k) \prod_{l\in J\setminus(L\cup K)} F_0^{(l)}(c_l) \prod_{i\in L\cap K} \widehat{F}_n^{(i)}(x_i-) \prod_{j\in(J\setminus L)\cap K} F_0^{(j)}(x_j-)
\end{aligned}$$

holds for every nonempty subset $K \subseteq J$ and every $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$, where the distribution functions $F_0^{(i)}$ and $\widehat{F}_n^{(i)}$ for $i = 1, \dots, d$ are $\mathbb{P}$-a.s. monotonically increasing and hence $\mathbb{P}$-a.s. $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}))$-measurable. For $\mathbb{P}$-almost every ω, the map $\boldsymbol{x}_K \mapsto \lim_{\boldsymbol{u}_K \nearrow \boldsymbol{x}_K} \big(\prod_{j\in J}(\widehat{F}_n^{(j)} - F_0^{(j)})\big)_\pm^{\boldsymbol{c}_J;K}(\boldsymbol{u}_K)$ is thus as product of measurable functions $(\mathcal{B}(\mathbb{R}^{|K|}), \mathcal{B}(\mathbb{R}))$-measurable for every nonempty subset $K \subseteq J$ and every $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$.

Fifth, we observe that $\mathbb{P}$-a.s.

$$\begin{aligned}
&\int_{(\boldsymbol{a}_K,\boldsymbol{b}_K]} \Big| \Big(\prod_{j\in J} \big(\widehat{F}_n^{(j)} - F_0^{(j)}\big)\Big)_{\pm}^{\boldsymbol{c}_J;K} (\boldsymbol{x}_K-) \Big| \, \mu^{\pm}_{(h_{n;J,F_0})^{\boldsymbol{c}_J;K}}(d((x_k)_{k\in K})) \\
&= \sum_{\substack{L\subseteq J,\\ |J\setminus L| \text{ even/odd}}} \prod_{k\in L\cap(J\setminus K)} \widehat{F}_n^{(k)}(c_k) \prod_{i\in J\setminus(L\cup K)} F_0^{(i)}(c_i) \\
&\qquad \cdot \int_{(\boldsymbol{a}_K,\boldsymbol{b}_K]} \Big(\prod_{l\in L\cap K} \widehat{F}_n^{(l)}(x_l-) \prod_{j\in(J\setminus L)\cap K} F_0^{(j)}(x_j-)\Big) \, \mu^{\pm}_{(h_{n;J,F_0})^{\boldsymbol{c}_J;K}}(d((x_k)_{k\in K}))
\end{aligned}$$

holds for all nonempty subsets $K \subseteq J$, $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$ and every finite interval $(\boldsymbol{a}_K, \boldsymbol{b}_K] \subsetneq \mathbb{R}^{|K|}$ in view of (3.6) and (3.7). The latter integral exists for every nonempty subset $K \subseteq J$, $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$ and for every finite interval $(\boldsymbol{a}_K, \boldsymbol{b}_K] \subsetneq \mathbb{R}^{|K|}$, $\mathbb{P}$-a.s., because for $\mathbb{P}$-almost every ω the functions $\widehat{F}_n^{(j)}$ and $F_0^{(j)}$ are as distribution functions bounded by 1 for each $j = 1, \ldots, d$ and $\mu^{\pm}_{(h_{n;J,F_0})^{\boldsymbol{c}_J;K}}((\boldsymbol{a}_K, \boldsymbol{b}_K]) < \infty$ (Recall that $\mu^{\pm}_{(h_{n;J,F_0})^{\boldsymbol{c}_J;K}}((\boldsymbol{a}_K, \boldsymbol{b}_K]) = \Delta_{\boldsymbol{a}_K}^{\boldsymbol{b}_K}(h_{n;J,F_0})_{\pm}^{\boldsymbol{c}_J;K}$ by definition). Moreover, we have $\mathbb{P}$-a.s.

$$\begin{aligned}
&\int_{(\boldsymbol{a}_K,\boldsymbol{b}_K]} (h_{n;J,F_0})_{\pm}^{\boldsymbol{c}_J;K}(\boldsymbol{x}_K) \, \mu^{\pm}_{(\prod_{j\in J}(\widehat{F}_n^{(j)}-F_0^{(j)}))^{\boldsymbol{c}_J;K}}(d((x_k)_{k\in K})) \\
&= \sum_{\substack{L\subseteq J,\\ |J\setminus L| \text{ even/odd}}} \prod_{k\in(J\setminus K)\cap L} \widehat{F}_n^{(k)}(c_k) \prod_{i\in J\setminus(K\cup L)} F_0^{(i)}(c_i) \\
&\qquad \cdot \int_{\boldsymbol{I}_K^{\boldsymbol{a},\boldsymbol{b}}} (h_{n;J,F_0})_{\pm}^{\boldsymbol{c}_J;K}(\boldsymbol{x}_K) \Big(\bigotimes_{l\in L\cap K} \mu_{\widehat{F}_n^{(l)}} \otimes \bigotimes_{j\in(J\setminus L)\cap K} \mu_{F_0^{(j)}} \Big)(d((x_l)_{l\in L\cap K}, (x_j)_{j\in(J\setminus L)\cap K}))
\end{aligned}$$

for every nonempty subset $K \subseteq J$, $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$ and every finite interval $(\boldsymbol{a}_K, \boldsymbol{b}_K] \subsetneq \mathbb{R}^{|K|}$, where $\boldsymbol{I}_K^{\boldsymbol{a},\boldsymbol{b}}$ is defined as in assumption (c). The fact that the latter integral exists for every nonempty subset $K \subseteq J$, $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$ and every finite interval $(\boldsymbol{a}_K, \boldsymbol{b}_K] \subsetneq \mathbb{R}^{|K|}$, $\mathbb{P}$-a.s., is thus ensured by assumption (c).

Finally, in view of Assumption 3.2.1, $\int_{\mathbb{R}^J} h_{n;J,F_0}(\boldsymbol{x}_J) \, \mu_{(\prod_{j\in J}(\widehat{F}_n^{(j)}-F_0^{(j)}))}(d((x_j)_{j\in J}))$ exists $\mathbb{P}$-a.s. because

$$\begin{aligned}
&\int_{\mathbb{R}^J} |h_{n;J,F_0}(\boldsymbol{x}_J)| \, \mu^{\pm}_{(\prod_{j\in J}(\widehat{F}_n^{(j)}-F_0^{(j)}))}(d((x_j)_{j\in J})) \\
&= \sum_{\substack{L\subseteq J,\\ |J\setminus L| \text{ even/odd}}} \int_{\mathbb{R}^J} |h_{n;J,F_0}(\boldsymbol{x}_J)| \Big(\bigotimes_{l\in L} \mu_{\widehat{F}_n^{(l)}} \otimes \bigotimes_{j\in J\setminus L} \mu_{F_0^{(j)}} \Big)(d((x_l)_{l\in L}, (x_j)_{j\in J\setminus L})) \\
&\leq \sum_{\substack{L\subseteq J,\\ |J\setminus L| \text{ even/odd}}} \int_{\mathbb{R}^d} |h_n(x_1, \ldots, x_d)| \Big(\bigotimes_{l\in L} \mu_{\widehat{F}_n^{(l)}} \otimes \bigotimes_{j\in L^c} \mu_{F_0^{(j)}} \Big)(d((x_l)_{l\in L}, (x_j)_{j\in L^c}))
\end{aligned}$$

$\mathbb{P}$-a.s. by Fubini's theorem. Since all the other remaining conditions of the integration by parts formula are satisfied by assumptions (a)-(d), this finishes the proof. □

Especially in the case where $\widehat{F}_n$ is the empirical distribution function condition (d) in Lemma 3.2.3 is essentially easier.

Remark 3.2.4 If for every $j \in \{1, \ldots, d\}$, $n \in \mathbb{N}$ and for $\mathbb{P}$-almost every ω there exist real numbers $x_{j,\ell}(\omega, n)$, $x_{j,u}(\omega, n)$ such that $\widehat{F}_n^{(j)}(\omega, x) - F_0^{(j)}(x) = -F_0^{(j)}(x)$ for all $x \leq x_{j,\ell}(\omega, n)$ and $\widehat{F}_n^{(j)}(\omega, x) - F_0^{(j)}(x) = 1 - F_0^{(j)}(x)$ for all $x \geq x_{j,u}(\omega, n)$, then assumption (d) in Lemma 3.2.3 boils down to the assumption that $\|h_{n;J,F_0}\|_\infty < \infty$ and $\sup_{(x_i)_{i\in J\setminus L}\in\mathbb{R}^{|J\setminus L|}} \int_{\mathbb{R}^L} \mu^{\pm}_{(h_{n;J,F_0})^{((x_i)_{i\in J});L}}(d((x_i)_{i\in L})) < \infty$ for all $n \in \mathbb{N}$ and all nonempty subsets $L, J \subseteq \{1, \ldots, d\}$ with $L \subsetneq J$. $\diamond$

Proof of Remark 3.2.4 On the one hand, we have for every $n \in \mathbb{N}$, for $\mathbb{P}$-almost every ω and for $b_j \geq x_{j,u}(\omega, n)$ and $a_j \leq x_{j,\ell}(\omega, n)$ for each $j = 1, \ldots, d$ that

$$\begin{aligned}
&\Big| \sum_{K\subseteq J} (-1)^{|J|-|K|} \prod_{j\in J} \Big(\widehat{F}_n^{(j)}\big(b_j^{\boldsymbol{a}_J;K}\big) - F_0^{(j)}\big(b_j^{\boldsymbol{a}_J;K}\big)\Big)\, h_{n;J,F_0}\big(\boldsymbol{b}_J^{\boldsymbol{a}_J;K}\big)\Big| \\
&\quad \leq \sum_{K\subseteq J} \|h_{n;J,F_0}\|_\infty \prod_{k\in K} \big(1 - F_0^{(k)}(b_k)\big) \prod_{j\in J\setminus K} \big(-F_0^{(j)}(a_j)\big)
\end{aligned}$$

holds. The latter bound converges to zero as $\{-a_j,\, b_j\}_{j\in J} \to +\infty$ under the assumption that $\|h_{n;J,F_0}\|_\infty < \infty$ for every $n \in \mathbb{N}$ and every nonempty subset $J \subseteq \{1, \ldots, d\}$.

On the other hand, for every $n \in \mathbb{N}$, for $\mathbb{P}$-almost every ω and for $b_j \geq x_{j,u}(\omega, n)$ and $a_j \leq x_{j,\ell}(\omega, n)$, $j = 1, \ldots, d$, we observe

$$\begin{aligned}
&\Big| \sum_{\substack{L,K\subsetneq J \text{ disjoint},\\ L\neq\emptyset}} (-1)^{|J|-|K|} \prod_{k\in K} \big(\widehat{F}_n^{(k)}(b_k) - F_0^{(k)}(b_k)\big) \prod_{j\in J\setminus(L\cup K)} \big(\widehat{F}_n^{(j)}(a_j) - F_0^{(j)}(a_j)\big) \\
&\qquad \cdot \int_{(\boldsymbol{a}_L,\boldsymbol{b}_L]} \prod_{j\in L} \big(\widehat{F}_n^{(j)}(x_j-) - F_0^{(j)}(x_j-)\big)\, \mu_{h_{n;J,F_0}}^{(\boldsymbol{b}_J^{\boldsymbol{a}_J;K});L}(d((x_j)_{j\in L}))\Big| \\
&\leq \sum_{\substack{L,K\subsetneq J \text{ disjoint},\\ L\neq\emptyset}} \prod_{k\in K} \big(1 - F_0^{(k)}(b_k)\big) \prod_{j\in J\setminus(L\cup K)} \big(-F_0^{(j)}(a_j)\big) \Big| \int_{\mathbb{R}^L} \mu_{h_{n;J,F_0}}^{(\boldsymbol{b}_J^{\boldsymbol{a}_J;K});L}(d((x_j)_{j\in L}))\Big| \\
&\leq \sum_{\substack{L,K\subsetneq J \text{ disjoint},\\ L\neq\emptyset}} \prod_{k\in K} \big(1 - F_0^{(k)}(b_k)\big) \prod_{j\in J\setminus(L\cup K)} \big(-F_0^{(j)}(a_j)\big) \\
&\qquad \cdot \sup_{(x_j)_{j\in J\setminus L}\in\mathbb{R}^{|J\setminus L|}} \int_{\mathbb{R}^L} \big(\mu^+_{(h_{n;J,F_0})^{((x_j)_{j\in J});L}} + \mu^-_{(h_{n;J,F_0})^{((x_j)_{j\in J});L}}\big)(d((x_j)_{j\in L})).
\end{aligned}$$

We note that $L \subsetneq J$ such that $K \neq \emptyset$ or $(L \cup K)^{\mathsf{c}} \neq \emptyset$. Hence, the latter bound converges to zero as $\{-a_j,\, b_j\}_{j\in J} \to +\infty$, if the latter supremum over integrals is finite for all $n \in \mathbb{N}$ and all nonempty subsets $L, J \subseteq \{1, \ldots, d\}$ with $L \subsetneq J$. $\square$

Let $\phi : \mathbb{R} \to [1, \infty)$ be any weight function, i.e. any continuous function being non-increasing on $(-\infty, 0]$ and non-decreasing on $[0, \infty)$. Subsequently, we denote by $\boldsymbol{D}_\phi$

the space of all bounded càdlàg functions on $\mathbb{R}$ satisfying $\|f\|_\phi := \|f\phi\|_\infty < \infty$ and $\lim_{|x|\to\infty} |f(x)| = 0$. We equip $\boldsymbol{D}_\phi$ with the weighted sup-metric $d_\phi(f,g) := \|f-g\|_\phi$, and we refer to $\mathcal{B}_\phi^\circ$ as the σ-algebra on $\boldsymbol{D}_\phi$ generated by the open balls with respect to d_ϕ.

If $\widehat{F}_n^{(j)}$ is the empirical distribution function and additionally $(\widehat{F}_n^{(j)} - F_0^{(j)}) \in \boldsymbol{D}_\phi$ $\mathbb{P}$-a.s. for every $j \in \{1,\dots,d\}$ and $n \in \mathbb{N}$, then we get weaker conditions on the function $h_{n;J,F_0}$ to ensure the validity of assumption (d) of Lemma 3.2.3 in comparison to Remark 3.2.4 by using the fact that $\widehat{F}_n^{(j)}(x) - F_0^{(j)}(x) \le 1/\phi(x)\|\widehat{F}_n^{(j)} - F_0^{(j)}\|_\phi$ for all $x \in \mathbb{R}$.

More generally, when $\widehat{F}_n^{(j)}$ is not necessarily the empirical distribution function but $(\widehat{F}_n^{(j)} - F_0^{(j)}) \in \boldsymbol{D}_\phi$ $\mathbb{P}$-a.s. for every $j \in \{1,\dots,d\}$ and $n \in \mathbb{N}$, the following remark holds.

Remark 3.2.5 Let $(\widehat{F}_n^{(j)} - F_0^{(j)}) \in \boldsymbol{D}_\phi$ $\mathbb{P}$-a.s. for every $j \in \{1,\dots,d\}$ and $n \in \mathbb{N}$. If for any nonempty subset $J \subseteq \{1,\dots,d\}$ and $n \in \mathbb{N}$

(d') $\lim_{\{|x_j|\}_{j\in J}\to\infty} \prod_{j\in J} 1/\phi(x_j)\, h_{n;J,F_0}((x_j)_{j\in J}) = 0$,

(d") $\lim_{\{|x_k|\}_{k\in J\setminus L}\to\infty} \prod_{k\in J\setminus L} 1/\phi(x_k) \int_{\mathbb{R}^L} \prod_{j\in L} 1/\phi(y_j)\, \mu^{\pm}_{(h_{n;J,F_0})^{((x_k)_{k\in J});L}}(d((y_j)_{j\in L})) = 0$ for all nonempty subsets $L \subsetneq J$,

then condition (d) of Lemma 3.2.3 holds. ◇

3.2.2 Examples

The following example for the set-up of a one-sample V-statistic of degree 2 is already discussed as Example 3.11 in [11].

Example 3.2.6 (Variance) If $h_2(x_1,x_2) = \frac{1}{2}(x_1 - x_2)^2$ and F has a finite second moment, then $\mathcal{V}_{h_2}(F,F)$ equals the variance of a random variable X with distribution function F. We assume that $\widehat{F}_n$ is an estimator of F such that $(\widehat{F}_n, \widehat{F}_n)$ is (ω-wise) an element of $\mathbf{F}_{h_2}$ for every $n \in \mathbb{N}$, Assumption 3.2.1 is fulfilled, and $d_\phi(\widehat{F}_n, F)$ is $\mathbb{P}$-a.s. finite for all $n \in \mathbb{N}$ and for some weight function ϕ satisfying $\int |x|/\phi(x)\,dx < \infty$. Then

$$\mu_{h_2}(d(x_1,x_2)) = -dx_1dx_2 \qquad \text{and} \qquad \mu_{h_{2\{i\},F}}(dx) = (x - \mathbb{E}[X])\,dx, \quad i = 1,2,$$

and the assumptions of Lemma 3.2.3 hold true. ◇

Further examples for set-ups of one-sample V-statistics of degree 2, under which representation (3.4) holds, can be found in [11]; for instance, Gini's mean difference (Example 3.10 and Example 3.12 in [11]), Cramér-von Mises Goodness-of-fit test statistics (Example 3.13 in [11]) and Arcones-Giné test statistics for symmetry (Example 3.14 in [11]).

The following example is a generalization of Example 3.2.6 to one-sample V-statistics of degree m.

Example 3.2.7 (Central moments of any order) If F has a finite moment of order $m \geq 2$ and if

$$h_m(x_1, \ldots, x_m) = \frac{1}{m!} \sum_{i_1=1}^{m} \cdots \sum_{i_m=1}^{m} x_{i_1}^m - \binom{m}{1} x_{i_1}^{m-1} x_{i_2} + \binom{m}{2} x_{i_1}^{m-2} x_{i_2} x_{i_3} - \cdots$$
$$+(-1)^m \Big(\binom{m}{m-1} - 1 \Big) x_{i_1} \cdots x_{i_m}, \tag{3.8}$$

then

$$\mathcal{V}_{h_m}(F, \ldots, F) = \mathbb{E}[(X - \mathbb{E}[X])^m]$$

for some random variable X with distribution function F, see Example 1.1.4 in [48]. For $m = 3$, for instance, the function h_3 has the following form $h_3(x_1, x_2, x_3) = \frac{1}{3}(x_1^3 + x_2^3 + x_3^3) - \frac{1}{2}(x_1^2 x_2 + x_1 x_2^2 + x_1^2 x_3 + x_1 x_3^2 + x_2^2 x_3 + x_2 x_3^2) + 2x_1 x_2 x_3 = \frac{1}{6}(2x_1 - x_2 - x_3)(-x_1 + 2x_2 - x_3)(-x_1 - x_2 + 2x_3)$ with corresponding measures $\mu_{h_3}(d(x_1, x_2, x_3)) = 2\, dx_1 dx_2 dx_3$ and

$$\mu_{h_{3\{i\},F}}(dx) = \big((x - \mathbb{E}[X])^2 - \mathbb{V}\mathrm{ar}[X]\big)\, dx, \qquad i = 1, 2, 3. \tag{3.9}$$

The assumptions of Lemma 3.2.3 hold true, if $\widehat{F}_n$ is an estimator of F such that $(\widehat{F}_n, \ldots, \widehat{F}_n)$ is (ω-wise) an element of $\mathbf{F}_{h_m}$ for every $n \in \mathbb{N}$, Assumption 3.2.1 is fulfilled, and $d_\phi(\widehat{F}_n, F)$ is $\mathbb{P}$-a.s. finite for all $n \in \mathbb{N}$ and for some weight function ϕ satisfying $\int_{\mathbb{R}} |x|^m / \phi(x)\, dx < \infty$. $\diamond$

Proof of Example 3.2.7 We now verify in two steps that the assumptions of Lemma 3.2.3 are fulfilled for all $m \geq 2$ and that $\mu_{h_3}(d(x_1, x_2, x_3)) = 2\, dx_1 dx_2 dx_3$ as well as (3.9) holds true.

Step 1. Let $J \subseteq \{1, \ldots, m\}$ be any nonempty set. Since Assumption 3.2.1 is fulfilled by assumption we only need to show (b)–(d).

(b) The function $h_{mJ,F}$ is given by a polynomial of degree m in the variables $x_j, j \in J$. Hence, $h_{mJ,F}$ is $|J|$ times continuously differentiable, which implies that the latter function is right continuous and locally of bounded $|J|$-fold variation by part (iv) of Proposition 2.4.4. Analogously, $(h_{mJ,F})^{\boldsymbol{c}_J;K}$ is $|K|$ times continuously differentiable and therefore locally of bounded $|K|$-fold variation for every nonempty subset $K \subseteq J$ and every fixed $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$, see part (iv) of Proposition 2.4.4.

(c) We first note that for any function $g : \mathbb{R}^k \to \mathbb{R}$ that is k times continuously differentiable we have that $g_{\widetilde{\mathbf{c}},+} : \mathbb{R}^k \to \mathbb{R}$ and $g_{\widetilde{\mathbf{c}},-} : \mathbb{R}^k \to \mathbb{R}$ defined by

$$g_{\widetilde{\mathbf{c}},+}(\boldsymbol{x}) \tag{3.10}$$
$$:= \int_{\widetilde{c}_k}^{x_k} \cdots \int_{\widetilde{c}_1}^{x_1} \Big(\frac{\partial^k g}{\partial y_k \cdots \partial y_1}(y_1, \ldots, y_k) \Big)^+ dy_1 \cdots dy_k - \frac{1}{2} \sum_{J \subsetneq \{1,\ldots,k\}} (-1)^{k-|J|} g(\boldsymbol{x}^{\widetilde{\mathbf{c}};J})$$

and

$$g_{\widetilde{\mathbf{c}},-}(\boldsymbol{x}) \tag{3.11}$$

$$:= \int_{\widetilde{c}_k}^{x_k} \cdots \int_{\widetilde{c}_1}^{x_1} \Big(\frac{\partial^k g}{\partial y_k \cdots \partial y_1}(y_1, \ldots, y_k)\Big)^- \, dy_1 \cdots dy_k + \frac{1}{2} \sum_{J \subsetneq \{1,\ldots,k\}} (-1)^{k-|J|} g(\boldsymbol{x}^{\widetilde{\boldsymbol{c}};J})$$

for $\boldsymbol{x} = (x_1, \ldots, x_k) \in \mathbb{R}^k$ and some $\widetilde{\boldsymbol{c}} = (\widetilde{c}_1, \ldots, \widetilde{c}_k) \in \mathbb{R}^k$ are $|K|$-fold monotonically increasing and right continuous functions that satisfy $g = g_{\widetilde{\boldsymbol{c}},+} - g_{\widetilde{\boldsymbol{c}},-}$. Indeed, for any $\widetilde{\boldsymbol{c}} \in \mathbb{R}^k$ we observe

$$\begin{aligned} g(\boldsymbol{x}) &= \Delta_{(\widetilde{c}_1,\ldots,\widetilde{c}_k)}^{(x_1,\ldots,x_k)} g - \big(\Delta_{(\widetilde{c}_1,\ldots,\widetilde{c}_k)}^{(x_1,\ldots,x_k)} g - g(\boldsymbol{x})\big) \\ &= \int_{\widetilde{c}_k}^{x_k} \cdots \int_{\widetilde{c}_1}^{x_1} \frac{\partial^k g}{\partial y_k \cdots \partial y_1}(y_1, \ldots, y_k) \, dy_1 \cdots dy_k - \sum_{J \subsetneq \{1,\ldots,k\}} (-1)^{k-|J|} g(\boldsymbol{x}^{\widetilde{\boldsymbol{c}};J}) \\ &= g_{\widetilde{\boldsymbol{c}},+}(\boldsymbol{x}) - g_{\widetilde{\boldsymbol{c}},-}(\boldsymbol{x}), \end{aligned}$$

where we used (2.12) and (2.1) in the second step. The k-fold monotonicity of $g_{\widetilde{\boldsymbol{c}},+}$ and $g_{\widetilde{\boldsymbol{c}},-}$ follows from part (vi) of Proposition 2.2.7, if we can show that $g_{\widetilde{\boldsymbol{c}},+}$ and $g_{\widetilde{\boldsymbol{c}},-}$ are k times continuously differentiable with $\frac{\partial^k}{\partial x_k \cdots \partial x_1} g_{\widetilde{\boldsymbol{c}},\pm}(\boldsymbol{x}) \geq 0$. In that case, $g_{\widetilde{\boldsymbol{c}},+}$ and $g_{\widetilde{\boldsymbol{c}},-}$ are additionally right continuous as differentiable functions. Now, g is k times continuously differentiable by assumption and also $\int_{\widetilde{c}_k}^{x_k} \cdots \int_{\widetilde{c}_1}^{x_1} \big(\frac{\partial^k}{\partial y_k \cdots \partial y_1} g(y_1, \ldots, y_k)\big)^{\pm} dy_1 \cdots dy_k$ is k times differentiable with continuous derivative $(\frac{\partial^k}{\partial x_k \cdots \partial x_1} g(x_1, \ldots, x_k))^{\pm}$ so that the functions $g_{\widetilde{\boldsymbol{c}},\pm}$ are indeed k times continuously differentiable. We note that the sum $\sum_{J \subsetneq \{1,\ldots,k\}} (-1)^{k-|J|} g(\boldsymbol{x}^{\widetilde{\boldsymbol{c}};J})$ has no summand depending on all $x_1, \ldots, x_k$. As a consequence,

$$\frac{\partial^k g_{\widetilde{\boldsymbol{c}},\pm}}{\partial x_k \cdots \partial x_1}(\boldsymbol{x}) = \Big(\frac{\partial^k g}{\partial x_k \cdots \partial x_1}(\boldsymbol{x})\Big)^{\pm}, \tag{3.12}$$

which is nonnegative by definition. With $g_{\widetilde{\boldsymbol{c}},+}$ and $g_{\widetilde{\boldsymbol{c}},-}$ we have thus found a suitable decomposition of g into two right continuous and k-fold monotonically increasing functions.

Now, let $K \subseteq J$ be any nonempty subset, $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$ and $\widetilde{\boldsymbol{c}}_K \in \mathbb{R}^{|K|}$. Then $(h_{mJ,F})^{\boldsymbol{c}_J;K}$ is given by a polynomial of degree m in each x_k, $k \in K$, and $|K|$ times continuously differentiable. In particular, the functions $((h_{mJ,F})^{\boldsymbol{c}_J;K})_{\widetilde{\boldsymbol{c}}_K,\pm}$ defined as in (3.10) and (3.11) are $|K|$-fold monotonically increasing and right continuous and satisfy $(h_{mJ,F})^{\boldsymbol{c}_J;K} = ((h_{mJ,F})^{\boldsymbol{c}_J;K})_{\widetilde{\boldsymbol{c}}_K,+} - ((h_{mJ,F})^{\boldsymbol{c}_J;K})_{\widetilde{\boldsymbol{c}}_K,-}$ as we have seen above.

The $(\mathcal{B}(\mathbb{R}^{|K|}), \mathcal{B}(\mathbb{R}))$-measurability of the functions $((h_{mJ,F})^{\boldsymbol{c}_J;K})_{\widetilde{\boldsymbol{c}}_K,\pm}$ is thus a direct consequence of the fact that the latter functions are $|K|$ times continuously differentiable (see above).

Moreover, the functions $((h_{mJ,F})^{\boldsymbol{c}_J;K})_{\widetilde{\boldsymbol{c}}_K,\pm}$ are bounded on every finite $(\boldsymbol{a}_K, \boldsymbol{b}_K] \subsetneq \mathbb{R}^{|K|}$ due to the continuous differentiability. As a result, we have that

$$\begin{aligned} &\int_{\boldsymbol{I}_K^{\boldsymbol{a},\boldsymbol{b}}} \big|((h_{mJ,F})^{\boldsymbol{c}_J;K})_{\widetilde{\boldsymbol{c}}_K,\pm}(\boldsymbol{x}_K)\big| \Big(\mu_{\widehat{F}_n}^{\otimes|L\cap K|} \otimes \mu_F^{\otimes|(J\setminus L)\cap K|}\Big)(d((x_l)_{l\in L\cap K}, (x_j)_{j\in(J\setminus L)\cap K})) \\ &\leq \; C_{\boldsymbol{a}_K,\boldsymbol{b}_K,\boldsymbol{c}_J} \prod_{l\in L\cap K} \mu_{\widehat{F}_n}((a_l, b_l]) \prod_{j\in(J\setminus L)\cap K} \mu_F((a_j, b_j]) \quad \leq \quad C_{\boldsymbol{a}_K,\boldsymbol{b}_K,\boldsymbol{c}_J} \end{aligned}$$

holds for all subsets $K, L \subseteq J$ with $K \neq \emptyset$, $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$ and every finite interval $\boldsymbol{I}_K^{\boldsymbol{a},\boldsymbol{b}} \subsetneq \mathbb{R}^{|K|}$, as defined in assumption (c) of Lemma 3.2.3, where $C_{\boldsymbol{a}_K,\boldsymbol{b}_K,\boldsymbol{c}_J}$ is some finite constant satisfying $\sup_{\boldsymbol{y}\in(\boldsymbol{a}_K,\boldsymbol{b}_K]} |((h_{m\,J,F})^{\boldsymbol{c}_J;K})_{\widetilde{\boldsymbol{c}}_K,\pm}(\boldsymbol{y})| \leq C_{\boldsymbol{a}_K,\boldsymbol{b}_K,\boldsymbol{c}_J}$. Since the latter bound is finite for every $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$ and every finite interval $(\boldsymbol{a}_K, \boldsymbol{b}_K] \subsetneq \mathbb{R}^{|K|}$, the integral on the left-hand side exists for any subsets $K, L \subseteq J$ with $K \neq \emptyset$, for every $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$ and $\boldsymbol{I}_K^{\boldsymbol{a},\boldsymbol{b}} \subsetneq \mathbb{R}^{|K|}$.

(d) To prove that the first limit in assumption (d) of Lemma 3.2.3 exists and equals zero $\mathbb{P}$-a.s. we have to show, according to Remark 3.2.5, that $\lim_{\{|x_j|\}_{j\in J}\to\infty} \prod_{j\in J} 1/\phi(x_j)\, h_{m\,J,F}((x_j)_{j\in J})$ equals zero. We note that $\int_{\mathbb{R}} |x|^m/\phi(x)\,dx < \infty$ by our assumptions, which implies that $|x|^m/\phi(x)$ and thus each polynomial of degree at most m in x divided by $\phi(x)$ converges to 0 as $|x| \to \infty$. Since $h_{m\,J,F}((x_j)_{j\in J})$ is a polynomial of degree m in the variables x_j, $j \in J$, this proves the first assertion.

For the proof that the second limit exists and equals zero $\mathbb{P}$-a.s. it suffices to show that $\prod_{k\in J\setminus L} 1/\phi(x_k) \int_{\mathbb{R}^L} \prod_{j\in L} 1/\phi(y_j)\, \mu^{\pm}_{(h_{m\,J,F})^{((x_k)_{k\in J});L}}(d((y_j)_{j\in L})$ converges to zero as $\{|x_k|\}_{k\in J\setminus L} \to \infty$ for any nonempty subset $L \subsetneq J$ by Remark 3.2.5.

Recall that if $g : \mathbb{R}^k \to \mathbb{R}$ is k-times continuously differentiable with $g_{\widetilde{\boldsymbol{c}},\pm}$ as defined in (3.10)-(3.11) for some $\widetilde{\boldsymbol{c}} \in \mathbb{R}^k$, the measures μ_g^+ and μ_g^- are generated by $g_{\widetilde{\boldsymbol{c}},+}$ and $g_{\widetilde{\boldsymbol{c}},-}$, respectively. Since $g_{\widetilde{\boldsymbol{c}},+}$ and $g_{\widetilde{\boldsymbol{c}},-}$ are k times continuously differentiable, as we have seen in the proof of (c), we obtain in view of (3.12)

$$\mu_g^{\pm}(d\boldsymbol{x}) \;=\; \frac{\partial^k g_{\widetilde{\boldsymbol{c}},\pm}}{\partial x_k \cdots \partial x_1}(\boldsymbol{x})\,d\boldsymbol{x} \;=\; \Big(\frac{\partial^k g}{\partial x_k \cdots \partial x_1}(\boldsymbol{x})\Big)^{\pm} d\boldsymbol{x} \tag{3.13}$$

for $\boldsymbol{x} \in \mathbb{R}^k$.

Now, for any nonempty subset $L \subsetneq J$ the function $(h_{m\,J,F})^{((x_k)_{k\in J});L}$ is given by the polynomial $(h_{m\,J,F})^{((x_k)_{k\in J});L}((y_j)_{j\in L})$, which is of degree m in each y_j, $j \in L$, and in each x_k, $k \in J \setminus L$. We note that the polynomial has the property that the derivative $(\frac{\partial^{|L|}}{\partial((y_j)_{j\in L})}(h_{m\,J,F})^{(x_k)_{k\in J};L})((y_j)_{j\in L})$ is a polynomial of degree $m - |L|$ in y_k, $k \in L$, and in x_k, $k \in J \setminus L$ so that $((\frac{\partial^{|L|}}{\partial((y_j)_{j\in L})}(h_{m\,J,F})^{(x_k)_{k\in J};L})((y_j)_{j\in L}))^{\pm}$ is piecewise composed of polynomials of degree at most $m-1$ in y_k, $k \in L$, and x_k, $k \in J \setminus L$. Along the lines of the proof of the first limit, the assertion is thus a direct consequence of $\int_{\mathbb{R}} |x|^m/\phi(x)\,dx < \infty$ in view of (3.13). This finishes the proof of assumption (d). So all assumptions of Lemma 3.2.3 are fulfilled.

Step 2. To prove $\mu_{h_3}(d(x_1, x_2, x_3)) = 2\,dx_1dx_2dx_3$, we note that the following statement can be derived from part (iii) of Proposition 2.2.7: Any two right continuous functions $f_1, f_2 : \mathbb{R}^d \to \mathbb{R}$ that are locally of bounded d-fold variation generate the same measure on $\mathbb{R}^d$ if and only if $f_1(x_1, \ldots, x_d) = f_2(x_1, \ldots, x_d) + \sum_{\emptyset\neq J\subsetneq\{1,\ldots,d\}} g_J((x_j)_{j\in J})$ for some functions $g_J : \mathbb{R}^J \to \mathbb{R}$. In our specific case, h_3 and $2x_1x_2x_3$ generate the same measure so that up to the coefficient 2 the measure μ_{h_3} coincides with the Lebesgue measure.

For the proof of (3.9), we observe that

$$\begin{aligned} h_{3\{1\},F}(x) &= \int_{\mathbb{R}^2} h_3(x,x_2,x_3)\,(\mu_F\otimes\mu_F)(d(x_2,x_3)) \\ &= \frac{1}{3}x^3+\frac{2}{3}\mathbb{E}[X^3]-x^2\mathbb{E}[X]-x\big(\mathbb{E}[X^2]-2\mathbb{E}[X]^2\big)-\mathbb{E}[X]\,\mathbb{E}[X^2], \end{aligned}$$

which yields $\frac{\partial}{\partial x}h_{3\{1\},F}(x) = (x^2-2x\,\mathbb{E}[X]+2\mathbb{E}[X]^2-\mathbb{E}[X^2])$. Hence, (3.9) holds for $i=1$ in view of (3.12). Due to the symmetry of h_3, assertion (3.9) is indeed valid for each $i=1,2,3$. □

As we have seen in Subsection 3.2.1, the kernel $h_{n;J,F_0}$ has to be locally of $|J|$-fold variation and $(h_{n;J,F_0})^{\boldsymbol{c}_J;K}$ has to be locally of bounded $|K|$-fold variation for any nonempty subsets $K,J\subseteq\{1,\dots,d\}$ with $K\subsetneq J$ and any $\boldsymbol{c}_J\in\mathbb{R}^{|J|}$. If the kernel has too many discontinuities, this might be not fulfilled. Beutner and Zähle showed in Remark 1.1 in [12] that the Wilcoxon and Mann-Whitney two-sample test statistics and an asymptotically equivalent (one-sample) statistic to the Wilcoxon signed rank test statistics do not fulfill the required assumptions for the generalized von Mises decomposition.

3.3 Weak (central) limit theorems

Let $d_1,\dots,d_k\in\mathbb{N}$, and let $h_{n,1}:\mathbb{R}^{d_1}\to\mathbb{R},\dots,h_{n,k}:\mathbb{R}^{d_k}\to\mathbb{R}$ be Borel measurable functions for every $n\in\mathbb{N}_0$. As before, for every $i=1,\dots,k$ we denote by $\mathbf{F}_{h_{n,i}}$ the set of all d_i-tuples $(F^{(i1)},\dots,F^{(id_i)})$ of distribution functions on $\mathbb{R}$ for which $\mathcal{V}_{h_{n,i}}(F^{(i1)},\dots,F^{(id_i)})$ exists. For each $i=1,\dots,k$ let $(F_0^{(i1)},\dots,F_0^{(id_i)})\in\mathbf{F}_{h_{0,i}}$ be fixed, and let $\widehat{F}_n^{(ij)}$ be an estimator of $F_0^{(ij)}$ for every $j=1,\dots,d_i$ and $n\in\mathbb{N}$ such that $(\widehat{F}_n^{(i1)},\dots,\widehat{F}_n^{(id_i)})\in\mathbf{F}_{h_{n,i}}$ (ω-wise) for every $n\in\mathbb{N}$.

Throughout the entire section, we use $\|\boldsymbol{v}\|$ to denote the Euclidean norm of some vector $\boldsymbol{v}\in\mathbb{R}^k$.

3.3.1 Weak limit theorem for one-sample V-statistics

In this subsection we focus on one-sample V-statistics, i.e. $F_0^{(ij)}=F_0$ and $\widehat{F}_n^{(ij)}=\widehat{F}_n$ for each $i=1,\dots,k$ and $j=1,\dots,d_i$. Let $(\mathbf{V},d_{\mathbf{V}})$ be any metric space. We equip $\mathbf{V}$ with the σ-algebra $\mathcal{B}^\circ$ generated by the open balls with respect to the metric $d_{\mathbf{V}}$.

Theorem 3.3.1 *Let (a_n) be a sequence in $(0,\infty)$ with $a_n\to\infty$ as $n\to\infty$. Suppose that*

(a) *for each $i=1,\dots,k$ the assumptions of Lemma 3.2.3 with h_n, $(F_0^{(1)},\dots,F_0^{(d)})$ and $(\widehat{F}_n^{(1)},\dots,\widehat{F}_n^{(d)})$ replaced by $h_{n,i}$, $(F_0,\dots,F_0)$ and $(\widehat{F}_n,\dots,\widehat{F}_n)$, respectively, are fulfilled,*

(b) *the process* $a_n(\widehat{F}_n - F_0)$ *is a* $(\mathbf{V}, \mathcal{B}^\circ)$*-valued random variable on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *for all* $n \in \mathbb{N}$*, and there exists a* $(\mathbf{V}, \mathcal{B}^\circ)$*-valued random variable* B *on* $(\Omega', \mathcal{F}', \mathbb{P}')$ *such that* $B(\Omega') \subseteq S$ *for some separable* $S \in \mathcal{B}^\circ$ *and*

$$a_n(\widehat{F}_n - F_0) \rightsquigarrow^\circ B \qquad \text{in } (\mathbf{V}, \mathcal{B}^\circ, d_{\mathbf{V}}),$$

(c) *for each* $n \in \mathbb{N}$ *the map* $\widetilde{\Phi}^n : \mathbf{V} \to \mathbb{R}^k$ *defined by*

$$\widetilde{\Phi}^n(f) := \begin{bmatrix} \sum_{\emptyset \neq J \subseteq \{1,\ldots,d_1\}} (-1)^{|J|}\, a_n^{1-|J|} \int_{\mathbb{R}^J} \prod_{j \in J} f(x_j-)\, \mu_{h_{n,1;J,F_0}}(d((x_j)_{j\in J})) \\ \vdots \\ \sum_{\emptyset \neq J \subseteq \{1,\ldots,d_k\}} (-1)^{|J|}\, a_n^{1-|J|} \int_{\mathbb{R}^J} \prod_{j \in J} f(x_j-)\, \mu_{h_{n,k;J,F_0}}(d((x_j)_{j\in J})) \end{bmatrix}$$

is well-defined with $\widetilde{\Phi}^n\big(a_n(\widehat{F}_n - F_0)\big) : \Omega \to \mathbb{R}^k$ *being* $(\mathcal{F}, \mathcal{B}(\mathbb{R}^k))$*-measurable, and the map* $\widetilde{\Phi}^0 : S \to \mathbb{R}^k$ *(with* S *as in (b)) defined by*

$$\widetilde{\Phi}^0(f) := \begin{bmatrix} -\sum_{j=1}^{d_1} \int_{\mathbb{R}} f(x_j-)\, \mu_{h_{0,1;\{j\},F_0}}(dx_j) \\ \vdots \\ -\sum_{j=1}^{d_k} \int_{\mathbb{R}} f(x_j-)\, \mu_{h_{0,k;\{j\},F_0}}(dx_j) \end{bmatrix}$$

is well-defined and $(\mathcal{B}^\circ \cap S, \mathcal{B}(\mathbb{R}^k))$*-measurable,*

(d) *for any sequence* $(f_n)_n \subseteq \mathbf{V}$ *we have* $\big\|\widetilde{\Phi}^n(f_n) - \widetilde{\Phi}^0(f_0)\big\| \to 0$ *when* $d_{\mathbf{V}}(f_n, f_0) \to 0$.

Then we have

$$a_n\left(\begin{bmatrix} \mathcal{V}_{h_{n,1}}(\widehat{F}_n, \ldots, \widehat{F}_n) \\ \vdots \\ \mathcal{V}_{h_{n,k}}(\widehat{F}_n, \ldots, \widehat{F}_n) \end{bmatrix} - \begin{bmatrix} \mathcal{V}_{h_{n,1}}(F_0, \ldots, F_0) \\ \vdots \\ \mathcal{V}_{h_{n,k}}(F_0, \ldots, F_0) \end{bmatrix} \right)$$
$$\rightsquigarrow \begin{bmatrix} -\sum_{j=1}^{d_1} \int_{\mathbb{R}} B(x_j-) \mu_{h_{0,1;\{j\},F_0}}(dx_j) \\ \vdots \\ -\sum_{j=1}^{d_k} \int_{\mathbb{R}} B(x_j-) \mu_{h_{0,k;\{j\},F_0}}(dx_j) \end{bmatrix} \qquad \text{in } (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k)).$$

Recall that $\rightsquigarrow^\circ$ denotes convergence in distribution with respect to the open-ball σ-algebra $\mathcal{B}^\circ$. In separable metric spaces such as $\mathbb{R}^k$ the open-ball σ-algebra coincides with the Borel-σ-algebra. In this case, we simply write $\rightsquigarrow$ for the convergence in distribution.

Proof of Theorem 3.3.1 By Lemma 3.2.3 and assumption (a) we obtain

$$a_n\left(\begin{bmatrix} \mathcal{V}_{h_{n,1}}(\widehat{F}_n, \ldots, \widehat{F}_n) \\ \vdots \\ \mathcal{V}_{h_{n,k}}(\widehat{F}_n, \ldots, \widehat{F}_n) \end{bmatrix} - \begin{bmatrix} \mathcal{V}_{h_{n,1}}(F_0, \ldots, F_0) \\ \vdots \\ \mathcal{V}_{h_{n,k}}(F_0, \ldots, F_0) \end{bmatrix} \right)$$

$$\begin{aligned} &= a_n \begin{bmatrix} \sum_{\emptyset\neq J\subseteq\{1,\dots,d_1\}}(-1)^{|J|}\int_{\mathbb{R}^J}\prod_{j\in J}\big(\widehat{F}_n(x_j-)-F_0(x_j-)\big)\,\mu_{h_{n,1;J,F_0}}(d((x_j)_{j\in J}))\\ \vdots\\ \sum_{\emptyset\neq J\subseteq\{1,\dots,d_k\}}(-1)^{|J|}\int_{\mathbb{R}^J}\prod_{j\in J}\big(\widehat{F}_n(x_j-)-F_0(x_j-)\big)\,\mu_{h_{n,k;J,F_0}}(d((x_j)_{j\in J}))\end{bmatrix}\\ &= \widetilde{\Phi}^n\big(a_n(\widehat{F}_n-F_0)\big)\end{aligned}$$

$\mathbb{P}$-a.s. for all $n\in\mathbb{N}$ so that the claim follows from the extended continuous mapping theorem (cf. Theorem C.1 in [13]) in view of assumptions (b)–(d). □

If the metric space is given by $(\boldsymbol{D}_\phi, d_\phi)$ and $h_n \equiv h_0$ for all $n\in\mathbb{N}$, then assumptions (c) and (d) of Theorem 3.3.1 boil down to a condition on the weight function ϕ and the kernel functions $h_{0,1},\dots,h_{0,k}$:

Remark 3.3.2 (i) Let the metric space $(\mathbf{V}, d_{\mathbf{V}})$ be given by $(\boldsymbol{D}_\phi, d_\phi)$, and let $h_n\equiv h_0$ for all $n\in\mathbb{N}$. If $\int_{\mathbb{R}^J}\prod_{j\in J}1/\phi(x_j)\,\mu^{\pm}_{h_{0,i;J,F_0}}(d((x_j)_{j\in J}))<\infty$ for all nonempty subsets $J\subseteq\{1,\dots,d_i\}$ and each $i=1,\dots,k$, then assumption (d) of Theorem 3.3.1 is fulfilled.

(ii) Assumption (c) of Theorem 3.3.1 holds under the conditions of part (i) of this remark, if additionally assumption (b) is valid. ◇

Proof of Remark 3.3.2 (i) For any sequence $(f_n)_n\subseteq\boldsymbol{D}_\phi$ and any $f_0\in\boldsymbol{D}_\phi$ we observe

$$\begin{aligned} &\big\|\widetilde{\Phi}^n(f_n)-\widetilde{\Phi}^0(f_0)\big\|\\ &\le \Big\|\Big[\sum_{j=1}^{d_i}\int_{\mathbb{R}}\big|f_n(x_j-)-f_0(x_j-)\big|\,\big(\mu^+_{h_{0,i;\{j\},F_0}}+\mu^-_{h_{0,i;\{j\},F_0}}\big)(dx_j)\Big]_{i\in\{1,\dots,k\}}\Big\|\\ &\quad+\Big\|\Big[\sum_{J\subseteq\{1,\dots,d_i\},|J|\ge2}a_n^{1-|J|}\int_{\mathbb{R}^J}\prod_{j\in J}|f_n(x_j-)|\,\big(\mu^+_{h_{0,i;J,F_0}}+\mu^-_{h_{0,i;J,F_0}}\big)(d((x_j)_{j\in J}))\Big]_{i\in\{1,\dots,k\}}\Big\|\\ &=: S_1(n)+S_2(n)\end{aligned}$$

for any $n\in\mathbb{N}$, where we use $[b_i]_{i\in\{1,\dots,k\}}$ to denote the column vector $\big[b_1\ \ \dots\ \ b_k\big]'$.

For the first summand we obtain

$$S_1(n)\ \le\ \Big\|\Big[\sum_{j=1}^{d_i}\|f_n-f_0\|_\phi\int_{\mathbb{R}}1/\phi(x_j)\,\big(\mu^+_{h_{0,i;\{j\},F_0}}+\mu^-_{h_{0,i;\{j\},F_0}}\big)(dx_j)\Big]_{i\in\{1,\dots,k\}}\Big\|$$

for any $n\in\mathbb{N}$, where $\phi(x)=\phi(x-)$ holds because of the continuity of ϕ. Since for any $j\in\{1,\dots,d_i\}$ and $i\in\{1,\dots,k\}$ we have $\int_{\mathbb{R}}1/\phi(x_j)\,|\mu_{h_{0,i;\{j\},F_0}}|(dx_j)<\infty$ by our assumptions, the latter bound converges to 0 when $\|f_n-f_0\|_\phi\to0$.

For the second summand we have

$$S_2(n)\ \le\ \Big\|\Big[\sum_{J\subseteq\{1,\dots,d_i\},|J|\ge2}a_n^{1-|J|}\sum_{K\subseteq J}\prod_{j\in K}\|f_n-f_0\|_\phi\prod_{\ell\in J\setminus K}\|f_0\|_\phi$$

$$\cdot \int_{\mathbb{R}^J} \prod_{j\in J} 1/\phi(x_j)\,\big(\mu^+_{h_{0,i;J,F_0}} + \mu^-_{h_{0,i;J,F_0}}\big)(d((x_j)_{j\in J}))\Big]_{i\in\{1,\dots,k\}}\Big\|$$

$$\leq \Big\|\Big[\sum_{J\subseteq\{1,\dots,d_i\},|J|\geq 2} a_0^{1-|J|} \sum_{K\subseteq J} \|f_n - f_0\|_\phi^{|K|}\,\|f_0\|_\phi^{|J\setminus K|}$$

$$\cdot \int_{\mathbb{R}^J} \prod_{j\in J} 1/\phi(x_j)\,\big(\mu^+_{h_{0,i;J,F_0}} + \mu^-_{h_{0,i;J,F_0}}\big)(d((x_j)_{j\in J}))\Big]_{i\in\{1,\dots,k\}}\Big\| \tag{3.14}$$

for any $n \in \mathbb{N}$ being sufficiently large, where we used in the first step that

$$\begin{aligned}\prod_{j\in J} |f_n(x_j-)| \;&\leq\; \prod_{j\in J} \big\{\big|f_n(x_j-) - f_0(x_j-)\big| + |f_0(x_j-)|\big\}\\ &=\; \sum_{K\subseteq J} \Big\{\prod_{j\in K} |f_n(x_j-) - f_0(x_j-)|\Big\}\Big\{\prod_{j\in J\setminus K} |f_0(x_j-)|\Big\}\end{aligned}$$

holds and in the second step that $a_n^{1-|J|} \to 0$ as $n \to \infty$. The integral on the right-hand side of (3.14) is finite by our assumptions, and $\|f_0\|_\phi < \infty$ because $f_0 \in \boldsymbol{D}_\phi$. The right-hand side of (3.14) thus converges to 0 when $\|f_n - f_0\|_\phi \to 0$. Hence, assumption (d) of Theorem 3.3.3 holds.

(ii) The map $\widetilde{\Phi}^n$ is well-defined because for any $f \in \boldsymbol{D}_\phi$, $n \in \mathbb{N}$, $i \in \{1,\dots,k\}$ and for every nonempty subset $J \subseteq \{1,\dots,d_i\}$ we have

$$\int_{\mathbb{R}^J} \big|\prod_{j\in J} f(x_j-)\big|\,\mu^\pm_{h_{0,i;J,F_0}}(d((x_j)_{j\in J})) \;\leq\; \|f\|_\phi^{|J|} \int_{\mathbb{R}^J} \prod_{j\in J} 1/\phi(x_j)\,\mu^\pm_{h_{0,i;J,F_0}}(d((x_j)_{j\in J}))$$

and the latter bound is finite by our assumptions. Similarly, we observe for any $f \in \boldsymbol{D}_\phi$, $i \in \{1,\dots,k\}$ and $j \in \{1,\dots,d_i\}$

$$\int_{\mathbb{R}} |f(x_j-)|\,\mu^\pm_{h_{0,i;\{j\},F_0}}(dx_j) \;\leq\; \|f\|_\phi \int_{\mathbb{R}} 1/\phi(x_j)\,\mu^\pm_{h_{0,i;\{j\},F_0}}(dx_j),$$

which is finite and, therefore, yields the well-definedness of $\widetilde{\Phi}^0$.

To show for every $n \in \mathbb{N}$ the $(\mathcal{F},\mathcal{B}(\mathbb{R}^k))$- and $(\mathcal{B}^\circ_\phi \cap S,\mathcal{B}(\mathbb{R}^k))$-measurability of $\widetilde{\Phi}^n(a_n(\widehat{F}_n - F_0))$ and $\widetilde{\Phi}^0$, respectively, we note that $\mathcal{B}(\mathbb{R}^k) = \mathcal{B}(\mathbb{R})^{\otimes k}$ (see e.g. Theorem 14.8. in [47]), so that the σ-algebra $\mathcal{B}(\mathbb{R}^k)$ is generated by the coordinate projections $\pi_i : \mathbb{R}^k \to \mathbb{R}$ for $i = 1,\dots,k$. It thus suffices to show that $\pi_i(\widetilde{\Phi}^n(a_n(\widehat{F}_n - F_0)))$ and $\pi_i(\widetilde{\Phi}^0)$ are $(\mathcal{F},\mathcal{B}(\mathbb{R}))$- and $(\mathcal{B}^\circ_\phi \cap S,\mathcal{B}(\mathbb{R}))$-measurable, respectively, for all $i \in \{1,\dots,k\}$ and $n \in \mathbb{N}$. The latter measurability holds because $\pi_i(\widetilde{\Phi}^0)$ is obviously $(d_\phi,\|\cdot\|)$-continuous under the assumption that $\int_{\mathbb{R}^J}\prod_{j\in J} 1/\phi(x_j)\,\mu^\pm_{h_{0,i;J,F_0}}(d((x_j)_{j\in J})) < \infty$. For the $(\mathcal{F},\mathcal{B}(\mathbb{R}))$-measurability of $\pi_i(\widetilde{\Phi}^n(a_n(\widehat{F}_n - F_0)))$ we observe that $\omega \mapsto a_n(\widehat{F}_n(\omega,\cdot) - F_0(\cdot))$ is $(\mathcal{F},\mathcal{B}^\circ_\phi)$-measurable for every $n \in \mathbb{N}$ by assumption (b) of Theorem 3.3.1. Since $\mathcal{B}^\circ_\phi$ coincides with the σ-algebra generated by the one-dimensional coordinate projections $\pi_x : \boldsymbol{D}_\phi \to \mathbb{R}, v \mapsto v(x)$ with $x \in \mathbb{R}$ by Lemma 4.1 in [13], we obtain in particular the

$(\mathcal{F}, \mathcal{B}(\mathbb{R}))$-measurability of $\omega \mapsto a_n(\widehat{F}_n(\omega, x) - F_0(x))$ for every $x \in \mathbb{R}$, $n \in \mathbb{N}$. Therefore, $\pi_i(\widetilde{\Phi}^n(a_n(\widehat{F}_n - F_0)))$ approximated by sums is $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$-measurable for every $n \in \mathbb{N}$ as limit and composition of measurable functions. □

We have seen in Theorem 3.3.1 that we mainly need a weak convergence theorem for $a_n(\widehat{F}_n - F)$ in assumption (b) to derive the asymptotics of V-statistics by means of this extended continuous mapping approach, provided the generalized von Mises decomposition is valid. If $\widehat{F}_n$ is the empirical distribution function, weak convergence theorems for the empirical process $\sqrt{n}(\widehat{F}_n - F)$ have extensively been studied. In the metric space $(\boldsymbol{D}_\phi, \|\cdot\|_\phi)$, where weak convergence of the empirical process with respect to $\|\cdot\|_\phi$ means weak convergence of the weighted version $\sqrt{n}(\widehat{F}_n - F)\phi$ with respect to $\|\cdot\|_\infty$, a weak convergence theorem for independent identically distributed data can be found for instance in Shorack and Wellner [70, Theorem 6.2.1]. For weakly dependent data, Shao and Yu studied in [69, Theorems 2.2, 2.3 and 2.4] the asymptotics of the weighted empirical process for stationary α-, ρ-mixing and associated sequences of random variables and Arcones and Yu in [4, Theorem 2.1] the one for stationary β-mixing sequences of random variable; see Section 3.2 in [10] and Example 4.4 and Section 5.2 of [13] for details. We refer to [17] and [33] for definitions and examples of the different mixing conditions and the relations between them. For stationary sequences of strongly dependent random variables (data with long memory) a weak convergence theorem for the weighted empirical process was proven for instance in Beutner et al. [8].

In [78, Theorem 1], Wu studied the asymptotic distribution of the weighted empirical process of stationary sequences by supposing a weak dependency condition similar to assumption (A8) in Subsection 1.2.2. In Chapter 1 we investigated the asymptotic distribution of the weighted empirical process for non-stationary time series and proved a variant of Theorem 1 in [78] for locally stationary time series. Combined with the extended continuous mapping approach, this enables to determine the asymptotic distribution of *weighted* V-statistics of degree d. We will come back to this application in Section 3.4.

If we marginally adjust the maps $\widetilde{\Phi}^n$ and $\widetilde{\Phi}^0$ in assumptions (c)–(d) and in the proof of Theorem 3.3.1, then the weak limit theorem for one-sample V-statistics can be generalized to multi-sample V-statistics.

3.3.2 Weak limit theorem for multi-sample V-statistics

As before, let $(\mathbf{V}, d_{\mathbf{V}})$ be any metric space that is equipped with the σ-algebra $\mathcal{B}^\circ$ generated by the open balls with respect to the metric $d_{\mathbf{V}}$. For $d := d_1 + \cdots + d_k$ we set $\mathbf{V}^d := \mathbf{V} \times \cdots \times \mathbf{V}$ and denote by $\mathcal{B}^{\circ,d}$ the σ-algebra on $\mathbf{V}^d$ generated by the open balls with respect to the metric $d^d_{\mathbf{V}}$ defined by $d^d_{\mathbf{V}}\big((x_1, \ldots, x_d), (y_1, \ldots, y_d)\big) :=$

$\max_{i\in\{1,\dots,d\}}\{d_{\mathbf{V}}(x_i,y_i)\}$. We note that $d^d_{\mathbf{V}}\big((x_1,\dots,x_d),(y_1,\dots,y_d)\big)\to 0$ if and only if $d_{\mathbf{V}}(x_i,y_i)\to 0$ for all $i\in\{1,\dots,d\}$.

For any nonempty subset $J\subseteq\{1,\dots,d\}$, the metric space $(\mathbf{V}^{|J|},\mathcal{B}^{\circ,|J|},d^{|J|}_{\mathbf{V}})$ is defined in the same way.

Theorem 3.3.3 *Let $d:=d_1+\cdots+d_k$ and let (a_n) be a sequence in $(0,\infty)$ with $a_n\to\infty$ as $n\to\infty$. Suppose that*

(a) *for each $i=1,\dots,k$ the assumptions of Lemma 3.2.3 with h_n, $(F_0^{(1)},\dots,F_0^{(d)})$ and $(\widehat{F}_n^{(1)},\dots,\widehat{F}_n^{(d)})$ replaced by $h_{n,i}$, $(F_0^{(i1)},\dots,F_0^{(id_i)})$ and $(\widehat{F}_n^{(i1)},\dots,\widehat{F}_n^{(id_i)})$, respectively, are fulfilled,*

(b) *the process $(a_n(\widehat{F}_n^{(i1)}-F_0^{(i1)}),\dots,a_n(\widehat{F}_n^{(id_i)}-F_0^{(id_i)}))_{i\in\{1,\dots,k\}}$ is a $(\mathbf{V}^d,\mathcal{B}^{\circ,d})$-valued random variable on the probability space $(\Omega,\mathcal{F},\mathbb{P})$ for all $n\in\mathbb{N}$, and there exists a $(\mathbf{V}^d,\mathcal{B}^{\circ,d})$-valued random variable $B:=(B_{i1},\dots,B_{id_i})_{i\in\{1,\dots,k\}}$ on $(\Omega',\mathcal{F}',\mathbb{P}')$ such that $B(\Omega')\subseteq S$ for some separable $S\in\mathcal{B}^{\circ,d}$ and*

$$\big(a_n(\widehat{F}_n^{(i1)}-F_0^{(i1)}),\dots,a_n(\widehat{F}_n^{(id_i)}-F_0^{(id_i)})\big)_{i\in\{1,\dots,k\}}\rightsquigarrow^{\circ}(B_{i1},\dots,B_{id_i})_{i\in\{1,\dots,k\}}$$

in $(\mathbf{V}^d,\mathcal{B}^{\circ,d},d^d_{\mathbf{V}})$,

(c) *for each $n\in\mathbb{N}$ the map $\Phi^n:\mathbf{V}^d\to\mathbb{R}^k$ defined by*

$$\Phi^n\big((f_{ij})_{i\in\{1,\dots,k\},j\in\{1,\dots,d_i\}}\big) := \begin{bmatrix}\sum_{\emptyset\neq J\subseteq\{1,\dots,d_1\}}(-1)^{|J|}\,a_n^{1-|J|}\int_{\mathbb{R}^J}\prod_{j\in J}f_{1j}(x_j-)\,\mu_{h_{n,1;J,F_0}}(d((x_j)_{j\in J}))\\ \vdots\\ \sum_{\emptyset\neq J\subseteq\{1,\dots,d_k\}}(-1)^{|J|}\,a_n^{1-|J|}\int_{\mathbb{R}^J}\prod_{j\in J}f_{kj}(x_j-)\,\mu_{h_{n,k;J,F_0}}(d((x_j)_{j\in J}))\end{bmatrix}$$

is well-defined with $\Phi^n\big((a_n(\widehat{F}_n^{(ij)}-F_0^{(ij)}))_{i\in\{1,\dots,k\},j\in\{1,\dots,d_i\}}\big):\Omega\to\mathbb{R}^k$ being $(\mathcal{F},\mathcal{B}(\mathbb{R}^k))$-measurable, and the map $\Phi^0:S\to\mathbb{R}^k$ (with S as in (b)) defined by

$$\Phi^0\big((f_{ij})_{i\in\{1,\dots,k\},j\in\{1,\dots,d_i\}}\big) := \begin{bmatrix}-\sum_{j=1}^{d_1}\int_{\mathbb{R}}f_{1j}(x_j-)\,\mu_{h_{0,1;\{j\},F_0}}(dx_j)\\ \vdots\\ -\sum_{j=1}^{d_k}\int_{\mathbb{R}}f_{kj}(x_j-)\,\mu_{h_{0,k;\{j\},F_0}}(dx_j)\end{bmatrix}$$

is well-defined and $(\mathcal{B}^{\circ,d}\cap S,\mathcal{B}(\mathbb{R}^k))$-measurable,

(d) *for any sequence $((f_{ij}^n)_{i\in\{1,\dots,k\},j\in\{1,\dots,d_i\}})_n\subseteq\mathbf{V}^d$ we have*

$$\big\|\Phi^n\big((f^n_{ij})_{i\in\{1,\dots,k\},j\in\{1,\dots,d_i\}}\big)-\Phi^0\big((f^0_{ij})_{i\in\{1,\dots,k\},j\in\{1,\dots,d_i\}}\big)\big\|\to 0$$

when $d_{\mathbf{V}}(f^n_{ij},f^0_{ij})\to 0$ for every $i\in\{1,\dots,k\}$ and $j\in\{1,\dots,d_i\}$.

Then we have

$$a_n\left(\begin{bmatrix}\mathcal{V}_{h_{n,1}}\big(\widehat{F}_n^{(11)},\ldots,\widehat{F}_n^{(1d_1)}\big)\\ \vdots\\ \mathcal{V}_{h_{n,k}}\big(\widehat{F}_n^{(k1)},\ldots,\widehat{F}_n^{(kd_k)}\big)\end{bmatrix}-\begin{bmatrix}\mathcal{V}_{h_{n,1}}\big(F_0^{(11)},\ldots,F_0^{(1d_1)}\big)\\ \vdots\\ \mathcal{V}_{h_{n,k}}\big(F_0^{(k1)},\ldots,F_0^{(kd_k)}\big)\end{bmatrix}\right)$$

$$\rightsquigarrow\quad\begin{bmatrix}-\sum_{j=1}^{d_1}\int_{\mathbb{R}}B_{1j}(x_j-)\,\mu_{h_{0,1;\{j\}},F_0}(dx_j)\\ \vdots\\ -\sum_{j=1}^{d_k}\int_{\mathbb{R}}B_{kj}(x_j-)\,\mu_{h_{0,k;\{j\}},F_0}(dx_j)\end{bmatrix}\qquad \textit{in } (\mathbb{R}^k,\mathcal{B}(\mathbb{R}^k)).$$

Proof By Lemma 3.2.3 and assumption (a) we obtain

$$\begin{aligned}&a_n\left(\begin{bmatrix}\mathcal{V}_{h_{n,1}}\big(\widehat{F}_n^{(11)},\ldots,\widehat{F}_n^{(1d_1)}\big)\\ \vdots\\ \mathcal{V}_{h_{n,k}}\big(\widehat{F}_n^{(k1)},\ldots,\widehat{F}_n^{(kd_k)}\big)\end{bmatrix}-\begin{bmatrix}\mathcal{V}_{h_{n,1}}\big(F_0^{(11)},\ldots,F_0^{(1d_1)}\big)\\ \vdots\\ \mathcal{V}_{h_{n,k}}\big(F_0^{(k1)},\ldots,F_0^{(kd_k)}\big)\end{bmatrix}\right)\\
&= a_n\begin{bmatrix}\sum_{\emptyset\neq J\subseteq\{1,\ldots,d_1\}}(-1)^{|J|}\int_{\mathbb{R}^J}\prod_{j\in J}\big(\widehat{F}_n^{(1j)}(x_j-)-F_0^{(1j)}(x_j-)\big)\,\mu_{h_{n,1;J,F_0}}(d((x_j)_{j\in J}))\\ \vdots\\ \sum_{\emptyset\neq J\subseteq\{1,\ldots,d_k\}}(-1)^{|J|}\int_{\mathbb{R}^J}\prod_{j\in J}\big(\widehat{F}_n^{(kj)}(x_j-)-F_0^{(kj)}(x_j-)\big)\,\mu_{h_{n,k;J,F_0}}(d((x_j)_{j\in J}))\end{bmatrix}\\
&= \Phi^n\big((a_n(\widehat{F}_n^{(ij)}-F_0^{(ij)}))_{i\in\{1,\ldots,k\},j\in\{1,\ldots,d_i\}}\big)\end{aligned}$$

$\mathbb{P}$-a.s. for all $n\in\mathbb{N}$ so that the claim follows from the extended continuous mapping theorem (cf. Theorem C.1 in [13]) in view of assumptions (b)–(d). □

In a next step we show that Theorem 3.3.3 can indeed be seen as generalization of Theorem 3.3.1. For one-sample V-statistics the statement of Theorem 3.3.1 can be derived from Theorem 3.3.3 in view of the following result.

Lemma 3.3.4 *Let (a_n) be a sequence in $(0,\infty)$ with $a_n\to\infty$ as $n\to\infty$, and suppose that*

($\tilde{\text{b}}$) *the process $a_n(\widehat{F}_n-F_0)$ is a $(\mathbf{V},\mathcal{B}^\circ)$-valued random variable on the probability space $(\Omega,\mathcal{F},\mathbb{P})$ for all $n\in\mathbb{N}$, and there exists a $(\mathbf{V},\mathcal{B}^\circ)$-valued random variable B on $(\Omega',\mathcal{F}',\mathbb{P}')$ such that $B(\Omega')\subseteq S$ for some separable $S\in\mathcal{B}^\circ$ and*

$$a_n(\widehat{F}_n-F_0)\rightsquigarrow^\circ B\qquad \textit{in } (\mathbf{V},\mathcal{B}^\circ,d_{\mathbf{V}}).$$

Then assumption (b) of Theorem 3.3.3 holds (for $F_0^{(ij)}=F_0$ and $\widehat{F}_n^{(ij)}=\widehat{F}_n$ for all $i=1,\ldots,k$ and $j=1,\ldots,d_i$).

Proof For simplicity, let $\mathcal{E}_n:=(a_n(\widehat{F}_n-F_0),\ldots,a_n(\widehat{F}_n-F_0))$ and $\overline{B}:=(B,\ldots,B)$. In the following we show in three steps the measurability of $\mathcal{E}_n$ and $\overline{B}$, the existence of a separable space $\overline{S}\in\mathcal{B}^{\circ,d}$ so that $\overline{B}(\Omega')\subseteq\overline{S}$, and the convergence $\mathcal{E}_n\rightsquigarrow^\circ\overline{B}$ in $(\mathbf{V}^d,\mathcal{B}^{\circ,d},d_{\mathbf{V}}^d)$.

Step 1. We first prove the $(\mathcal{F}, \mathcal{B}^{\circ,d})$-measurability of $\mathcal{E}_n : \Omega \to \mathbf{V}^d$. Since every open ball $B_{d^d_{\mathbf{V}}}((x_1, \ldots, x_d), r)$ with respect to the metric $d^d_{\mathbf{V}}$ can be written as product $B_{d_{\mathbf{V}}}(x_1, r) \times \cdots \times B_{d_{\mathbf{V}}}(x_d, r)$ of open balls with respect to the metric $d_{\mathbf{V}}$, we have $\mathcal{B}^{\circ,d} \subset (\mathcal{B}^\circ)^{\otimes d}$. It therefore suffices to show that $\mathcal{E}_n$ is $(\mathcal{F}, (\mathcal{B}^\circ)^{\otimes d})$-measurable. Now, the σ-algebra $(\mathcal{B}^\circ)^{\otimes d}$ coincides with the σ-algebra on $\mathbf{V}^d$ generated by the coordinate projections π_i, $i \in \{1, \ldots, d\}$, given by $\pi_i(v_1, \ldots, v_d) := v_i$. The map $\mathcal{E}_n$ is then $(\mathcal{F}, (\mathcal{B}^\circ)^{\otimes d})$-measurable if and only if $\pi_i(\mathcal{E}_n)$ is $(\mathcal{F}, \mathcal{B}^\circ)$-measurable for any $i \in \{1, \ldots, d\}$, see Theorem 7.4. in [6]. This implies the $(\mathcal{F}, \mathcal{B}^{\circ,d})$-measurability of $\mathcal{E}_n$ because $a_n(\widehat{F}_n - F_0)$ is $(\mathcal{F}, \mathcal{B}^\circ)$-measurable by assumption $(\tilde{\text{b}})$. Analogously, we may deduce the $(\mathcal{F}', \mathcal{B}^{\circ,d})$-measurability of $\overline{B} : \Omega' \to \mathbf{V}^d$ from the $(\mathcal{F}', \mathcal{B}^\circ)$-measurability of $B : \Omega' \to \mathbf{V}$.

Step 2. We now prove the existence of a separable space $\overline{S} \in \mathcal{B}^{\circ,d}$ such that $\overline{B}(\Omega') \subseteq \overline{S}$. Let us use $S \in \mathcal{B}^\circ$ to denote the separable space that fulfills $B(\Omega') \subseteq S$. Then $\overline{B}(\Omega') \subseteq S \times \cdots \times S$. It, therefore, suffices to show that $S \times \cdots \times S$ is separable.

As separable space, S contains a countable dense subset I. Obviously, $I \times \cdots \times I$ is also countable. It thus remains to show that $I \times \cdots \times I$ is a dense subset of $S \times \cdots \times S$. Let $\boldsymbol{x} := (x_1, \ldots, x_d)$ be any point in $S \times \cdots \times S$ and $r > 0$. Since I is a dense subset of S, there exist $y_1, \ldots, y_d \in I$ such that $d_{\mathbf{V}}(x_j, y_j) < r$ for each $j = 1, \ldots, d$. Let $\boldsymbol{y} := (y_1, \ldots, y_d)$. Then $\boldsymbol{y} \in I \times \cdots \times I$ and fulfills $d^d_{\mathbf{V}}((x_1, \ldots, x_d), (y_1, \ldots, y_d)) = \max_{j \in \{1, \ldots, d\}}\{d_{\mathbf{V}}(x_j, y_j)\} < r$, so that $I \times \cdots \times I$ is indeed dense in $S \times \cdots \times S$.

Step 3. For the proof of $(a_n(\widehat{F}_n - F), \ldots, a_n(\widehat{F}_n - F))$ converging in distribution to $(B, \ldots, B)$ with respect to the open-ball σ-algebra, it suffices by Portmanteau's theorem (in form of Theorem A.3 in [13]) to show that

$$\int_{\mathbf{V}^d} f(\boldsymbol{x})\, \mathbb{P}_{(a_n(\widehat{F}_n - F), \ldots, a_n(\widehat{F}_n - F))}(d\boldsymbol{x}) \quad \longrightarrow \quad \int_{\mathbf{V}^d} f(\boldsymbol{x})\, \mathbb{P}'_{(B, \ldots, B)}(d\boldsymbol{x}) \tag{3.15}$$

for all uniformly continuous functions $f \in C^\circ_{\mathrm{b}}(\mathbf{V}^d)$, where $C^\circ_{\mathrm{b}}(\mathbf{V}^d)$ denotes the set of all bounded, continuous and $(\mathcal{B}^{\circ,d}, \mathcal{B}(\mathbb{R}))$-measurable real-valued functions on $\mathbf{V}^d$.

Let $f \in C^\circ_{\mathrm{b}}(\mathbf{V}^d)$ be uniformly continuous. Then we have

$$\begin{aligned}
&\Big| \int_{\mathbf{V}^d} f(\boldsymbol{x})\, \mathbb{P}_{(a_n(\widehat{F}_n - F), \ldots, a_n(\widehat{F}_n - F))}(d\boldsymbol{x}) - \int_{\mathbf{V}^d} f(\boldsymbol{x})\, \mathbb{P}'_{(B, \ldots, B)}(d\boldsymbol{x}) \Big| \\
&= \Big| \int_\Omega f\big(a_n(\widehat{F}_n(\omega) - F), \ldots, a_n(\widehat{F}_n(\omega) - F)\big)\, \mathbb{P}(d\omega) - \int_{\Omega'} f\big(B(\omega), \ldots, B(\omega)\big)\, \mathbb{P}'(d\omega) \Big| \\
&= \Big| \int_\Omega f_1\big(a_n(\widehat{F}_n(\omega) - F)\big)\, \mathbb{P}(d\omega) - \int_{\Omega'} f_1\big(B(\omega)\big)\, \mathbb{P}'(d\omega) \Big| \\
&= \Big| \int_{\mathbf{V}} f_1(x)\, \mathbb{P}_{(a_n(\widehat{F}_n - F))}(dx) - \int_{\mathbf{V}} f_1(x)\, \mathbb{P}'_B(dx) \Big|,
\end{aligned} \tag{3.16}$$

where $f_1 : \mathbf{V} \to \mathbb{R}$ is defined as $f_1(x) := f(x, \ldots, x)$. We note that f_1 is uniformly continuous and bounded, which follows immediately from the uniform continuity and boundedness of f. Moreover, f_1 is $(\mathcal{B}^\circ, \mathcal{B}(\mathbb{R}))$-measurable. We subsequently show that

$\alpha : \mathbf{V} \to \mathbf{V}^d, x \mapsto (x, \ldots, x)$ is $(\mathcal{B}^\circ, \mathcal{B}^{\circ,d})$-measurable so that f_1 is measurable as composition of the measurable functions f and α. Recall that $\mathcal{B}^{\circ,d} \subseteq (\mathcal{B}^\circ)^{\otimes d}$. For the measurability of α it therefore suffices to show that α is $(\mathcal{B}^\circ, (\mathcal{B}^\circ)^{\otimes d})$-measurable. Since $(\mathcal{B}^\circ)^{\otimes d}$ coincides with the σ-algebra generated by the projections $\pi_i : \mathbf{V}^d \to \mathbf{V}, i = 1, \ldots, d$, every $B \in (\mathcal{B}^\circ)^{\otimes d}$ can be identified with $\pi_i^{-1}(A)$ for some $A \in \mathcal{B}^\circ$ and $i \in \{1, \ldots, d\}$. Hence, for the pre-image of B under α we obtain $\alpha^{-1}(B) = \alpha^{-1}(\pi_i^{-1}(A)) = (\pi_i \circ \alpha)^{-1}(A) = \mathrm{id}^{-1}(A) = A \in \mathcal{B}^\circ$, which proves the measurability of α.

Now, the right-hand side in (3.16) converges to 0 in view of assumption ($\tilde{\text{b}}$) and Portmanteau's theorem, which proves (3.15). □

At the other extreme, in the case of multi-sample V-statistics of degree d with independent estimators instead of one-sample V-statistics of degree d with identical distribution functions and identical estimators, an analogue of Lemma 3.3.4 is valid. More precisely, assumption (b) of Theorem 3.3.3 can be replaced by an analogous condition on the components of $(a_n(\widehat{F}_n^{(i1)} - F_0^{(i1)}), \ldots, a_n(\widehat{F}_n^{(id_i)} - F_0^{(id_i)}))_{i \in \{1,\ldots,k\}}$, if the components are independent for every $n \in \mathbb{N}$.

Lemma 3.3.5 *Let (a_n) be a sequence in $(0, \infty)$ with $a_n \to \infty$ as $n \to \infty$, and suppose that*

(b') *the processes $(a_n(\widehat{F}_n^{(11)} - F_0^{(11)}))_{n \in \mathbb{N}}, \ldots, (a_n(\widehat{F}_n^{(1d_1)} - F_0^{(1d_1)}))_{n \in \mathbb{N}}, \ldots, (a_n(\widehat{F}_n^{(k1)} - F_0^{(k1)}))_{n \in \mathbb{N}}, \ldots, (a_n(\widehat{F}_n^{(kd_k)} - F_0^{(kd_k)}))_{n \in \mathbb{N}}$ are independent sequences of (potentially dependent) random variables,*

(b") *for each $i = 1, \ldots, k$ and $j = 1, \ldots, d_i$, the processes $a_n(\widehat{F}_n^{(ij)} - F_0^{(ij)})$ are $(\mathbf{V}, \mathcal{B}^\circ)$-valued random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ for all $n \in \mathbb{N}$, and there exist $(\mathbf{V}, \mathcal{B}^\circ)$-valued random variables B_{ij} on $(\Omega', \mathcal{F}', \mathbb{P}')$ such that $B_{ij}(\Omega') \subseteq S_{ij}$ for some separable $S_{ij} \in \mathcal{B}^\circ$ and*

$$a_n(\widehat{F}_n^{(ij)} - F_0^{(ij)}) \rightsquigarrow^\circ B_{ij} \qquad \text{in } (\mathbf{V}, \mathcal{B}^\circ, d_{\mathbf{V}}).$$

Then assumption (b) of Theorem 3.3.3 holds.

Proof For simplicity, let $\mathcal{E}_n = (\mathcal{E}_n^{(1)}, \ldots, \mathcal{E}_n^{(d)}) := (a_n(\widehat{F}_n^{(i1)} - F_0^{(i1)}), \ldots, a_n(\widehat{F}_n^{(id_i)} - F_0^{(id_i)}))_{i \in \{1,\ldots,k\}}$ and $B = (B_1, \ldots, B_d) := (B_{i1}, \ldots, B_{id_i})_{i \in \{1,\ldots,k\}}$. The $(\mathcal{F}, \mathcal{B}^{\circ,d})$- and $(\mathcal{F}', \mathcal{B}^{\circ,d})$-measurability of $\mathcal{E}_n : \Omega \to \mathbf{V}^d$ and $B : \Omega' \to \mathbf{V}^d$, respectively, and the existence of a separable space $\overline{S} := S_1 \times \cdots \times S_d \in \mathcal{B}^{\circ,d}$ with $B(\Omega') \subseteq \overline{S}$ can be proven just as in Step 1 and 2 in the proof of Lemma 3.3.4.

For the proof of $\mathcal{E}_n$ converging in distribution to B with respect to the open-ball σ-algebra, we adopt some arguments from the proof of Theorem 3.1(ii) in [13] (see also the proof of Theorem 2.2 in [49]). By the implication (f)⇒(a) of Theorem A.3 in [13] it suffices to show that

$$\int_{\mathbf{V}^d} f(x_1,\ldots,x_d)\,\mathbb{P}_{\mathcal{E}_n}(d(x_1,\ldots,x_d)) \quad\longrightarrow\quad \int_{\mathbf{V}^d} f(x_1,\ldots,x_d)\,\mathbb{P}'_B(d(x_1,\ldots,x_d)) \tag{3.17}$$

for all $f \in \mathrm{BL}_1^{\circ,d}$, where $\mathrm{BL}_1^{\circ,d}$ denotes the set of all $(\mathcal{B}^{\circ,d},\mathcal{B}(\mathbb{R}))$-measurable functions $f:\mathbf{V}^d\to\mathbb{R}$ satisfying $\sup_{\boldsymbol{x}\in\mathbf{V}^d}|f(\boldsymbol{x})|\le 1$ and $|f(\boldsymbol{x})-f(\boldsymbol{y})|\le d_{\mathbf{V}}^d(\boldsymbol{x},\boldsymbol{y})$ for all $\boldsymbol{x},\boldsymbol{y}\in\mathbf{V}^d$. In order to prove this, we identify $\mathcal{E}_n$ and B with their canonical processes. Let $\pi_{j,n}:\mathbf{V}^{\mathbb{N}}\times\cdots\times\mathbf{V}^{\mathbb{N}}\to\mathbf{V}$ and $\pi_j:\mathbf{V}^d\to\mathbf{V}$ be the projections defined by $\pi_{j,n}((x_{1,1},x_{1,2},\ldots),\ldots,(x_{d,1},x_{d,2},\ldots)):=x_{j,n}$ and $\pi_j(x_1,\ldots,x_d):=x_j$, respectively, for $j=1,\ldots,d$ and $n\in\mathbb{N}$. Further we denote by $(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}}):=((\mathbf{V}^{\mathbb{N}})^d,(\mathcal{B}^{\otimes\mathbb{N}})^{\otimes d},\overline{\mathbb{P}}_1\otimes\cdots\otimes\overline{\mathbb{P}}_d)$ and $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}}):=(\mathbf{V}^d,\mathcal{B}^{\otimes d},\widetilde{\mathbb{P}}_1\otimes\cdots\otimes\widetilde{\mathbb{P}}_d)$ the corresponding probability spaces with $\overline{\mathbb{P}}_j:=\mathbb{P}_{(\mathcal{E}_n^{(j)})_{n\in\mathbb{N}}}$ and $\widetilde{\mathbb{P}}_j:=\mathbb{P}'_{B_j}$ for $j=1,\ldots,d$. In the definitions of $(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}})$ and $\pi_{j,n}$ we used that $(\mathcal{E}_n^{(1)})_{n\in\mathbb{N}},\ldots,(\mathcal{E}_n^{(d)})_{n\in\mathbb{N}}$ are independent so that the measure $\mathbb{P}_{((\mathcal{E}_n^{(1)})_{n\in\mathbb{N}},\ldots,(\mathcal{E}_n^{(d)})_{n\in\mathbb{N}})}$ coincides with the product measure $\mathbb{P}_{(\mathcal{E}_n^{(1)})_{n\in\mathbb{N}}}\otimes\cdots\otimes\mathbb{P}_{(\mathcal{E}_n^{(d)})_{n\in\mathbb{N}}}$. In the following let $\widetilde{B}:=(\widetilde{B}_1,\ldots,\widetilde{B}_d)$ with $\widetilde{B}_1:=\pi_1(B),\ldots,\widetilde{B}_d:=\pi_d(B)$ and $\overline{\mathcal{E}}_n:=(\overline{\mathcal{E}}_n^{(1)},\ldots,\overline{\mathcal{E}}_n^{(d)})$ with $\overline{\mathcal{E}}_n^{(1)}:=\pi_{1,n}((\mathcal{E}_n)_n),\ldots,\overline{\mathcal{E}}_n^{(d)}:=\pi_{d,n}((\mathcal{E}_n)_n)$, where we note that $(\mathcal{E}_n)_n=(\mathcal{E}_n^{(1)},\ldots,\mathcal{E}_n^{(d)})_{n\in\mathbb{N}}=((\mathcal{E}_n^{(1)})_{n\in\mathbb{N}},\ldots,(\mathcal{E}_n^{(d)})_{n\in\mathbb{N}})$. Then we have for any $f\in\mathrm{BL}_1^{\circ,d}$

$$\begin{aligned}
&\Big|\int_{\mathbf{V}^d} f(\boldsymbol{x})\,\mathbb{P}_{(\mathcal{E}_n^{(1)},\ldots,\mathcal{E}_n^{(d)})}(d\boldsymbol{x})-\int_{\mathbf{V}^d} f(\boldsymbol{x}')\,\mathbb{P}'_{(B_1,\ldots,B_d)}(d\boldsymbol{x}')\Big|\\
&=\Big|\int_{\mathbf{V}^d} f(\boldsymbol{x})\,\overline{\mathbb{P}}_{(\overline{\mathcal{E}}_n^{(1)},\ldots,\overline{\mathcal{E}}_n^{(d)})}(d\boldsymbol{x})-\int_{\mathbf{V}^d} f(\boldsymbol{x}')\,\widetilde{\mathbb{P}}_{(\widetilde{B}_1,\ldots,\widetilde{B}_d)}(d\boldsymbol{x}')\Big|\\
&=\Big|\int_{\overline{\Omega}} f\big(\overline{\mathcal{E}}_n^{(1)}(\omega),\ldots,\overline{\mathcal{E}}_n^{(d)}(\omega)\big)\,\overline{\mathbb{P}}(d\omega)-\int_{\widetilde{\Omega}} f\big(\widetilde{B}_1(\omega'),\ldots,\widetilde{B}_d(\omega')\big)\,\widetilde{\mathbb{P}}(d\omega')\Big|\\
&=\Big|\int_{\mathbf{V}^{\mathbb{N}}\times\cdots\times\mathbf{V}^{\mathbb{N}}} f\big(\overline{\mathcal{E}}_n^{(1)}(\omega_1),\ldots,\overline{\mathcal{E}}_n^{(d)}(\omega_d)\big)\Big(\bigotimes_{j=1}^d\overline{\mathbb{P}}_j\Big)(d(\omega_1,\ldots,\omega_d))\\
&\qquad-\int_{\mathbf{V}^d} f\big(\widetilde{B}_1(\omega_1'),\ldots,\widetilde{B}_d(\omega_d')\big)\Big(\bigotimes_{j=1}^d\widetilde{\mathbb{P}}_j\Big)(d(\omega_1',\ldots,\omega_d'))\Big|.
\end{aligned}$$

Subsequently, let $\boldsymbol{\omega}:=(\omega_1,\ldots,\omega_d)$ and $\boldsymbol{\omega}':=(\omega_1',\ldots,\omega_d')$. Moreover, we use the following (slightly misleading) notation that $\overline{\mathcal{E}}_n(\boldsymbol{\omega}):=(\overline{\mathcal{E}}_n^{(1)}(\omega_1),\ldots,\overline{\mathcal{E}}_n^{(d)}(\omega_d))$ and $\widetilde{B}(\boldsymbol{\omega}'):=(B_1(\omega_1'),\ldots,B_d(\omega_d'))$. Adding telescoping sums to the latter difference yields

$$\begin{aligned}
&\Big|\int_{\mathbf{V}^d} f(\boldsymbol{x})\,\mathbb{P}_{(\mathcal{E}_n^{(1)},\ldots,\mathcal{E}_n^{(d)})}(d\boldsymbol{x})-\int_{\mathbf{V}^d} f(\boldsymbol{x}')\,\mathbb{P}'_{(B_1,\ldots,B_d)}(d\boldsymbol{x}')\Big|\\
&\le\Big|\int_{(\mathbf{V}^{\mathbb{N}})^d} f\big(\overline{\mathcal{E}}_n(\boldsymbol{\omega})\big)\Big(\bigotimes_{j=1}^d\overline{\mathbb{P}}_j\Big)(d\boldsymbol{\omega})\\
&\qquad-\int_{\mathbf{V}\times(\mathbf{V}^{\mathbb{N}})^{d-1}} f\big((\widetilde{B}(\boldsymbol{\omega}'))^{\overline{\mathcal{E}}_n(\boldsymbol{\omega});\{1\}}\big)\Big(\widetilde{\mathbb{P}}_1\otimes\bigotimes_{j=2}^d\overline{\mathbb{P}}_j\Big)(d((\boldsymbol{\omega}')^{\boldsymbol{\omega};\{1\}}))\Big|
\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{k=2}^{d-1} \Big| \int_{\mathbf{V}^{k-1}\times(\mathbf{V}^{\mathbb{N}})^{d-k+1}} f\big((\widetilde{B}(\boldsymbol{\omega}'))^{\overline{\mathcal{E}}_n(\boldsymbol{\omega});\{1,\dots,k-1\}}\big) \Big(\bigotimes_{j=1}^{k-1}\widetilde{\mathbb{P}}_j \otimes \bigotimes_{j=k}^{d}\overline{\mathbb{P}}_j\Big)(d((\boldsymbol{\omega}')^{\boldsymbol{\omega};\{1,\dots,k-1\}}))\\
&\quad - \int_{\mathbf{V}^{k}\times(\mathbf{V}^{\mathbb{N}})^{d-k}} f\big((\widetilde{B}(\boldsymbol{\omega}'))^{\overline{\mathcal{E}}_n(\boldsymbol{\omega});\{1,\dots,k\}}\big) \Big(\bigotimes_{j=1}^{k}\widetilde{\mathbb{P}}_j \otimes \bigotimes_{j=k+1}^{d}\overline{\mathbb{P}}_j\Big)(d((\boldsymbol{\omega}')^{\boldsymbol{\omega};\{1,\dots,k\}}))\Big|\\
&+ \Big| \int_{\mathbf{V}^{d-1}\times\mathbf{V}^{\mathbb{N}}} f\big((\widetilde{B}(\boldsymbol{\omega}'))^{\overline{\mathcal{E}}_n(\boldsymbol{\omega});\{1,\dots,d-1\}}\big) \Big(\Big(\bigotimes_{j=1}^{d-1}\widetilde{\mathbb{P}}_j\Big) \otimes \overline{\mathbb{P}}_d\Big)(d((\boldsymbol{\omega}')^{\boldsymbol{\omega};\{1,\dots,d-1\}}))\\
&\quad - \int_{\mathbf{V}^{d}} f\big(\widetilde{B}(\boldsymbol{\omega}')\big) \Big(\bigotimes_{j=1}^{d}\widetilde{\mathbb{P}}_j\Big)(d\boldsymbol{\omega}')\Big|\\
&=: S_1(n) + \sum_{k=2}^{d-1} S_k(n) + S_d(n),
\end{aligned}
$$

where the expression $(\boldsymbol{x}')^{\boldsymbol{x};J}$ is defined analogously to (2.2) for any $\boldsymbol{x} := (x_1,\dots,x_d)$ in $(\mathbf{V}^{\mathbb{N}})^d$, $\boldsymbol{x}' := (x_1',\dots,x_d')$ in $\mathbf{V}^d$ and $\emptyset \neq J \subseteq \{1,\dots,d\}$.

For the last summand, Fubini's theorem yields

$$
\begin{aligned}
S_d(n) &\le \int_{\mathbf{V}^{d-1}} \Big| \int_{\mathbf{V}^{\mathbb{N}}} f\big((\widetilde{B}(\boldsymbol{\omega}'))^{\overline{\mathcal{E}}_n(\boldsymbol{\omega});\{1,\dots,d-1\}}\big)\, \overline{\mathbb{P}}_d(d\omega_d)\\
&\qquad - \int_{\mathbf{V}} f\big(\widetilde{B}(\boldsymbol{\omega}')\big)\, \widetilde{\mathbb{P}}_d(d\omega_d')\Big| \Big(\bigotimes_{j=1}^{d-1}\widetilde{\mathbb{P}}_j\Big)(d(\omega_1',\dots,\omega_{d-1}'))\\
&= \int_{\mathbf{V}^{d-1}} \Big| \int_{\mathbf{V}^{\mathbb{N}}} f^{\widetilde{B}(\boldsymbol{\omega}');\{d\}}\big(\overline{\mathcal{E}}_n^{(d)}(\omega_d)\big)\, \overline{\mathbb{P}}_d(d\omega_d)\\
&\qquad - \int_{\mathbf{V}} f^{\widetilde{B}(\boldsymbol{\omega}');\{d\}}\big(\widetilde{B}_d(\omega_d')\big)\, \widetilde{\mathbb{P}}_d(d\omega_d')\Big| \Big(\bigotimes_{j=1}^{d-1}\widetilde{\mathbb{P}}_j\Big)(d(\omega_1',\dots,\omega_{d-1}')), \qquad (3.18)
\end{aligned}
$$

where we used that for any $\boldsymbol{x}' \in \mathbf{V}^d$ the function $f^{\boldsymbol{x}';\{d\}} : \mathbf{V} \to \mathbb{R}$, defined analogously to (2.4), is $(\mathcal{B}^\circ, \mathcal{B}(\mathbb{R}))$-measurable. Indeed, the function f is $(\mathcal{B}^{\circ,d}, \mathcal{B}(\mathbb{R}))$-measurable by assumption. In other words $f^{-1}(B)$ lies in $\mathcal{B}^{\circ,d}$ for every $B \in \mathcal{B}(\mathbb{R})$ and in particular $f^{-1}(B) \in (\mathcal{B}^\circ)^{\otimes d}$ in view of $\mathcal{B}^{\circ,d} \subseteq (\mathcal{B}^\circ)^{\otimes d}$. Now, Lemma 23.1 of [6] states that $A_{x_2,\dots,x_d} := \{x_1 \in \mathbf{V} : (x_1,\dots,x_d) \in A\}$ lies in $\mathcal{B}^\circ$ for every $A \in (\mathcal{B}^\circ)^{\otimes d}$, which implies that $(f^{\boldsymbol{x}';\{d\}})^{-1}(B) = (f^{-1}(B))_{x_2,\dots,x_d} := \{x_1 \in \mathbf{V} : (x_1,\dots,x_d) \in f^{-1}(B)\}$ lies in $\mathcal{B}^\circ$ for every $B \in \mathcal{B}(\mathbb{R})$. Therefore, the function $f^{\boldsymbol{x}';\{d\}}$ is $(\mathcal{B}^\circ, \mathcal{B}(\mathbb{R}))$-measurable. Since in addition $f^{\boldsymbol{x}';\{d\}}$ is bounded and Lipschitz-continuous, we obtain that $f^{\boldsymbol{x}';\{d\}} \in \mathrm{BL}_1^{\circ,1}$ for any $\boldsymbol{x}' \in \mathbf{V}^d$. By assumption (b") and Portmanteau's theorem (in form of Theorem A.3 in [13]), we have that $|\int_{\mathbf{V}} f^{\widetilde{B}(\boldsymbol{\omega}');\{d\}}(x_d)\,\mathbb{P}_{\mathcal{E}_n^{(d)}}(dx_d) - \int_{\mathbf{V}} f^{\widetilde{B}(\boldsymbol{\omega}');\{d\}}(x_d')\,\mathbb{P}'_{B_d}(dx_d')| \to 0$ and, consequently, the integrand of the outer integral of (3.18) converges to 0 for $\mathbb{P}$-almost every $\omega_1',\dots,\omega_{d-1}'$. Then the summand $S_d(n)$ also converges to 0 because $\sup_{x\in\mathbf{V}} |f^{\boldsymbol{x}';\{d\}}(x)| \le \sup_{\boldsymbol{x}\in\mathbf{V}^d} |f(\boldsymbol{x})| \le 1$ for any $\boldsymbol{x}' \in \mathbf{V}^d$ with the result that the dominated convergence theorem is applicable.

For the first summand, we obtain by Fubini's theorem

$$\begin{aligned} S_1(n) &\le \int_{(\mathbf{V}^{\mathbb{N}})^{d-1}} \Big| \int_{\mathbf{V}^{\mathbb{N}}} f^{\overline{\mathcal{E}}_n(\boldsymbol{\omega});\{1\}}\big(\overline{\mathcal{E}}_n^{(1)}(\omega_1)\big)\, \overline{\mathbb{P}}_1(d\omega_1) \\ &\qquad - \int_{\mathbf{V}} f^{\overline{\mathcal{E}}_n(\boldsymbol{\omega});\{1\}}\big(\widetilde{B}_1(\omega_1')\big)\, \widetilde{\mathbb{P}}_1(d\omega_1') \Big| \Big(\bigotimes_{j=2}^{d} \overline{\mathbb{P}}_j\Big)(d(\omega_2,\ldots,\omega_d)) \\ &\le \int_{(\mathbf{V}^{\mathbb{N}})^{d-1}} \sup_{m\in\mathbb{N}} \Big| \int_{\mathbf{V}^{\mathbb{N}}} f^{\overline{\mathcal{E}}_m(\boldsymbol{\omega});\{1\}}\big(\overline{\mathcal{E}}_n^{(1)}(\omega_1)\big)\, \overline{\mathbb{P}}_1(d\omega_1) \\ &\qquad - \int_{\mathbf{V}} f^{\overline{\mathcal{E}}_m(\boldsymbol{\omega});\{1\}}\big(\widetilde{B}_1(\omega_1')\big)\, \widetilde{\mathbb{P}}_1(d\omega_1') \Big| \Big(\bigotimes_{j=2}^{d} \overline{\mathbb{P}}_j\Big)(d(\omega_2,\ldots,\omega_d)) \\ &\le \int_{(\mathbf{V}^{\mathbb{N}})^{d-1}} \sup_{f\in \mathrm{BL}_1^{\circ,1}} \Big| \int_{\mathbf{V}^{\mathbb{N}}} f\big(\overline{\mathcal{E}}_n^{(1)}(\omega_1)\big)\, \overline{\mathbb{P}}_1(d\omega_1) \\ &\qquad - \int_{\mathbf{V}} f\big(\widetilde{B}_1(\omega_1')\big)\, \widetilde{\mathbb{P}}_1(d\omega_1') \Big| \Big(\bigotimes_{j=2}^{d} \overline{\mathbb{P}}_j\Big)(d(\omega_2,\ldots,\omega_d)) \\ &= d_{\mathrm{BL}}^{\circ}\big(\mathbb{P}_{\mathcal{E}_n^{(1)}}, \mathbb{P}'_{B_1}\big), \end{aligned}$$

where $d_{\mathrm{BL}}^{\circ}(\mathbb{P}_{\mathcal{E}_n^{(1)}}, \mathbb{P}'_{B_1}) := \sup_{f\in\mathrm{BL}_1^{\circ,1}} |\int_{\mathbf{V}} f(x)\,\mathbb{P}_{\mathcal{E}_n^{(1)}}(dx) - \int_{\mathbf{V}} f(x')\,\mathbb{P}'_{B_1}(dx')|$ is referred to as bounded Lipschitz distance. Here, we used that for any $n\in\mathbb{N}$ and $\overline{\mathbb{P}}_j$-almost every ω_j, $j=2,\ldots,d$, the function $f^{\overline{\mathcal{E}}_n(\boldsymbol{\omega});\{1\}}$ lies in the set $\mathrm{BL}_1^{\circ,1}$ of all bounded, Lipschitz continuous and $(\mathcal{B}^\circ,\mathcal{B}(\mathbb{R}))$-measurable functions by the same argumentation as for $f^{\boldsymbol{x}';\{d\}}$ above. The fact that the latter bound converges to 0 then follows from assumption (b") and the implication (a)⇒(g) of Portmanteau's theorem A.3 in [13].

For the summands $S_2(n),\ldots,S_{d-1}(n)$ one can argue just as for the first summand. For any $k\in\{2,\ldots,d-1\}$, Fubini's theorem yields

$$\begin{aligned} &S_k(n) \\ &\le \int_{\mathbf{V}^{k-1}\times(\mathbf{V}^{\mathbb{N}})^{d-k}} \Big| \int_{\mathbf{V}^{\mathbb{N}}} f^{(\widetilde{B}(\boldsymbol{\omega}'))^{\overline{\mathcal{E}}_n(\boldsymbol{\omega});\{1,\ldots,k-1\}};\{k\}}\big(\overline{\mathcal{E}}_n^{(k)}(\omega_k)\big)\, \overline{\mathbb{P}}_k(d\omega_k) \\ &\quad - \int_{\mathbf{V}} f^{(\widetilde{B}(\boldsymbol{\omega}'))^{\overline{\mathcal{E}}_n(\boldsymbol{\omega});\{1,\ldots,k-1\}};\{k\}}\big(\widetilde{B}_k(\omega_k')\big)\, \widetilde{\mathbb{P}}_k(d\omega_k') \Big| \Big(\bigotimes_{j=1}^{k-1}\widetilde{\mathbb{P}}_j \otimes \bigotimes_{j=k+1}^{d} \overline{\mathbb{P}}_j\Big)(d((\boldsymbol{\omega}')^{\boldsymbol{\omega};\{1,\ldots,k-1\}})) \\ &\le \int_{\mathbf{V}^{k-1}\times(\mathbf{V}^{\mathbb{N}})^{d-k}} \sup_{m\in\mathbb{N}} \Big| \int_{\mathbf{V}^{\mathbb{N}}} f^{(\widetilde{B}(\boldsymbol{\omega}'))^{\overline{\mathcal{E}}_m(\boldsymbol{\omega});\{1,\ldots,k-1\}};\{k\}}\big(\overline{\mathcal{E}}_n^{(k)}(\omega_k)\big)\, \overline{\mathbb{P}}_k(d\omega_k) \\ &\quad - \int_{\mathbf{V}} f^{(\widetilde{B}(\boldsymbol{\omega}'))^{\overline{\mathcal{E}}_m(\boldsymbol{\omega});\{1,\ldots,k-1\}};\{k\}}\big(\widetilde{B}_k(\omega_k')\big)\, \widetilde{\mathbb{P}}_k(d\omega_k') \Big| \Big(\bigotimes_{j=1}^{k-1}\widetilde{\mathbb{P}}_j \otimes \bigotimes_{j=k+1}^{d} \overline{\mathbb{P}}_j\Big)(d((\boldsymbol{\omega}')^{\boldsymbol{\omega};\{1,\ldots,k-1\}})) \\ &\le \int_{\mathbf{V}^{k-1}\times(\mathbf{V}^{\mathbb{N}})^{d-k}} \sup_{f\in\mathrm{BL}_1^{\circ,1}} \Big| \int_{\mathbf{V}^{\mathbb{N}}} f\big(\overline{\mathcal{E}}_n^{(k)}(\omega_k)\big)\, \overline{\mathbb{P}}_k(d\omega_k) \end{aligned}$$

$$
\begin{aligned}
&- \int_{\mathbf{V}} f\big(\widetilde{B}_k(\omega_k')\big)\, \widetilde{\mathbb{P}}_k(d\omega_k')\Big|\Big(\bigotimes_{j=1}^{k-1} \widetilde{\mathbb{P}}_j \otimes \bigotimes_{j=k+1}^{d} \overline{\mathbb{P}}_j\Big)(d((\boldsymbol{\omega}')^{\boldsymbol{\omega};\{1,\ldots,k-1\}}))\\
=\;& d_{\mathrm{BL}}^{\circ}\big(\mathbb{P}_{\mathcal{E}_n^{(k)}}, \mathbb{P}'_{B_k}\big)
\end{aligned}
$$

with $f^{(\widetilde{B}(\boldsymbol{\omega}'))^{\overline{\mathcal{E}}_m(\boldsymbol{\omega});\{1,\ldots,k-1\}};\{k\}} \in \mathrm{BL}_1^{\circ,1}$ for any $n \in \mathbb{N}$, $\widetilde{\mathbb{P}}_j$-almost every ω_j', $j = 1, \ldots, k-1$, and $\overline{\mathbb{P}}_j$-almost every ω_j, $j = k+1, \ldots, d$, by the same argumentation as above. Now, the bounded Lipschitz distance $d_{\mathrm{BL}}^{\circ}(\mathbb{P}_{\mathcal{E}_n^{(k)}}, \mathbb{P}'_{B_k})$ converges to 0 by assumption (b") and the implication (a)⇒(g) of Portmanteau's theorem (in form of Theorem A.3 in [13]), which completes the proof of (3.17). □

3.3.3 Example: Skewness

The skewness measures the asymmetry of the probability distribution of a real-valued random variable. For any distribution function F of a random variable X with finite third moment, the skewness is defined by

$$v(F) := c_3(F)/(c_2(F))^{3/2}$$

with $c_k(F) := \mathbb{E}[(X - \mathbb{E}[X])^k]$. Hence, in view of Example 3.2.7, the skewness can be expressed in terms of V-statistics as

$$\frac{\mathcal{V}_{h_3}(F,F,F)}{(\mathcal{V}_{h_2}(F,F))^{3/2}} =: \mathcal{V}(F)$$

with $h_2 : \mathbb{R}^2 \to \mathbb{R}$ and $h_3 : \mathbb{R}^3 \to \mathbb{R}$ as defined in (3.8). Let $\widehat{F}_n$ be an estimator for F. Then $\mathcal{V}(\widehat{F}_n)$ is a natural estimator for $\mathcal{V}(F)$. In particular, the asymptotic behavior of $\mathcal{V}(\widehat{F}_n)$ can easily be derived from the asymptotic behavior of $\widehat{F}_n$ by an application of Theorem 3.3.1 in combination with the delta-method.

Corollary 3.3.6 *Let (a_n) be a sequence in $(0,\infty)$ with $a_n \to \infty$ as $n \to \infty$, and let ϕ be any weight function that fulfills $\int_{\mathbb{R}} |x|^3/\phi(x)\,dx < \infty$. Assume $(F,F) \in \mathbf{F}_{h_2}$ and $(F,F,F) \in \mathbf{F}_{h_3}$, and assume that $(\widehat{F}_n, \widehat{F}_n)$ and $(\widehat{F}_n, \widehat{F}_n, \widehat{F}_n)$ are ω-wise for every $n \in \mathbb{N}$ elements of $\mathbf{F}_{h_2}$ and $\mathbf{F}_{h_3}$, respectively, such that the following assumptions are fulfilled:*

(a) *Assumption 3.2.1 is fulfilled for both triples $(h_2, (F,F), (\widehat{F}_n, \widehat{F}_n))$ and $(h_3, (F,F,F), (\widehat{F}_n, \widehat{F}_n, \widehat{F}_n))$.*

(b) *The process $a_n(\widehat{F}_n - F)$ is a $(\boldsymbol{D}_\phi, \mathcal{B}_\phi^\circ)$-valued random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ for all $n \in \mathbb{N}$, and there exists a $(\boldsymbol{D}_\phi, \mathcal{B}_\phi^\circ)$-valued random variable B on $(\Omega', \mathcal{F}', \mathbb{P}')$ such that $B(\Omega') \subseteq S$ for some separable $S \in \mathcal{B}_\phi^\circ$ and*

$$a_n(\widehat{F}_n - F) \rightsquigarrow^\circ B \qquad \text{in } (\boldsymbol{D}_\phi, \mathcal{B}_\phi^\circ, d_\phi).$$

Then

$$a_n\big(\mathcal{V}(\widehat{F}_n)-\mathcal{V}(F)\big)\rightsquigarrow\int_{\mathbb{R}}B(x-)\,\mu_{h_{2,3,F}}(dx)\qquad in\ (\mathbb{R},\mathcal{B}(\mathbb{R})),\tag{3.19}$$

where $\mu_{h_{2,3,F}}$ *is the measure generated by the function* $h_{2,3,F}(x):=JH_{(\mathcal{V}_{h_3}(F,F,F),\mathcal{V}_{h_2}(F,F))}\cdot[-3\,h_{3\{1\},F}(x),-2\,h_{2\{1\},F}(x)]'$ *and* $JH_{(x,y)}:=\big[y^{-3/2}\ \ -\frac{3}{2}\,xy^{-5/2}\big]$ *is the Jacobian matrix of the function* $H(x,y)=x/y^{3/2}$.

In particular, if B *is a continuous Gaussian process with zero mean and covariance function* γ*, then the right-hand side of (3.19) is a centered normally distributed random variable with variance* $\int_{\mathbb{R}}\int_{\mathbb{R}}\gamma(x,y)\,\mu_{h_{2,3,F}}(dx)\,\mu_{h_{2,3,F}}(dy)$.

Proof The function $H:\mathbb{R}^2\to\mathbb{R}$ defined by $H(x,y):=x/y^{3/2}$ is continuously differentiable with Jacobian $JH_{(x,y)}=\big[y^{-3/2}\ \ -\frac{3}{2}\,xy^{-5/2}\big]$. In terms of this function H, the convergence in distribution in (3.19) reads

$$\begin{aligned}&a_n\Big(H\big(\mathcal{V}_{h_3}(\widehat{F}_n,\widehat{F}_n,\widehat{F}_n),\mathcal{V}_{h_2}(\widehat{F}_n,\widehat{F}_n)\big)-H\big(\mathcal{V}_{h_3}(F,F,F),\mathcal{V}_{h_2}(F,F)\big)\Big)\\&\quad\rightsquigarrow\ -JH_{(\mathcal{V}_{h_3}(F,F,F),\mathcal{V}_{h_2}(F,F))}\cdot\begin{bmatrix}3\int_{\mathbb{R}}B(x-)\,\mu_{h_{3\{1\},F}}(dx)\\2\int_{\mathbb{R}}B(x-)\,\mu_{h_{2\{1\},F}}(dx)\end{bmatrix}\end{aligned}\tag{3.20}$$

with $\mu_{h_{3\{1\},F}}$ and $\mu_{h_{2\{1\},F}}$ as defined in (3.9) and in Example 3.2.6, respectively. In the following, we will show that

$$a_n\Big(\begin{bmatrix}\mathcal{V}_{h_3}(\widehat{F}_n,\widehat{F}_n,\widehat{F}_n)\\\mathcal{V}_{h_2}(\widehat{F}_n,\widehat{F}_n)\end{bmatrix}-\begin{bmatrix}\mathcal{V}_{h_3}(F,F,F)\\\mathcal{V}_{h_2}(F,F)\end{bmatrix}\Big)\rightsquigarrow-\begin{bmatrix}3\int_{\mathbb{R}}B(x-)\,\mu_{h_{3\{1\},F}}(dx)\\2\int_{\mathbb{R}}B(x-)\,\mu_{h_{2\{1\},F}}(dx)\end{bmatrix}.\tag{3.21}$$

Then (3.20) results from (3.21) by an application of the delta-method in form of Theorem 3.1 in [72].

For the proof of (3.21) we obtain by Theorem 3.3.1

$$a_n\Big(\begin{bmatrix}\mathcal{V}_{h_3}(\widehat{F}_n,\widehat{F}_n,\widehat{F}_n)\\\mathcal{V}_{h_2}(\widehat{F}_n,\widehat{F}_n)\end{bmatrix}-\begin{bmatrix}\mathcal{V}_{h_3}(F,F,F)\\\mathcal{V}_{h_2}(F,F)\end{bmatrix}\Big)\rightsquigarrow-\begin{bmatrix}\sum_{j=1}^{3}\int_{\mathbb{R}}B(x_j-)\,\mu_{h_{3\{j\},F}}(dx_j)\\\sum_{j=1}^{2}\int_{\mathbb{R}}B(x_j-)\,\mu_{h_{2\{j\},F}}(dx_j)\end{bmatrix},$$

provided the assumptions of Theorem 3.3.1 are fulfilled. The measures $\mu_{h_{3\{j\},F}}$ and $\mu_{h_{2\{j\},F}}$, respectively, coincide for all j because of the symmetry of the corresponding functions (see (3.9) and Example 3.2.6) so that the limit process of the latter convergence is indeed the same limit process as asserted in (3.21).

We now prove that Theorem 3.3.1 is applicable. As shown in Example 3.2.7, the assumptions of Lemma 3.2.3 are fulfilled. In Remark 3.3.2, we proved that under assumption (b) conditions (c) and (d) of Theorem 3.3.1 follow from the assumption that $\int_{\mathbb{R}^J}\prod_{j\in J}1/\phi(x_j)\,\mu^{\pm}_{h_{mJ,F}}(d((x_j)_{j\in J}))<\infty$ for all nonempty subsets $J\subseteq\{1,\dots,m\}$ and $m=2,3$. To show that the latter integral is finite under the given assumptions, we note that $h_{mJ,F}$ is $|J|$ times continuously differentiable so that

$$\mu^{\pm}_{h_{mJ,F}}(d((x_j)_{j\in J}))=\Big(\frac{\partial^{|J|}h_{mJ,F}}{\partial((x_j)_{j\in J})}((x_j)_{j\in J})\Big)^{\pm}d((x_j)_{j\in J})$$

by (3.13). Analogously to the argumentation for the second limit in the proof of assumption (d) within Example 3.2.7, the positive and negative part of the derivative $(\frac{\partial^{|J|}}{\partial((x_j)_{j\in J})}h_{m\,J,F})((x_j)_{j\in J})$ are piecewise composed of polynomials of degree at most $m-1$ in each x_j, $j \in J$, so that the claim follows from the assumption $\int_{\mathbb{R}} |x|^m/\phi(x)\,dx < \infty$.

It remains to show that the limit process is a centered normally distributed random variable with variance $\int_{\mathbb{R}}\int_{\mathbb{R}} \gamma(x,y)\,\mu_{h_{2,3,F}}(dx)\mu_{h_{2,3,F}}(dy)$, if B is a continuous centered Gaussian process with covariance function γ. Since $h_{2,3,F}$ is continuously differentiable, we have that $\int_{\mathbb{R}} B(x)\,\mu_{h_{2,3,F}}(dx) = \int_{\mathbb{R}} B(x)\,(\frac{\partial}{\partial x}h_{2,3,F})(x)\,dx$ holds. For every $a,b\in\mathbb{R}$ with $a \le b$ let $a = t_0 < t_1 < \cdots < t_n = b$ be a partition of the interval $[a,b]$, and set $\Delta t := \max_{i=0,\dots,n-1}\{t_{i+1}-t_i\}$. Then $\int_{[a,b]} B(x)\,\mu_{h_{2,3,F}}(dx)$ can be approximated by the Riemann sum $B^{a,b}_{\Delta t} := \sum_{i=0}^{n-1} B(t_i)\,(\frac{\partial}{\partial x}h_{2,3,F})(t_i)\cdot(t_{i+1}-t_i)$, where B is a Gaussian process and, therefore, $(B(t_0),\dots,B(t_{n-1}))$ is multivariate normally distributed for every $t_0,\dots,t_{n-1}\in[a,b]$. As a consequence, the random variable $B^{a,b}_{\Delta t}$ is normally distributed with zero mean and variance $v^{a,b}_{\Delta t} := \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}(\frac{\partial}{\partial x}h_{2,3,F})(t_i)(\frac{\partial}{\partial x}h_{2,3,F})(t_j)\cdot(t_{i+1}-t_i)(t_{j+1}-t_j)\,\gamma(t_i,t_j)$ and the characteristic function of $B^{a,b}_{\Delta t}$ is given by $\varphi_{B^{a,b}_{\Delta t}}(\alpha) = e^{-\alpha^2/2\cdot v^{a,b}_{\Delta t}}$. By means of the continuous mapping theorem and the dominated convergence theorem (with 1 being the dominating function) we determine the characteristic function of the integral $\int_{\mathbb{R}} B(x)\,\mu_{h_{2,3,F}}(dx)$ and obtain $\varphi(\alpha) = \mathbb{E}[\lim_{a\to-\infty,b\to\infty}\lim_{\Delta t\to 0} e^{i\alpha B^{a,b}_{\Delta t}}] = \lim_{a\to-\infty,b\to\infty}\lim_{\Delta t\to0}\varphi_{B^{a,b}_{\Delta t}}(\alpha) = e^{-v/2\cdot\alpha^2}$, where $v := \lim_{a\to-\infty,\,b\to\infty}\lim_{\Delta t\to 0} v^{a,b}_{\Delta t} = \int_{\mathbb{R}}\int_{\mathbb{R}}\gamma(x,y)\,\mu_{h_{2,3,F}}(dx)\,\mu_{h_{2,3,F}}(dy)$. This proves that the integral $\int_{\mathbb{R}} B(x)\,\mu_{h_{2,3,F}}(dx)$ is indeed centered normally distributed with variance $\int_{\mathbb{R}}\int_{\mathbb{R}}\gamma(x,y)\,\mu_{h_{2,3,F}}(dx)\,\mu_{h_{2,3,F}}(dy)$. □

3.4 The case of non-stationary time series

Let $X_{n,1},\dots,X_{n,n}$ be a non-stationary time series of the form (1.2) and recall that $F_{p,n}$ denotes the distribution function of $X_{n,i_{p,n}}$ with $i_{p,n} := \lfloor pn \rfloor$ for some fixed $p \in (0,1)$. Moreover, let $\widehat{F}_{p,n}$ be defined as in (1.3). For some given Borel measurable function $h:\mathbb{R}^d\to\mathbb{R}$ it can be reasonable to use $\mathcal{V}_h(\widehat{F}_{p,n},\dots,\widehat{F}_{p,n})$ as an estimator for $\mathcal{V}_h(F_{p,n},\dots,F_{p,n})$. We note that $\mathcal{V}_h(\widehat{F}_{p,n},\dots,\widehat{F}_{p,n})$ can be seen as a weighted V-statistic as it admits the representation

$$\mathcal{V}_h(\widehat{F}_{p,n},\dots,\widehat{F}_{p,n}) = \sum_{i_1=1}^{n}\cdots\sum_{i_d=1}^{n} w_n(i_1,\dots,i_d)\,h(X_{n,i_1},\dots,X_{n,i_d}) \tag{3.22}$$

with $w_n(i_1,\dots,i_d) := c_n^d\,\kappa\big(\frac{i_1-i_{p,n}}{nb_n}\big)\cdots\kappa\big(\frac{i_d-i_{p,n}}{nb_n}\big)$.

Applying Theorem 3.3.1 yields that under suitable assumptions (see Theorem 3.4.2 below)

$$\sqrt{nb_n}\big(\mathcal{V}_h(\widehat{F}_{p,n},\dots,\widehat{F}_{p,n}) - \mathcal{V}_h(F_{p,n},\dots,F_{p,n})\big) \rightsquigarrow Z \tag{3.23}$$

for some centered normally distributed random variable Z with variance

$$\mathbb{V}\mathrm{ar}[Z] = \int_{\mathbb{R}}\int_{\mathbb{R}} \gamma_p(x,y)\,\mu_{h_{F_p}}(dx)\mu_{h_{F_p}}(dy),$$

where $\gamma_p(x,y)$ is the covariance function defined in (1.8), and $\mu_{h_{F_p}} := \sum_{j=1}^{d} \mu_{h_{\{j\},F_p}}$. Recall that F_p denotes the distribution function of $\xi_p := \sum_{j=0}^{\ell} G_j(p,\boldsymbol{\epsilon}_0)\mathbb{1}_{(p_j,p_{j+1}]}(p)$ and that the map $h_{J,F_p} : \mathbb{R}^J \to \mathbb{R}$ is defined by $h_{\{1,\ldots,d\},F_p}(x_1,\ldots,x_d) := h(x_1,\ldots,x_d)$ for $J = \{1,\ldots,d\}$ and by $h_{J,F_p}((x_j)_{j\in J}) := \int_{\mathbb{R}^{J^c}} h^{(x_1,\ldots,x_d);J^c}((y_j)_{j\in J^c})\,(\bigotimes_{j\in J^c}\mu_{F_p})(d((y_j)_{j\in J^c}))$ for $\emptyset \neq J \subsetneq \{1,\ldots,d\}$, see Section 3.1. We now collect the required assumptions for the generalized von Mises decomposition of the left-hand side of (3.23), and we prove (3.23) in Theorem 3.4.2 below.

Let $\phi_s : \mathbb{R} \to [1,\infty)$ be the specific weight function, defined by $\phi_s(x) := (1+|x|)^s$ for some $s \in \mathbb{R}$, that we already know from Chapter 1. For brevity, we subsequently write $(\boldsymbol{D}_{(s)}, \mathcal{B}^\circ_{(s)}, \|\cdot\|_{(s)})$ instead of $(\boldsymbol{D}_{\phi_s}, \mathcal{B}^\circ_{\phi_s}, \|\cdot\|_{\phi_s})$.

Lemma 3.4.1 *Let the following assumptions be fulfilled.*

(a) *Let* $(F_p,\ldots,F_p) \in \mathbf{F}_h$ *and* $(F_{p,n},\ldots,F_{p,n}) \in \mathbf{F}_h$ *for every* $n \in \mathbb{N}$*, and let for all* $n \in \mathbb{N}$ *and for every subset* $J \subseteq \{1,\ldots,d\}$

$$\textstyle\int_{\mathbb{R}^d} \big|h(x_1,\ldots,x_d)\big|\,\big(\mu_{F_{p,n}}^{\otimes|J|} \otimes \mu_{F_p}^{\otimes|J^c|}\big)(d((x_j)_{j\in J},(x_j)_{j\in J^c})) < \infty.$$

(b) *For every nonempty subset* $J \subseteq \{1,\ldots,d\}$ *the function* h_{J,F_p} *is right continuous and* $(h_{J,F_p})^{\boldsymbol{c}_J;K}$ *is locally of bounded* $|K|$*-fold variation for every nonempty subset* $K \subseteq J$ *and every fixed* $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$.

(c) *For every nonempty subset* $J \subseteq \{1,\ldots,d\}$ *the function* $(h_{J,F_p})^{\boldsymbol{c}_J;K}_{\pm}$ *is* $(\mathcal{B}(\mathbb{R}^{|K|}), \mathcal{B}(\mathbb{R}))$*-measurable for every nonempty subset* $K \subseteq J$ *and every fixed* $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$*, and the integral*

$$\int_{\boldsymbol{I}_K^{\boldsymbol{a},\boldsymbol{b}}} (h_{J,F_p})^{\boldsymbol{c}_J;K}_{\pm}(\boldsymbol{x}_K)\left(\mu_{F_{p,n}}^{\otimes|L\cap K|} \otimes \mu_{F_p}^{\otimes|(J\setminus L)\cap K|}\right)(d((x_l)_{l\in L\cap K},(x_j)_{j\in(J\setminus L)\cap K}))$$

exists for all subsets $K, L \subseteq J$ *with* $K \neq \emptyset$, $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$, $n \in \mathbb{N}$ *and for every finite interval* $\boldsymbol{I}_K^{\boldsymbol{a},\boldsymbol{b}} := (\boldsymbol{a}_{L\cap K}, \boldsymbol{b}_{L\cap K}] \times (\boldsymbol{a}_{(J\setminus L)\cap K}, \boldsymbol{b}_{(J\setminus L)\cap K}] \subsetneq \mathbb{R}^{|K|}$, $\mathbb{P}$*-a.s., where* $(h_{J,F_p})^{\boldsymbol{c}_J;K}_{+}$ *and* $(h_{J,F_p})^{\boldsymbol{c}_J;K}_{-}$ *are* $|K|$*-fold monotonically increasing and right continuous functions satisfying* $(h_{J,F_p})^{\boldsymbol{c}_J;K} = (h_{J,F_p})^{\boldsymbol{c}_J;K}_{+} - (h_{J,F_p})^{\boldsymbol{c}_J;K}_{-}$.

(d) *For every* $n \in \mathbb{N}$ *and some* $\lambda \in [0,\infty)$*, let* $(F_{p,n}-F_p) \in \boldsymbol{D}_{(\lambda)}$ *and* $(\widehat{F}_{p,n}-F_p) \in \boldsymbol{D}_{(\lambda)}$, $\mathbb{P}$*-a.s., and*

$$\textstyle\lim_{\{|x_j|\}_{j\in J}\to\infty} \prod_{j\in J} \phi_{-\lambda}(x_j)\, h_{J,F_p}((x_j)_{j\in J}) = 0,$$
$$\textstyle\lim_{\{|x_k|\}_{k\in J\setminus L}\to\infty} \prod_{k\in J\setminus L} \phi_{-\lambda}(x_k) \int_{\mathbb{R}^L} \prod_{j\in L} \phi_{-\lambda}(y_j)\, \mu^{\pm}_{(h_{J,F_p})^{(x_k)_{k\in J};L}}(d((y_j)_{j\in L})) = 0$$

holds for all nonempty subsets $L, J \subseteq \{1,\dots,d\}$ with $L \subsetneq J$.

Then the representation

$$\begin{aligned}&\mathcal{V}_h(\widehat{F}_{p,n},\dots,\widehat{F}_{p,n}) - \mathcal{V}_h(F_p,\dots,F_p)\\ &\quad= \sum_{\emptyset\neq J\subseteq\{1,\dots,d\}} (-1)^{|J|}\int_{\mathbb{R}^J}\prod_{j\in J}\big(\widehat{F}_{p,n}(x_j-) - F_p(x_j-)\big)\,\mu_{h_{J,F_p}}(d((x_j)_{j\in J}))\end{aligned}$$

holds true $\mathbb{P}$-a.s. for all $n \in \mathbb{N}$, and for all $n \in \mathbb{N}$

$$\begin{aligned}&\mathcal{V}_h(F_{p,n},\dots,F_{p,n}) - \mathcal{V}_h(F_p,\dots,F_p)\\ &\quad= \sum_{\emptyset\neq J\subseteq\{1,\dots,d\}} (-1)^{|J|}\int_{\mathbb{R}^J}\prod_{j\in J}\big(F_{p,n}(x_j-) - F_p(x_j-)\big)\,\mu_{h_{J,F_p}}(d((x_j)_{j\in J})).\end{aligned}$$

Proof The claim follows from Lemma 3.2.3 and Remark 3.2.5 under the given assumptions, if we can show that assumptions (a) and (c) hold with $F_{p,n}$ replaced by $\widehat{F}_{p,n}$.

To show that $(\widehat{F}_{p,n},\dots,\widehat{F}_{p,n}) \in \mathbf{F}_h$ (ω-wise) for every $n \in \mathbb{N}$, we note that on the one hand $(\widehat{F}_{p,n},\dots,\widehat{F}_{p,n})$ is a tuple of distribution functions for every ω. On the other hand, the integral in (3.1) for $(\widehat{F}_{p,n},\dots,\widehat{F}_{p,n})$ in place of $(F^{(1)},\dots,F^{(d)})$ exists because it has the representation (3.22).

Analogously to representation (3.22), we have $\mathbb{P}$-a.s.

$$\begin{aligned}&\int_{\mathbb{R}^d}\big|h(x_1,\dots,x_d)\big|\,\Big(\mu_{\widehat{F}_{p,n}}^{\otimes|J|}\otimes\mu_{F_p}^{\otimes|J^{\mathsf{c}}|}\Big)(d((x_j)_{j\in J},(x_j)_{j\in J^{\mathsf{c}}}))\\ &\quad= \sum_{1\le i_j\le n,\ j\in J} c_n^{|J|}\prod_{j\in J}\kappa\Big(\frac{i_j - i_{p,n}}{nb_n}\Big)\\ &\qquad\cdot\int_{\mathbb{R}^{J^{\mathsf{c}}}}\big|h\big((X_{n,i_1},\dots,X_{n,i_d})^{(x_1,\dots,x_d);J}\big)\big|\;\mu_{F_p}^{\otimes|J^{\mathsf{c}}|}(d((x_j)_{j\in J^{\mathsf{c}}}))\end{aligned}\tag{3.24}$$

for all $n \in \mathbb{N}$ and for every nonempty subset $J \subsetneq \{1,\dots,d\}$, where $\boldsymbol{x}^{\boldsymbol{y};J}$ is defined as in (2.2) for all $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^d$. Since the integral on the right-hand side of (3.24) exists for $\mathbb{P}$-almost every ω, for all $n \in \mathbb{N}$ and every nonempty subset $J \subsetneq \{1,\dots,d\}$ by assumption (a), the integral on the left-hand side of (3.24) exists $\mathbb{P}$-a.s. for all $n \in \mathbb{N}$ and every nonempty subset $J \subsetneq \{1,\dots,d\}$.

In exactly the same way we can show that $\mathbb{P}$-a.s for all subsets $K, L \subseteq J$ with $K \neq \emptyset$, $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$, $n \in \mathbb{N}$ and for every finite interval $\boldsymbol{I}_K^{\boldsymbol{a},\boldsymbol{b}} \subsetneq \mathbb{R}^{|K|}$ the integral

$$\int_{\boldsymbol{I}_K^{\boldsymbol{a},\boldsymbol{b}}}(h_{J,F_p})_{\pm}^{\boldsymbol{c}_J;K}(\boldsymbol{x}_K)\,\Big(\mu_{\widehat{F}_{p,n}}^{\otimes|L\cap K|}\otimes\mu_{F_p}^{\otimes|(J\setminus L)\cap K|}\Big)(d((x_l)_{l\in L\cap K},(x_j)_{j\in(J\setminus L)\cap K}))$$

has a similar representation to the right-hand side of (3.24), so that the existence of the latter integral follows from assumption (c) for all subsets $K, L \subseteq J$ with $K \neq \emptyset$, $\boldsymbol{c}_J \in \mathbb{R}^{|J|}$, $n \in \mathbb{N}$ and for every finite interval $\boldsymbol{I}_K^{\boldsymbol{a},\boldsymbol{b}} \subsetneq \mathbb{R}^{|K|}$, $\mathbb{P}$-a.s. □

Theorem 3.4.2 *Suppose that for some* $\lambda \in [0,\infty)$ *the assumptions of Lemma 3.4.1 hold and* $\int_{\mathbb{R}^J} \prod_{j\in J} \phi_{-\lambda}(x_j)\, \mu^{\pm}_{h_{J,F_p}}(d((x_j)_{j\in J})) < \infty$ *for all subsets* $\emptyset \neq J \subseteq \{1,\ldots,d\}$*. If additionally* $\lim_{n\to\infty} nb_n = \infty$*,* $\sqrt{nb_n}\|F_{p,n} - F_p\|_{(\lambda)} \to 0$ *and* $\sqrt{nb_n}(\widehat{F}_{p,n}(\cdot) - F_p(\cdot)) \rightsquigarrow^{\circ} B_p$ *in* $(\boldsymbol{D}_{(\lambda)}, \mathcal{B}^{\circ}_{(\lambda)}, \|\cdot\|_{(\lambda)})$ *for some continuous centered Gaussian process with covariance function* γ_p*, then (3.23) is valid.*

Proof According to Lemma 3.4.1, we have that $\mathcal{V}_h(F_{p,n},\ldots,F_{p,n})$, $\mathcal{V}_h(F_p,\ldots,F_p)$ and $\mathcal{V}_h(\widehat{F}_{p,n},\ldots,\widehat{F}_{p,n})$ ($\mathbb{P}$-a.s.) exist for all $n \in \mathbb{N}$ so that

$$\begin{aligned}
&\sqrt{nb_n}\big(\mathcal{V}_h(\widehat{F}_{p,n},\ldots,\widehat{F}_{p,n}) - \mathcal{V}_h(F_{p,n},\ldots,F_{p,n})\big)\\
&\quad= \sqrt{nb_n}\big(\mathcal{V}_h(\widehat{F}_{p,n},\ldots,\widehat{F}_{p,n}) - \mathcal{V}_h(F_p,\ldots,F_p)\big)\\
&\qquad+\sqrt{nb_n}\big(\mathcal{V}_h(F_p,\ldots,F_p) - \mathcal{V}_h(F_{p,n},\ldots,F_{p,n})\big)\\
&\quad=: S_1(n) + S_2(n) \qquad (3.25)
\end{aligned}$$

$\mathbb{P}$-a.s. for all $n \in \mathbb{N}$. For the second summand we obtain by Lemma 3.4.1

$$\begin{aligned}
&|S_2(n)|\\
&= \Big|\sqrt{nb_n} \sum_{\emptyset\neq J\subseteq\{1,\ldots,d\}} (-1)^{|J|} \int_{\mathbb{R}^J} \prod_{j\in J} \big(F_{p,n}(x_j-) - F_p(x_j-)\big)\, \mu_{h_{J,F_p}}(d((x_j)_{j\in J}))\Big|\\
&\leq \sqrt{nb_n} \sum_{\emptyset\neq J\subseteq\{1,\ldots,d\}} \int_{\mathbb{R}^J} \prod_{j\in J} \Big(\|F_{p,n} - F_p\|_{(\lambda)}\, \phi_{-\lambda}(x_j-)\Big) \big(\mu^{+}_{h_{J,F_p}} + \mu^{-}_{h_{J,F_p}}\big)(d((x_j)_{j\in J}))\\
&= \sum_{\emptyset\neq J\subseteq\{1,\ldots,d\}} (nb_n)^{(1-|J|)/2}\\
&\quad \cdot\Big(\sqrt{nb_n}\|F_{p,n} - F_p\|_{(\lambda)}\Big)^{|J|} \int_{\mathbb{R}^J} \prod_{j\in J} \phi_{-\lambda}(x_j)\, \big(\mu^{+}_{h_{J,F_p}} + \mu^{-}_{h_{J,F_p}}\big)(d((x_j)_{j\in J})),
\end{aligned}$$

which converges to 0 because $\sqrt{nb_n}\|F_{p,n} - F_p\|_{(\lambda)} \to 0$ and the latter integral is finite by assumption.

For the summand $S_1(n)$ we note that $\sqrt{nb_n}(\widehat{F}_{p,n}(\cdot) - F_p(\cdot)) \rightsquigarrow^{\circ} B_p$ in the metric space $(\boldsymbol{D}_{(\lambda)}, \mathcal{B}^{\circ}_{(\lambda)}, \|\cdot\|_{(\lambda)})$ by our assumptions. By means of Theorem 3.3.1, the summand $S_1(n)$ converges in distribution to the expression $-\sum_{j=1}^{d} \int_{\mathbb{R}} B_p(x_j)\, \mu_{h_{\{j\},F_p}}(dx_j)$, if we can show that the assumptions of Theorem 3.3.1 hold true. In this case, we could conclude by Slutsky's theorem that

$$\sqrt{nb_n}\big(\mathcal{V}_h(\widehat{F}_{p,n},\ldots,\widehat{F}_{p,n}) - \mathcal{V}_h(F_{p,n},\ldots,F_{p,n})\big) \rightsquigarrow -\int_{\mathbb{R}} B_p(x)\, \mu_{h_{F_p}}(dx) \qquad (3.26)$$

with $\mu_{h_{F_p}} := \sum_{j=1}^{d} \mu_{h_{\{j\},F_p}}$ in view of (3.25). We now verify that the assumptions of Theorem 3.3.1 hold true. Since $\mathcal{B}^{\circ}_{(\lambda)}$ coincides with the σ-algebra generated by the one-dimensional coordinate projections $\pi_x : \boldsymbol{D}_{(\lambda)} \to \mathbb{R}, v \mapsto v(x)$ (recall Lemma 4.1 in [13]), the $(\mathcal{F}, \mathcal{B}^{\circ}_{(\lambda)})$-measurability of $\omega \mapsto \sqrt{nb_n}(\widehat{F}_{p,n}(\omega,\cdot) - F_{p,n}(\cdot))$ is a direct consequence of the $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$-measurability of $\omega \mapsto \pi_x\big(\sqrt{nb_n}(\widehat{F}_{p,n}(\omega,\cdot) - F_{p,n}(\cdot))\big)$. Along with our assumptions, this implies conditions (a) and (b) of Theorem 3.3.1. Assumptions (c) and (d) follow from the assumption that $\int_{\mathbb{R}^J} \prod_{j\in J} \phi_{-\lambda}(x_j)\, \mu^{\pm}_{h_{J,F_p}}(d((x_j)_{j\in J}))$ is finite for all nonempty subsets $J \subseteq \{1,\dots,d\}$ in view of Remark 3.3.2.

It thus remains to show that the limit process is a normally distributed random variable with zero mean and variance $\int_{\mathbb{R}}\int_{\mathbb{R}} \gamma_p(x,y)\, \mu_{h_{F_p}}(dx)\mu_{h_{F_p}}(dy)$, where $\gamma_p(x,y)$ is the covariance function defined in (1.8). For every $a, b \in \mathbb{R}$ with $a \le b$ the integral $\int_{[a,b]} -B_p(x)\, \mu_{h_{F_p}}(dx)$ can be approximated by $\sum_{i=0}^{n-1} -B_p(t_i)(h_{F_p}(t_{i+1}) - h_{F_p}(t_i))$, where $a = t_0 < t_1 < \cdots < t_n = b$ is a partition of the interval $[a,b]$. Since B_p is a Gaussian process, the random variable $(B_p(t_0),\dots,B_p(t_{n-1}))$ is multivariate normally distributed for every $t_0,\dots,t_n \in [a,b]$. Hence the sum $\sum_{i=0}^{n-1} -B_p(t_i)(h_{F_p}(t_{i+1})-h_{F_p}(t_i))$ is normally distributed with zero mean and variance $v^{a,b}_{\Delta t} := \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}(h_{F_p}(t_{i+1})-h_{F_p}(t_i))(h_{F_p}(t_{j+1})-h_{F_p}(t_j))\,\gamma_p(t_i,t_j)$, where $\gamma_p(t_i,t_j) = \mathbb{C}\mathrm{ov}\big(B_p(t_i), B_p(t_j)\big)$ and $\Delta t := \max_{i=0,\dots,n-1}\{t_{i+1} - t_i\}$. Note that $v := \lim_{a\to-\infty,\, b\to\infty} \lim_{\Delta t\to 0} v^{a,b}_{\Delta t} = \int_{\mathbb{R}}\int_{\mathbb{R}} \gamma_p(x,y)\, \mu_{h_{F_p}}(dx)\, \mu_{h_{F_p}}(dy)$. The claim now follows from the fact that the characteristic function of $\int_{\mathbb{R}} -B_p(x)\, \mu_{h_{F_p}}(dx)$ equals $\varphi(\alpha) = e^{-v/2\cdot\alpha^2}$, which can be shown by the same argumentation as in the proof of Corollary 3.3.6. □

In view of Lemma 1.2.2, Lemma 1.2.6 and Theorem 1.2.4, Theorem 3.4.2 leads to the following result.

Corollary 3.4.3 *Suppose that the assumptions (A1)–(A9), (B2) and (B4) in Subsection 1.2.2 hold for some common* $\lambda \in [0,\infty)$. *If for the same* λ *the assumptions of Lemma 3.4.1 are fulfilled and* $\int_{\mathbb{R}^J} \prod_{j\in J} \phi_{-\lambda}(x_j)\, \mu^{\pm}_{h_{J,F_p}}(d((x_j)_{j\in J})) < \infty$ *for all subsets* $\emptyset \ne J \subseteq \{1,\dots,d\}$*, then (3.23) is valid.*

Bibliography

[1] Aistleitner, C. and Dick, J. (2015). Functions of bounded variation, signed measures, and a general Koksma–Hlawka inequality. *Acta Arithmetica*, 167, 143–171.

[2] Aistleitner, C., Pausinger, F., Svane, A.M. and Tichy, R.F. (2017). On functions of bounded variation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 162, 405–418.

[3] Andrews, D.W.K. (1991). An empirical process central limit theorem for dependent non-identically distributed random variables. *Journal of Multivariate Analysis*, 38, 187–203.

[4] Arcones, M.A. and Yu, B. (1994). Central limit theorems for empirical and U-processes of stationary mixing sequences. *Journal of Theoretical Probability*, 7, 47–71.

[5] Basrak, B., Davis, R.A. and Mikosch, T. (2002). Regular variation of GARCH processes. *Stochastic Processes and their Applications*, 99, 95–115.

[6] Bauer, H. (2001). *Measure and integration theory*. de Gruyter, Berlin.

[7] Berghaus, B., Bücher, A. and Volgushev, S. (2017). Weak convergence of the empirical copula process with respect to weighted metrics. *Bernoulli*, 23, 743–772.

[8] Beutner, E., Wu, W.B. and Zähle, H. (2012). Asymptotics for statistical functionals of long-memory sequences. *Stochastic Processes and their Applications*, 122, 910–929.

[9] Beutner, E. and Zähle, H. (2010). A modified functional delta method and its application to the estimation of risk functionals. *J. Multivariate Anal.* 101, 2452–2463.

[10] Beutner, E., Zähle, H. (2012). Deriving the asymptotic distribution of U-statistics and V-statistics of dependent data using weighted empirical processes. *Bernoulli*, 18, 803–822.

[11] Beutner, E. and Zähle, H. (2014). Continuous mapping approach to the asymptotics of U- and V-statistics. *Bernoulli*, 20, 846–877.

[12] Beutner, E. and Zähle, H. (2014). Supplement to "Continuous mapping approach to the asymptotics of U- and V-statistics". DOI:10.3150/13-BEJ508SUPP.

[13] Beutner, E. and Zähle, H. (2016). Functional delta-method for the bootstrap of quasi-Hadamard differentiable functionals. *Electronic Journal of Statistics*, 10, 1181–1222.

[14] Beutner, E. and Zähle, H. (2018). Bootstrapping Average Value at Risk of single and collective risks. *Risks*, 6, online first.

[15] Billingsley, P. (1999). *Convergence of probability measures.* Wiley, New York.

[16] Bougerol, P. and Picard, N. (1992). Stationarity of GARCH processes and some nonnegative time series. *Journal of Econometrics*, 52, 115–127.

[17] Bradley, R.C. (2005). Basic properties of strong mixing conditions. A survey and some open problems. *Probability Surveys*, 2, 107–144.

[18] Chow, Y.S. and Teicher, H. (1988). *Probability Theory: Independence, Interchangeability, Martingales.* Second Edition. Springer-Verlag, New York.

[19] Dahlhaus, R. (1996). On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Processes and their Applications*, 62, 139–168.

[20] Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Annals of Statistics*, 25, 1–37.

[21] Dahlhaus, R. (2012). Locally stationary processes. *Handbook of statistics*, 30, 351–413.

[22] Dahlhaus, R., Richter, S. and Wu, W.B. (2019). Towards a general theory for nonlinear locally stationary processes. *Bernoulli*, in press.

[23] Dahlhaus, R. and Subba Rao, S. (2006). Statistical Inference for time-varying ARCH processes. *Annals of Statistics*, 34, 1074–1114.

[24] Dehling, H. (2006). *Limit theorems for dependent U-statistics.* Lecture Notes in Statistics, Springer-Verlag, New York.

[25] Dehling, H. and Taqqu, M.S. (1991). Bivariate symmetric statistics of long-range dependent observations. *J. Statist. Plann. Inference* 28, 153–165.

[26] Dehling, H. and Wendler, M. (2010). Central limit theorem and the bootstrap for U-statistics of strongly mixing data. *Journal of Multivariate Analysis*, 101(1), 126–137.

[27] Denker, M. (1985). *Asymptotic distribution theory in nonparametric statistics.* Vieweg & Sohn, Braunschweig/Wiesbaden.

[28] Denker, M. and Keller, G. (1986). Rigorous statistical procedures for data from dynamical systems. *Journal of Statistical Physics*, 44, 67–93.

[29] Dewan, I. and Prakasa Rao, B.L.S. (2001). Asymptotic normality for U-statistics of asscoiated sequences. *Journal of Statistical Planning and Inference*, 97, 201–225.

[30] Dewan, I. and Prakasa Rao, B.L.S. (2002). Central limit theorems of U-statistics of associated random variables. *Statistics and Probability Letters*, 57, 9–15.

[31] Dewan, I. and Prakasa Rao, B.L.S (2015). Corrigendum to "Central limit theorems of U-statistics of associated random variables" [Statistics and Probability Letters, 57 (2002), 9–15]. *Statistics and Probability Letters*, 106, 147–148.

[32] Doob, J.L. (1953). *Stochastic Processes.* Wiley, New York.

[33] Doukhan, P. (1994). Mixing: Properties and Examples. *Lecture Notes in Statistics*, Springer-Verlag, New York.

[34] Fermanian, J.-D. (1998). Contributions à l'Analyse Nonparamétrique des Fonctions de Hasard sur Données Multivariées et Censurées. *PhD thesis*, Université Paris VI.

[35] Fryzlewicz, P., Sapatinas, T., and Subba Rao, S. (2008). Normalized least-squares estimation in time-varying ARCH models. *Annals of Statistics*, 36, 742–786.

[36] Fryzlewicz, P. and Subba Rao, S. (2011). Mixing properties of ARCH and time-varying ARCH processes. *Bernoulli*, 17, 320–346.

[37] Garg, M. and Dewan, I. (2015). On asymptotic behavior of U-statistics for associated random variables. *Statistics and Probability Letters*, 105, 209–220.

[38] Garg, M. and Dewan, I. (2018). On limiting distribution of U-statistics based on associated random variables. *Statistics and Probability Letters*, 132, 7–16.

[39] Ghorpade, S.R. and Limaye, B.V. (2010). *A course in multivariable calculus and analysis.* Springer-Verlag, New York.

[40] Gill, R.D., van der Laan, M.J. and Wellner, J.A. (1995). Inefficient estimators of the bivariate survival function of three models. *Annales de l'Institut de Henri Poincaré (B)*, 31, 545–597.

[41] Giraitis, L., Kokoszka, P. and Leipus, R. (2000). Stationary ARCH models: dependence structure and central limit theorem. *Econometric Theory*, 16, 3–22.

[42] Halmos, P.R. (1946). The theory of unbiased estimation. *Annals of Mathematical Statistics* 17, 34–43.

[43] Hannan, E.J. (1973). Central limit theorems for time series regression. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 26, 157–170.

[44] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics*, 19, 293–325.

[45] Hoffmann-Jørgensen, J. (1984). Stochastic processes in Polish spaces. *Unpublished manuscript.*

[46] Kallenberg, O. (2002). *Foundations of modern probability.* Springer-Verlag, Berlin.

[47] Klenke, A. (2014). *Probability theory: A Comprehensive Course.* Springer-Verlag, London.

[48] Koroljuk, V.S. and Borovskich, Yu.V. (1994). *Theory of U-statistics.* Kluwer Academic Publishers, Dordrecht.

[49] Kosorok, M.R. (2008). Bootstrapping the Grenander estimator. *In* Balakrishnan, N., Peña, E. A., Silvapulle, M. J. (Eds.). *Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen.*, Institute of Mathematical Statistics: Hayward, CA., 282–292.

[50] Lee, A.J. (1990). *U-Statistics.* Marcel Dekker Inc., New York.

[51] Leonov, A.S. (1998). On the total variation for functions of several variables and a multidimensional analog of Helly's selection principle. *Mathematical Notes*, 63, 61–71.

[52] Leucht, A. (2012). Degenerate U- and V-statistics under weak dependence: Asymptotic theory and bootstrap consistency. *Bernoulli*, 18, 552–585.

[53] Leucht, A. and Neumann M.H. (2013). Degenerate U- and V-statistics under ergodicity: asymptotics, bootstrap and applications in statistics. *Annals of the Institute of Statistical Mathematics*, 65, 349–386.

[54] Leucht, A. and Neumann M.H. (2013). Dependent wild bootstrap for degenerate U- and V-statistics. *Journal of Multivariate Analysis*, 117, 257–280.

[55] Mayer, U., Zähle, H. and Zhou, Z. (2019). Functional weak limit theorem for a local empirical process of non-stationary time series and its application. *Submitted.*

[56] Nason, G. P., von Sachs, R. and Kroisandt, G. (2000). Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *Journal of the Royal Statistical Society. Series B*, 62, 271-292.

[57] Neumann, M.H. and von Sachs, R. (1997). Wavelet thresholding in anisotropic function classes and applications to adaptive estimation of evolutionary spectra. *Annals of Statistics*, 25, 38–76.

[58] Nourdin, I. (2012). *Selected aspects of fractional Brownian motion.* Springer-Verlag, Milan.

[59] Owen, A.B. (2005). Multidimensional variation for quasi-Monte Carlo. In *Contemporary multivariate analysis and design of experiments. Series in Biostatistics*, 2, 49–74. World Scientific Publishing, Hackensack, New Jersey.

[60] Pollard, D. (1984). *Convergence of stochastic processes.* Springer-Verlag, New York.

[61] Priestley, M.B. (1965). Evolutionary spectra and non-stationary processes. *Journal of the Royal Statistical Society. Series B*, 27, 204–237.

[62] Priestley, M.B. (1981). *Spectral Analysis and Time Series: Multivariate series, prediction and control.* Academic Press, London.

[63] Radulović, D., Wegkamp, M. and Zhao, Y. (2017). Weak convergence of empirical copula processes indexed by functions. *Bernoulli*, 23, 3346–3384.

[64] Richter, H. (1956). *Wahrscheinlichkeitstheorie.* Springer-Verlag, Berlin.

[65] Sen, P.K. (1972). Limiting behavior of regular functionals of empirical distributions for stationar *-mixing processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 25, 71–82.

[66] Sen, P.K. (1981). *Sequential nonparametrics: invariance principles and statistical inference.* Wiley, New York.

[67] Sen, P.K. (1992). Introduction to Hoeffding (1948): A class of statistics with asymptotically normal distribution. In Kotz, S., Johnson, N.L. *Breakthroughs in Statistics*, Vol. I, 299–307, Springer-Verlag.

[68] Serfling, R.J. (1980). *Approximation theorems of mathematical statistics*, Wiley, New York.

[69] Shao, Q.-M. and Yu, H. (1996). Weak convergence for weighted empirical processes of dependent sequences. *Ann. Probab.* 24, 2098–2127.

[70] Shorack, G.R. and Wellner, J.A. (1986). *Empirical processes with applications to statistics*. Wiley, New York.

[71] Straumann, D. and Mikosch, T. (2006). Quasi-maximum likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equation approach. *Annals of Statistics*, 34, 2449–2495.

[72] Van der Vaart, A.W. (1998). *Asymptotic statistics*. Cambridge University Press, Cambridge.

[73] Van der Vaart, A.W. and Wellner, J.A. (1996). *Weak convergence and empirical processes*. Springer-Verlag, New York.

[74] von Mises, R. (1947). On the asymptotic distribution of differentiable statistical functions. *Annals of Mathematical Statistics* 18, 309–348.

[75] Wu, W.B. (2003). Empirical processes of long-memory sequences. *Bernoulli*, 9, 809–831.

[76] Wu, W.B. (2005). Nonlinear system theory: Another look at dependence. *Proceedings of the National Academy of Sciences of the USA*, 102, 14150–14154.

[77] Wu, W.B. (2007). Strong invariance principles for dependent random variables. *Annals of Probability*, 35, 2294–2320.

[78] Wu, W.B. (2008). Empirical processes of stationary sequences. *Statistica Sinica*, 18, 313–333.

[79] Wu, W. and Zhou, Z. (2018). Gradient-based structural change detection for non-stationary time series M-estimation. *Annals of Statistics*, 46, 1197–1224.

[80] Yoshihara, K. (1976). Limiting behavior of U-statistics for stationary, absolutely regular processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 35, 237–252.

[81] Young, W.H. (1917). On multiple integration by parts and the second theorem of the mean. *Proceedings of the London Mathematical Society (2)*, 16, 273–293.

[82] Zhou, Z. (2013). Heteroscedasticity and autocorrelation robust structural change detection. *Journal of the American Statistical Association*, 108, 726–740.

[83] Zhou, Z. (2014). Inference of weighted V-statistics for non-stationary time series and its applications. *Annals of Statistics*, 42, 87–114.

[84] Zhou, Z. and Wu, W.B. (2009). Local linear quantile estimation for nonstationary time series. *Annals of Statistics*, 37, 2696–2729.