Least Squares Estimation in Multiple Change-Point Models

Von der

Fakultät Mathematik und Naturwissenschaften
der Technischen Universität Dresden

genehmigte

Dissertation

zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

von

René Mauer, M.Sc.

geboren am 21. Februar 1987 in Pirna

Die Dissertation wurde in der Zeit von Oktober 2014 bis April 2018 unter der Betreuung von Prof. Dr. Dietmar Ferger im Institut für Mathematische Stochastik angefertigt.

Eingereicht am: 17. April 2018
Tag der Disputation: 17. August 2018

Gutachter: Prof. Dr. Dietmar Ferger, TU Dresden
Prof. Dr. Ansgar Steland, RWTH Aachen

Contents

Index of notation

Analysis

$\lfloor\cdot\rfloor$	floor function, $\lfloor x\rfloor := \max_{y\in\mathbb{Z},y\leq x}\{y\}$
$\|\cdot\|$	maximum norm
$\xrightarrow[n\to\infty]{}$, $\lambda\downarrow 0$	point-wise convergence, convergence from above
$A^{\complement}$, $\lvert A\rvert$	complement, cardinality of the set A
$\lvert a\rvert$	absolute value of the real number a
$\mathbf{a}^T$	transpose of the vector $\mathbf{a}$
$\text{Argmin}(f)$, $\text{Argmax}(f)$, $\text{Arginf}(f)$	set of minimizers, maximizers, infimizers of f
$\text{argmin}(f)$, $\text{argmax}(f)$	an arbitrary minimizer, maximizer of f
$B_r(x)$	open ball, center x, radius r
$D(A)$	Skorokhod space on a set $A\subseteq\mathbb{R}^d$, $d\in\mathbb{N}$
$\mathbb{1}_A$	indicator function of the set A
$f\vert_A$	function f restricted on the domain A
f'	derivative of the function f
$\mathbb{N}_0$	natural numbers including 0
$\mathbb{N}_m$	natural numbers greater than or equal to m
$\mathbb{Q}_{>0}$, $\mathbb{R}_{>0}$	rational, real numbers greater than 0

Probability theory

$\sim$	distributed as
$\stackrel{\mathcal{L}}{=}$	equality in distribution
$\xrightarrow[n\to\infty]{a.s.}$	almost sure convergence (with respect to $\mathbb{P}$)
$\xrightarrow[n\to\infty]{\mathbb{P}}$	convergence in probability
$\xrightarrow[n\to\infty]{\mathcal{L}}$	convergence in distribution
$\bigotimes_{i=1}^{n} \mathbb{P}_{X_i}$, $\mathbb{P}_X \otimes \mathbb{P}_Y$	product measure
$a.s.$	almost surely (with respect to $\mathbb{P}$)
$B(m, p)$	Binomial distribution with parameters m, p
$\mathcal{B}(\mathbb{R})$	Borel σ-algebra on $\mathbb{R}$
$\mathbb{E}$	expectation
$\mathbb{E}[Y\|Z]$	conditional expectation of Y with respect to Z
$Exp(\lambda)$	Exponential distribution with parameter λ
$N(\mu, \sigma)$	Normal distribution with parameters μ, σ
$\mathbb{P}$	probability
$\mathbb{P}_Z$	distribution of a random variable/vector Z
$Poi(\lambda)$	Poisson distribution with parameter λ
$\mathbb{V}$	variance
Z^+	positive part of a random variable Z
$\sigma(Z_1, ..., Z_n)$	σ-algebra generated by $Z_1, ..., Z_n$

Multiple change-point model

a_1, a_2, $a^*_{n,1}$, $a^*_{n,2}$	auxiliary functions, p. 22, 105
$G_{n,x,\delta}$	set, p. 70
H_n, $H_{n,x,\delta}$	set, p. 38, 44
$\bar{M}_n$, $\hat{M}_n$, M^*_n	criterion functions, p. 22, 60, 105
M_p	maximum of p-th absolute moments, p. 5
S_n, $\bar{S}_n$, S^*_n	criterion functions, p. 60, 22, 104
$S_{u,v}$	sum of centered observations, p. 147
$\bar{X}_{u,v}$	mean value of observations, p. 59
$\boldsymbol{\alpha} = (\alpha, \beta, \gamma)$	expectations, p. 5
$\hat{\boldsymbol{\alpha}}_n = \left(\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n\right)$	estimator of $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)$, p. 62
Γ, Γ_1, Γ_2	limit process, p. 39
$\bar{\Gamma}_n$, $\bar{\Gamma}_{n,1}$, $\bar{\Gamma}_{n,2}$	rescaled process, p. 38
Γ^*_n, $\Gamma^*_{n,1}$, $\Gamma^*_{n,2}$	rescaled process, p. 114
Δ_n	set, p. 4
$\hat{\delta}_n$, $\hat{\varrho}_n$, δ^*_n, ϱ^*_n	functions, decomposition of $\hat{\rho}_n$, ρ^*_n, p. 147, 149
Θ, Θ_n	sets, p. 4, 25
$\Theta^1, ..., \Theta^6$	sets, partition of Θ, p. 29
$\boldsymbol{\theta} = (\theta_1, \theta_2)$	multiple change-point, p. 4
$\bar{\boldsymbol{\theta}}_n$, $\hat{\boldsymbol{\theta}}_n$, $\boldsymbol{\theta}^*_n$	estimators of $\boldsymbol{\theta}$, p. 25, 63, 106
ρ, $\hat{\rho}$	limit functions, p. 32, 64
$\bar{\rho}_n$, $\hat{\rho}_n$, ρ^*_n	criterion functions, p. 25, 63, 106
$\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$	moments of change, p. 5
$\bar{\boldsymbol{\tau}}_n = (\bar{\tau}_n, \bar{\sigma}_n)$, $\hat{\boldsymbol{\tau}}_n = (\hat{\tau}_n, \hat{\sigma}_n)$, $\boldsymbol{\tau}^*_n = (\tau^*_n, \sigma^*_n)$	estimators of $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$, p. 23, 62, 106

Abbreviations

Cor.	Corollary
def.	definition
In.	Inequality
Lem.	Lemma
Por. Thm.	Portmanteau Theorem
Tr. In.	Triangle Inequality

Chapter 1

Preliminaries

1.1 Introduction

Change-point analysis is a field of mathematical statistics, which concerns itself with the detection and estimation of structural changes within a data set of time-ordered observations. To reach this target, approximately homogeneous observations are assembled into segments, which are established on the basis of some criteria such as expectations or variances of the underlying distributions. This type of problems appears in various scientific fields. One of the first applications was quality control in companies where the goal is to find out whether the quality of products is deteriorated from a certain point. Furthermore, in biology change-point models are used for segmentation of DNA sequences (see for instance Braun and Müller [3]). Some more applications are indicated in Chen and Gupta [6] and Fremdt [18] or more detailed in Basseville and Nikiforov [2].

We distinguish in principle the *sequential* and the *retrospective* change-point problem. In sequential problems we make decisions on the appearance of change-points simultaneously with the sequential process of data collection, i.e., we have to examine with every new observation whether a change occurs. However, in retrospective problems the entire data set is already available. We refer the reader to Brodsky [4], Brodsky and Darkhovsky [5] or Csörgö and Horváth [8] to get an overview on both approaches.

This work is concerned with the retrospective point of view. The mathematical formulation of these problems goes back to the 1950s, see for instance Page [22, 23]. In the past researchers have considered many different parametric and non-parametric models to detect and estimate change-points. Some in the literature discussed methods are mentioned here. The basic approach of using the maximum likelihood method can be found for instance in Hinkley [21]. Ferger [12, 13] proposed an estimator of a single change-point determined by weighted U-statistic-type processes and investigated the convergence in distribution and the almost sure convergence of such estimators. Döring [10, 11] generalized this approach to an arbitrary, but known, number of change-points. Another well-known method is the least squares method. A weighted least squares estimator to estimate a single change-point was introduced by Ferger [15]. Simultaneously, he established the connection to the maximum-likelihood estimator and the estimator determined by weighted U-statistic-type processes.

This thesis is intended to applying the least squares method to estimate two change-points, which is a generalization of the approach of Ferger [15]. The following non-parametric framework is handled. Let $X_i = X_{i,n}$, $1 \leq i \leq n$, $n \in \mathbb{N}$, be real-valued random variables and Q_1, Q_2 and Q_3 be unknown distributions. Assume that there exist unknown $0 < \theta_1 < \theta_2 < 1$ such that

$$X_i \sim \begin{cases} Q_1, & 1 \leq i \leq \lfloor n\theta_1 \rfloor, \\ Q_2, & \lfloor n\theta_1 \rfloor + 1 \leq i \leq \lfloor n\theta_2 \rfloor, \\ Q_3, & \lfloor n\theta_2 \rfloor + 1 \leq i \leq n \end{cases}$$

for all $n \in \mathbb{N}$. To obtain a well-defined model we further suppose that the expectations of random variables from adjacent segments are different. The parameter of interest is the so-called *multiple change-point* (θ_1, θ_2), which is to be estimated by the least squares method. The main focus of attention lies on the discussion of the consistency as well as the investigation of the convergence in distribution of such least squares estimators in the given multiple change-point model. For this purpose, we apply a similar approach used in Döring [10, 11] and Ferger [14, 15]. At the end of this work we give an outlook on the asymptotic properties of least squares estimators in the case of an arbitrary, but known, number of change-points.

We can often find tests to detect change-points in the literature where under the null hypothesis (no change) the distribution of some test statistic is examined to construct critical regions. However, note that we look at the alternative hypothesis (existence of change-points) where the investigation of the distribution of such estimators becomes more complex than under the null hypothesis. Indeed, if the number of change-points is unknown for any reason (for example in some practical applications), one has to previously detect the occurrence of multiple change-points with such tests (see for instance Brodsky [4] or Csörgö and Horváth [8]) or determine the number of change-points based on content-related considerations.

This work is organized as follows. We start with the accurate formulation of the multiple change-point model. Then we briefly sketch the essential steps to get the main results of the work. Chapter 2 provides the relevant mathematical tools for our purpose. For simplicity, Chapter 3 deals with the case of known expectations. This chapter is intended to present the fundamental recipe to estimate change-points and conclude asymptotic claims for the estimator. Based on the least squares estimator of the *moments of change* $(\lfloor n\theta_1 \rfloor, \lfloor n\theta_2 \rfloor)$ we construct the estimator of (θ_1, θ_2). Under different moment conditions we show weak and strong consistency. Furthermore, we investigate convergence in distribution and identify the limit variable, which is used to derive a confidence region for $(\lfloor n\theta_1 \rfloor, \lfloor n\theta_2 \rfloor)$. From Chapter 4 on, the expectations are assumed to be unknown. Here, we state and prove the main results of this work. Section 4.1 contains the simultaneous least squares estimation of the change-points and the expectations and discusses the consistency of the resulting estimators. Since convergence in distribution of the estimator of the multiple change-point is hard to show, we introduce another least squares estimator in Section 4.2. In the estimator from Chapter 3 the known expectations are replaced by their estimators. We thus obtain an estimator with the same structure as in Chapter 3. Consequently, we can proceed on a similar way to Chapter 3 but some proofs are more technical. We treat consistency and convergence in distribution. Based on these results, we derive a confidence region for the parameter $(\lfloor n\theta_1 \rfloor, \lfloor n\theta_2 \rfloor)$ in the case of unknown expectations. Chapter 5 indicates the performance of all estimators and several relations by a simulation study. The last chapter gives an outlook where we discuss our strong conjecture that all of the results of Chapter 4 can be generalized to an arbitrary, but known, number of change-points. Moreover, we specify ideas for further work on this field.

1.2 Model

This section presents the multiple change-point model.

Let $(X_{j,n})_{\substack{n\in\mathbb{N} \\ 1\le j\le n}}$ be a triangular array of random variables defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with values in the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Each row of the triangular array consists of independent random variables, i.e., $X_{1,n}, ..., X_{n,n}$ are independent for every $n \in \mathbb{N}$. Let us denote by Θ and Δ_n, $n \in \mathbb{N}$, the sets

$$\Theta := \left\{(s,t) \in \mathbb{R}^2 \middle| 0 < s < t < 1\right\} \quad \text{and} \quad \Delta_n := \left\{(k,l) \in \mathbb{N}^2 \middle| 1 \le k < l \le n-1\right\}.$$

We assume that there exists a vector $\boldsymbol{\theta} := (\theta_1, \theta_2) \in \Theta$ such that for all $n \in \mathbb{N}$

$$X_{i,n} \sim \begin{cases} Q_1, & 1 \le i \le \lfloor n\theta_1 \rfloor, \\ Q_2, & \lfloor n\theta_1 \rfloor + 1 \le i \le \lfloor n\theta_2 \rfloor, \\ Q_3, & \lfloor n\theta_2 \rfloor + 1 \le i \le n, \end{cases} \tag{1.1}$$

where Q_1, Q_2, Q_3 are arbitrary, but unknown, distributions. It is of interest to estimate the unknown so-called *multiple change-point* $\boldsymbol{\theta} = (\theta_1, \theta_2)$. The quantities $\theta_1, \theta_2 - \theta_1, 1 - \theta_2$ give ratios of how many observations belong to each segment in the statistical population. We call the also unknown parameters $\boldsymbol{\tau}_n := (\tau_n, \sigma_n) := (\lfloor n\theta_1 \rfloor, \lfloor n\theta_2 \rfloor) \in \Delta_n$, $n \in \mathbb{N}$, *moments of change*. These state the last indices before the first and second change of distribution. In order to get asymptotic results it is necessary to consider such a triangular array. For increasing $n \in \mathbb{N}$, the triangular array guarantees that more and more random variables arise from each distribution. Roughly speaking, therefore it is possible to estimate the true parameter $\boldsymbol{\theta} = (\theta_1, \theta_2)$, at all.

Furthermore, we suppose that the expectations $\boldsymbol{\alpha} := (\alpha, \beta, \gamma)$ defined by

$$\alpha := \int_{\mathbb{R}} x \; Q_1(\mathrm{d}x), \quad \beta := \int_{\mathbb{R}} x \; Q_2(\mathrm{d}x) \quad \text{and} \quad \gamma := \int_{\mathbb{R}} x \; Q_3(\mathrm{d}x)$$

exist, are finite and satisfy

$$\alpha \neq \beta \quad \text{and} \quad \beta \neq \gamma. \tag{1.2}$$

The last condition means that the distributions of adjacent segments differ in their first moments. It ensures that our multiple change-point model is well-defined.

For illustration, Figure 1.1 depicts the entire model where realizations of some random variables are represented by dots.

In the whole work it is crucial to consider moment estimates to conclude asymptotic results for all estimators. For this purpose, let

$$M_p := \max\left\{ \int_{\mathbb{R}} |x|^p \; Q_1(\mathrm{d}x), \int_{\mathbb{R}} |x|^p \; Q_2(\mathrm{d}x), \int_{\mathbb{R}} |x|^p \; Q_3(\mathrm{d}x) \right\}$$

denote the maximum of the p-th absolute moments, $p \in [1, \infty)$. Unless otherwise stated we assume that $M_1 < \infty$.

To simplify notation, we write $X_1, ..., X_n$ instead of $X_{1,n}, ..., X_{n,n}$, $n \in \mathbb{N}$, for the n-th row of the triangular array.

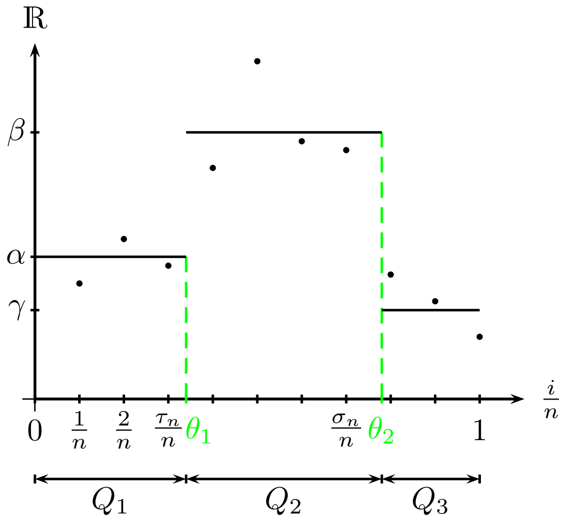

Figure 1.1: Multiple change-point model with $(\theta_1, \theta_2) = (0.34, 0.78)$ in the case $\gamma < \alpha < \beta$ for $n = 10$ (left) and $n = 25$ (right) observations.

Finally, let us introduce the argmin notation. For an arbitrary set A and a mapping $f : A \to \mathbb{R}$ we denote by

$$\operatorname{Argmin}(f) := \{k \in A | f(k) \leq f(l) \text{ for all } l \in A\} \qquad \text{and}$$
$$\operatorname{Argmax}(f) := \{k \in A | f(k) \geq f(l) \text{ for all } l \in A\}$$

the set of all minimizing and maximizing points of f, respectively. If we choose a specific minimizing or maximizing point of f, then we write $\operatorname{argmin}_{k\in A} f(k)$ or $\operatorname{argmax}_{k\in A} f(k)$.

1.3 Essential results

This section gives a brief exposition of the agenda in this work and summarizes the main results without proofs. From now on, the fact that some random functions depend on the sample $X_1, ..., X_n$ is omitted.

At first, the expectations $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)$ in our model are assumed to be known. To obtain an estimator of the multiple change-point $\boldsymbol{\theta} = (\theta_1, \theta_2)$, we have to estimate the moments of change $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ previously. For this purpose, by the least squares method, we are interested in minimizing the random criterion function

$$\bar{S}_n(k, l) := \sum_{i=1}^{k}(X_i - \alpha)^2 + \sum_{i=k+1}^{l}(X_i - \beta)^2 + \sum_{i=l+1}^{n}(X_i - \gamma)^2, \qquad (k, l) \in \Delta_n.$$

Since it is possible that $\bar{S}_n$ has several minimizers, we choose an arbitrary minimizer

$$\bar{\boldsymbol{\tau}}_n := \underset{(k,l)\in\Delta_n}{\operatorname{argmin}} \bar{S}_n(k, l)$$

as the estimator of $\boldsymbol{\tau}_n$. By properties of the floor function it is clear that $\frac{1}{n}\boldsymbol{\tau}_n \xrightarrow[n\to\infty]{} \boldsymbol{\theta}$. Therefore, a reasonable estimator of $\boldsymbol{\theta}$ is given by $\bar{\boldsymbol{\theta}}_n := \frac{1}{n}\bar{\boldsymbol{\tau}}_n$.

The first aim is to prove consistency of $\bar{\boldsymbol{\theta}}_n$. To this end, Theorem 2.1 provides conditions to show almost sure convergence and convergence in probability of random minimizers. To stay in the context of this theorem, $\bar{\boldsymbol{\theta}}_n$ must be represented in another form:

$$\bar{\boldsymbol{\theta}}_n = \underset{(s,t)\in\Theta_n}{\operatorname{argmin}} \bar{\rho}_n(s, t)$$

with some technical domain Θ_n and

$$\bar{\rho}_n(s, t) := \frac{1}{n}\bar{M}_n(\lfloor ns\rfloor, \lfloor nt\rfloor), \qquad (s, t) \in \Theta_n,$$

where $\bar{M}_n$ is some criterion function, which has the same minimizers as $\bar{S}_n$. We first have to show uniform convergence in probability and almost sure uniform convergence of $\bar{\rho}_n$ to a deterministic function ρ, i.e.,

$$\sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s, t) - \rho(s, t)| \xrightarrow[n\to\infty]{\mathbb{P}\ (a.s.)} 0.$$

Moreover, $\boldsymbol{\theta} = (\theta_1, \theta_2)$ must be a minimizing point of ρ, which is additionally well-separated, by definition,

$$\inf\{\rho(s,t) : \|(\theta_1,\theta_2) - (s,t)\| \geq \varepsilon,\ (s,t) \in \Theta\} - \rho(\theta_1,\theta_2) > 0$$

for all $\varepsilon > 0$, where $\|\cdot\|$ denotes the maximum norm. Applying Theorem 2.1 leads to the weak and strong consistency of $\bar{\boldsymbol{\theta}}_n$ under different moment conditions.

Theorem. *If $M_2 < \infty$, then*

$$\bar{\boldsymbol{\theta}}_n \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\theta}.$$

Theorem. *Suppose there is some $p \in (2,\infty)$ such that $M_p < \infty$. Then*

$$\bar{\boldsymbol{\theta}}_n \xrightarrow[n\to\infty]{a.s.} \boldsymbol{\theta}.$$

Our next objective is to investigate the convergence in distribution of $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$. For this purpose, let $(\xi_{i,r})_{i\in\mathbb{N}}$, $r \in \{1,2,3\}$, be three independent sequences, which for each r consist of independent and identically distributed random variables with common distribution Q_r. Write

$$\Gamma(k,l) := \Gamma_1(k) + \Gamma_2(l), \qquad (k,l) \in \mathbb{Z}^2,$$

where

$$\Gamma_1(k) := \begin{cases} 2(\beta-\alpha)\sum_{i=1}^{k}(\xi_{i,2}-\beta) + k(\alpha-\beta)^2, & k \geq 0, \\ -2(\beta-\alpha)\sum_{i=1}^{-k}(\xi_{i,1}-\alpha) - k(\alpha-\beta)^2, & k < 0 \end{cases} \qquad \text{and}$$

$$\Gamma_2(l) := \begin{cases} 2(\gamma-\beta)\sum_{i=1}^{l}(\xi_{i,3}-\gamma) + l(\beta-\gamma)^2, & l \geq 0, \\ -2(\gamma-\beta)\sum_{i=1}^{-l}(\xi_{i,2}-\beta) - l(\beta-\gamma)^2, & l < 0. \end{cases}$$

Note that the process Γ is a sum of random walks. The main idea to examine the convergence in distribution of $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$ is to introduce the so-called *rescaled process*

$$\bar{\Gamma}_n(k,l) := \bar{M}_n(\tau_n + k, \sigma_n + l) - \bar{M}_n(\tau_n, \sigma_n), \qquad (k,l) \in H_n,$$

where H_n is some technical domain. Since $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$ is a minimizer of $\bar{\Gamma}_n$ for each $n \in \mathbb{N}$, we are able to apply Theorem 2.3, which gives conditions to show convergence in distribution of random minimizers. The convergence of the finite-dimensional distributions of $\bar{\Gamma}_n$ to Γ, i.e., for all $m \in \mathbb{N}$ and $(k_1, l_1), ..., (k_m, l_m) \in \mathbb{Z}^2$ it holds

$$\big(\bar{\Gamma}_n(k_1, l_1), ..., \bar{\Gamma}_n(k_m, l_m)\big) \xrightarrow[n\to\infty]{\mathcal{L}} (\Gamma(k_1, l_1), ..., \Gamma(k_m, l_m)),$$

is the first assumption to check. Furthermore, by the Hájek-Rényi Inequality (see Lemma 2.8), we get an estimate of the error probability

$$\mathbb{P}[x \leq \|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta] \leq Cx^{-1},$$

where $C > 0$ is a constant, $\delta > 0$ sufficiently small, $n \in \mathbb{N}$ sufficiently large and $x \geq 2$. Combining this with the weak consistency of $\bar{\boldsymbol{\theta}}_n$ yields stochastic boundedness of $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$ (second assumption of Theorem 2.3), i.e.,

$$\lim_{x\to\infty} \limsup_{n\to\infty} \mathbb{P}[\|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \geq x] = 0.$$

If the distributions Q_1, Q_2 and Q_3 are continuous, the limit process Γ has almost surely exactly one minimizer. By application of Theorem 2.3, we obtain convergence in distribution of $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$ to the minimizer of a sum of random walks.

Theorem. *If $M_2 < \infty$, then*

$$\limsup_{n\to\infty} \mathbb{P}[\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n \in F] \leq \mathbb{P}[\mathrm{Argmin}(\Gamma) \cap F \neq \emptyset] \qquad \textit{for all } F \subseteq \mathbb{Z}^2.$$

In addition, if Q_1, Q_2 and Q_3 are continuous, then $\mathrm{Argmin}(\Gamma) = \{\boldsymbol{T}\}$ *almost surely and*

$$\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n \xrightarrow[n\to\infty]{\mathcal{L}} \boldsymbol{T} \quad \textit{in} \quad \mathbb{Z}^2.$$

Based on the last result and the Continuous Mapping Theorem for convergence in distribution, we derive an asymptotic confidence region for the parameter $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$. For this purpose, let $F^{-1}_{\|\boldsymbol{T}\|}(\vartheta)$, $\vartheta \in (0,1)$, denote the ϑ-quantile of the distribution function $F_{\|\boldsymbol{T}\|}$ of $\|\boldsymbol{T}\|$.

Theorem. *Suppose that $M_2 < \infty$. Let Q_1, Q_2 and Q_3 be continuous distributions and $\vartheta \in (0,1)$. For each $n \in \mathbb{N}$, the random interval*

$$I_n(\vartheta) := \left[\bar{\tau}_n - F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta), \bar{\tau}_n + F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta)\right] \times \left[\bar{\sigma}_n - F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta), \bar{\sigma}_n + F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta)\right]$$

is an asymptotic confidence region for $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ at level $1-\vartheta$.

We now proceed with the asymptotic behavior of least squares estimators in our multiple change-point model if the expectations $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)$ are assumed to be unknown. For abbreviation, we use $\bar{X}_{u,v} := \frac{1}{v-u}\sum_{i=u+1}^{v} X_i$ for $u, v \in \mathbb{N}_0$ with $u < v \leq n$. To simultaneously estimate the moments of change $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ and the expectations $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)$ by the least squares method, we minimize the criterion function

$$S_n(k,l,a,b,c) := \sum_{i=1}^{k}(X_i - a)^2 + \sum_{i=k+1}^{l}(X_i - b)^2 + \sum_{i=l+1}^{n}(X_i - c)^2, \quad (k,l) \in \Delta_n,$$
$$(a,b,c) \in \mathbb{R}^3.$$

To do this, set

$$\hat{M}_n(k,l) := k\bar{X}_{0,k}^2 + (l-k)\bar{X}_{k,l}^2 + (n-l)\bar{X}_{l,n}^2, \quad (k,l) \in \Delta_n,$$

and choose an arbitrary maximizing point

$$\hat{\boldsymbol{\tau}}_n := (\hat{\tau}_n, \hat{\sigma}_n) := \underset{(k,l)\in\Delta_n}{\operatorname{argmax}}\, \hat{M}_n(k,l).$$

We can show that $(\hat{\boldsymbol{\tau}}_n, \hat{\boldsymbol{\alpha}}_n)$ is a minimizer of S_n, where

$$\hat{\boldsymbol{\alpha}}_n := \left(\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n\right) := \left(\bar{X}_{0,\hat{\tau}_n}, \bar{X}_{\hat{\tau}_n,\hat{\sigma}_n}, \bar{X}_{\hat{\sigma}_n,n}\right).$$

So, we have found an estimator of $(\boldsymbol{\tau}_n, \boldsymbol{\alpha})$. Likewise as before, $\hat{\boldsymbol{\theta}}_n := \frac{1}{n}\hat{\boldsymbol{\tau}}_n$ is a reasonable estimator of the multiple change-point $\boldsymbol{\theta} = (\theta_1, \theta_2)$.
The strong consistency of $\hat{\boldsymbol{\theta}}_n$ was shown by Albrecht [1].

Theorem. *Suppose there is some $p \in (4, \infty)$ such that $M_p < \infty$. Then*

$$\hat{\boldsymbol{\theta}}_n \xrightarrow[n\to\infty]{a.s.} \boldsymbol{\theta}.$$

Based on the proof of the previous theorem, we conclude the weak consistency under a weaker moment condition.

Theorem. *Suppose there is some $p \in (2, \infty)$ such that $M_p < \infty$. Then*

$$\hat{\boldsymbol{\theta}}_n \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\theta}.$$

To get further results, we prove the stochastic boundedness of $\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$. By applications of Markov's Inequality (compare Lemma 2.4) and some maximal inequalities like Chow and Doob (see Lemmas 2.9 and 2.10), we obtain an estimate of the error probability. This and the weak consistency of $\hat{\boldsymbol{\theta}}_n$ lead to the stochastic boundedness of $\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$. Hence, we are able to show weak consistency of the estimator of expectations.

Theorem. *If $M_4 < \infty$, then*

$$\hat{\boldsymbol{\alpha}}_n \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\alpha}.$$

Likewise as before, the investigation of convergence in distribution of $\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$ is based on the introduction of the rescaled process with respect to $\hat{M}_n$. However, the calculation of the rescaled process is hard to handle. Therefore, we introduce another estimator of the moments of change $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ and examine the asymptotic behavior of the new estimator. In the criterion function $\bar{S}_n$ the known expectations $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)$ are replaced by their associated estimators $\hat{\boldsymbol{\alpha}}_n = \left(\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n\right)$. Write

$$S_n^*(k,l) := \sum_{i=1}^{k} (X_i - \hat{\alpha}_n)^2 + \sum_{i=k+1}^{l} \left(X_i - \hat{\beta}_n\right)^2 + \sum_{i=l+1}^{n} (X_i - \hat{\gamma}_n)^2, \qquad (k,l) \in \Delta_n.$$

It follows that another least squares estimator of $(\boldsymbol{\tau}_n, \boldsymbol{\alpha})$ is given by $(\boldsymbol{\tau}_n^*, \hat{\boldsymbol{\alpha}}_n)$, where

$$\boldsymbol{\tau}_n^* := \operatorname*{argmin}_{(k,l)\in\Delta_n} S_n^*(k,l)$$

is an arbitrary minimizer of S_n^*. The multiple change-point $\boldsymbol{\theta} = (\theta_1, \theta_2)$ is estimated by $\boldsymbol{\theta}_n^* := \frac{1}{n}\boldsymbol{\tau}_n^*$. We can now proceed on a very similar way to the case of known expectations, because both criterion functions S_n^* and $\bar{S}_n$ feature a very similar structure. Observe that many proofs become more technical, because S_n^* involves the estimators of expectations. To prove weak consistency of $\boldsymbol{\theta}_n^*$ and convergence in distribution of $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$, the main results of our work, we use the weak consistency of $\hat{\boldsymbol{\alpha}}_n$ permanently. Hence, these results can be proved only for the same moment condition as in the previous theorem.

Theorem. *If $M_4 < \infty$, then*

$$\boldsymbol{\theta}_n^* \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\theta}.$$

Observe below that $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$ converges in distribution to the same limit process as $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$ in the case of known expectations.

Theorem. *If $M_4 < \infty$, then*

$$\limsup_{n\to\infty} \mathbb{P}[\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n \in F] \leq \mathbb{P}[\operatorname{Argmin}(\Gamma) \cap F \neq \emptyset] \qquad \text{for all } F \subseteq \mathbb{Z}^2.$$

In addition, if Q_1, Q_2 and Q_3 are continuous, then $\operatorname{Argmin}(\Gamma) = \{\boldsymbol{T}\}$ *almost surely and*

$$\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n \xrightarrow[n\to\infty]{\mathcal{L}} \boldsymbol{T} \qquad \text{in} \quad \mathbb{Z}^2.$$

By the same steps as before, we get an asymptotic confidence region for $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ in the case of unknown expectations.

Theorem. *Suppose that $M_4 < \infty$. Let Q_1, Q_2 and Q_3 be continuous distributions and $\vartheta \in (0,1)$. For each $n \in \mathbb{N}$, the random interval*

$$I_n(\vartheta) := \left[\tau_n^* - F_{\|\boldsymbol{T}\|}^{-1}(1-\vartheta), \tau_n^* + F_{\|\boldsymbol{T}\|}^{-1}(1-\vartheta)\right] \times \left[\sigma_n^* - F_{\|\boldsymbol{T}\|}^{-1}(1-\vartheta), \sigma_n^* + F_{\|\boldsymbol{T}\|}^{-1}(1-\vartheta)\right]$$

is an asymptotic confidence region for $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ at level $1-\vartheta$.

Chapter 2

Fundamentals

This chapter provides the relevant tools to prove the results of this work. In fact, we gather theorems for convergence of random minimizing points and some inequalities.

2.1 Continuous Mapping Theorems for the argmin functional

This section deals with the convergence of random minimizing points. The following theorem, which is adapted from Ferger [16], gives criteria to prove almost sure convergence and convergence in probability. For further information about the multivariate Skorokhod space we refer the reader to Döring [9] or Ferger [16] and the references given there.

Theorem 2.1. *Let* $O \subseteq \mathbb{R}^q$, $q \in \mathbb{N}$, *be an open set and let* Z, Z_n, $n \in \mathbb{N}$, *be stochastic processes defined on* $(\Omega, \mathcal{A}, \mathbb{P})$ *with trajectories in the multivariate Skorokhod space* $D(O)$. *Let* $(T_n)_{n\in\mathbb{N}} \subseteq O$ *be a sequence such that* $T_n \subseteq T_{n+1}$ *for every* $n \in \mathbb{N}$ *with* $\bigcup_{n\in\mathbb{N}} T_n = O$. *Furthermore, let* $\boldsymbol{\sigma}_n$ *be a random variable with* $\boldsymbol{\sigma}_n \in \operatorname{Argmin}(Z_n)$ *for each* $n \in \mathbb{N}$. *If* *(i)* $\sup_{\mathbf{t}\in T_n} |Z_n(\mathbf{t}) - Z(\mathbf{t})| \xrightarrow[n\to\infty]{a.s.\ (\mathbb{P})} 0$ *and*

(ii) *Z has almost surely a minimizing point $\boldsymbol{\sigma}$ satisfying*

$$\inf\{Z(\mathbf{t}) : \|\boldsymbol{\sigma} - \mathbf{t}\| \geq \varepsilon,\ \mathbf{t} \in O\} > Z(\boldsymbol{\sigma}) \tag{2.1}$$

for all $\varepsilon > 0$, then

$$\boldsymbol{\sigma}_n \xrightarrow[n\to\infty]{a.s.\ (\mathbb{P})} \boldsymbol{\sigma}.$$

Proof. Ferger [16, p. 28, Theorem 3.3 and Remark 3.1] shows the assertion for infimizing points under the assumption that we have a sequence $(T_n)_{n\in\mathbb{N}}$ of open sets with $\liminf_{n\to\infty} T_n = O \subseteq \mathbb{R}^q$. This proof shows that under the new assumption to $(T_n)_{n\in\mathbb{N}}$ the assertion is still preserved. Note that every minimizing point is an infimizing point. □

We call a minimizing point satisfying condition (2.1) *well-seperated*. The previous theorem and the next remark help us to prove weak and strong consistency of several estimators of the multiple change-point.

Remark 2.2. To formulate Theorem 2.1 for maximizing points, replace „Argmin“ by „Argmax“, „minimizing“ by „maximizing“ and condition (2.1) by

$$\sup\{Z(\mathbf{t}) : \|\boldsymbol{\sigma} - \mathbf{t}\| \geq \varepsilon,\ \mathbf{t} \in O\} < Z(\boldsymbol{\sigma}) \quad \text{for all } \varepsilon > 0. \tag{2.2}$$

The following theorem states under which conditions the convergence in distribution of minimizers of discrete stochastic processes is ensured.

Theorem 2.3. *Let Z, Z_n, $n \in \mathbb{N}$, be stochastic processes indexed by $\mathbb{Z}^q$, $q \in \mathbb{N}$, and let* $\mathrm{Argmin}(Z)$ *and* $\mathrm{Argmin}(Z_n)$ *be non-empty. Furthermore, let $\boldsymbol{\sigma}_n$ be a random variable with* $\boldsymbol{\sigma}_n \in \mathrm{Argmin}(Z_n)$ *for each $n \in \mathbb{N}$. If* *(i) for all* $m \in \mathbb{N}$ *and* $\mathbf{k}_1, ..., \mathbf{k}_m \in \mathbb{Z}^q$ *it holds*

$$(Z_n(\mathbf{k}_1), ..., Z_n(\mathbf{k}_m)) \xrightarrow[n\to\infty]{\mathcal{L}} (Z(\mathbf{k}_1), ..., Z(\mathbf{k}_m)) \qquad \textit{and}$$

(ii) $\lim_{d\to\infty} \limsup_{n\to\infty} \mathbb{P}[\|\boldsymbol{\sigma}_n\| > d] = 0$*,*

then

$$\limsup_{n\to\infty} \mathbb{P}[\boldsymbol{\sigma}_n \in F] \leq \mathbb{P}[\mathrm{Argmin}(Z) \cap F \neq \emptyset]$$

for all $F \subseteq \mathbb{Z}^q$*. If in addition*

(iii) $\mathrm{Argmin}(Z) = \{\boldsymbol{\sigma}\}$ *almost surely,*

then

$$\boldsymbol{\sigma}_n \xrightarrow[n\to\infty]{\mathcal{L}} \boldsymbol{\sigma} \qquad \textit{in } \mathbb{Z}^q.$$

Proof. The proof can be found in Ferger [17]. □

2.2 Inequalities

In many proofs of this work it is crucial to estimate probabilities or moments. For the convenience of the reader, some in the literature well-known inequalities are recalled without proofs. After this we give some moment estimates for sums of observations from the model.

Lemma 2.4 (Markov Inequality)**.** *Let* Z *be a random variable and* $r \in (0, \infty)$*. Then for all* $\varepsilon > 0$

$$\mathbb{P}\left[|Z| \geq \varepsilon\right] \leq \varepsilon^{-r}\mathbb{E}[|Z|^r].$$

Lemma 2.5. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $B \in \mathcal{A}$ with $\mathbb{P}[B] > 0$. Let Z be a random variable. Then for all $\varepsilon > 0$*

$$\mathbb{P}[|Z| \geq \varepsilon | B] \leq \varepsilon^{-2} \mathbb{P}[B]^{-1} \mathbb{E}\left[\mathbb{1}_B Z^2\right].$$

Lemma 2.6 (Chebyshev Inequality)**.** *Let $Z_1, ..., Z_n$, $n \in \mathbb{N}$, be pairwise uncorrelated and centered random variables. Then for all $\varepsilon > 0$*

$$\mathbb{P}\left[\left|\sum_{i=1}^{n} Z_i\right| \geq \varepsilon\right] \leq \varepsilon^{-2} \sum_{i=1}^{n} \mathbb{V}[Z_i].$$

Lemma 2.7 (First Kolmogorov Inequality)**.** *Let $Z_1, ..., Z_n$, $n \in \mathbb{N}$, be independent and centered random variables. Then for all $\varepsilon > 0$*

$$\mathbb{P}\left[\max_{1 \leq k \leq n} \left|\sum_{i=1}^{k} Z_i\right| \geq \varepsilon\right] \leq \varepsilon^{-2} \sum_{i=1}^{n} \mathbb{V}[Z_i].$$

Lemma 2.8 (Hájek-Rényi Inequality)**.** *Let $Z_1, ..., Z_n$ be independent and centered random variables with finite variances. Let $c_1, ..., c_n$ be a non-increasing sequence of positive numbers. Then for any $\varepsilon > 0$ and for any $m \in \mathbb{N}$ with $1 \leq m \leq n$*

$$\mathbb{P}\left[\max_{m \leq k \leq n} c_k \left|\sum_{i=1}^{k} Z_i\right| \geq \varepsilon\right] \leq \varepsilon^{-2} \left(c_m^2 \sum_{k=1}^{m} \mathbb{V}[Z_k] + \sum_{k=m+1}^{n} c_k^2 \mathbb{V}[Z_k]\right).$$

Lemma 2.9 (Chow Inequality)**.** *Let $(S_k)_{k \in \mathbb{N}}$ be a submartingale with respect to the filtration $(\mathcal{F}_k)_{k \in \mathbb{N}}$. Then for each $\varepsilon > 0$ and $a_{m+1} \geq a_{m+2} \geq ... \geq a_n$, $m, n \in \mathbb{N}$ with $m < n$*

$$\mathbb{P}\left[\max_{m+1 \leq k \leq n} a_k S_k \geq \varepsilon\right] \leq \varepsilon^{-1} \left(a_n \mathbb{E}\left[S_n^+\right] + \sum_{k=m+1}^{n-1} (a_k - a_{k+1}) \mathbb{E}\left[S_k^+\right]\right).$$

Lemma 2.10 (Doob Inequalities)**.** *Let $(S_k)_{k \in \mathbb{N}}$ be a non-negative submartingale with respect to the filtration $(\mathcal{F}_k)_{k \in \mathbb{N}}$.*

(i) Let $1 \leq m \leq n$ and $\varepsilon > 0$. Then

$$\mathbb{P}\left[\max_{m \leq k \leq n} S_k \geq \varepsilon\right] \leq \varepsilon^{-1} \mathbb{E}[|S_n|].$$

(ii) Let $1 \leq m \leq n$ and $r \in (1, \infty)$. Then

$$\mathbb{E}\left[\max_{m \leq k \leq n} S_k^r\right] \leq \left(\frac{r}{r-1}\right)^r \mathbb{E}\left[S_n^r\right].$$

Lemma 2.11 (Cauchy–Schwarz Inequalities). *(i) For any $a_1, ..., a_n, b_1, ..., b_n \in \mathbb{R}$ it holds*

$$\sum_{i=1}^{n} |a_i b_i| \leq \left(\sum_{i=1}^{n} a_i^2\right)^{1/2} \left(\sum_{i=1}^{n} b_i^2\right)^{1/2}.$$

(ii) For any random variables X and Y it holds

$$\mathbb{E}[|XY|] \leq \mathbb{E}\left[X^2\right]^{1/2} \mathbb{E}\left[Y^2\right]^{1/2}.$$

Lemma 2.12 (Jensen Inequality). *Let $g : \mathbb{R} \longrightarrow \mathbb{R}$ be a convex function. Suppose that expectations of Z and $g(Z)$ exist. Then*

$$g(\mathbb{E}[Z]) \leq \mathbb{E}[g(Z)].$$

Lemma 2.13 (c_r-Inequality). *Let $Z_1, ..., Z_n$ be random variables and $r \in (0, \infty)$. Then*

$$\mathbb{E}\left[\left|\sum_{i=1}^{n} Z_i\right|^r\right] \leq c_r \sum_{i=1}^{n} \mathbb{E}\left[|Z_i|^r\right] \quad \text{with} \quad c_r = \begin{cases} 1, & r \leq 1, \\ n^{r-1}, & r > 1. \end{cases}$$

The next lemma ensures that we can apply some maximal inequalities to a sum of independent and centered random variables.

Lemma 2.14. *Let $(Z_i)_{i \in \mathbb{N}}$ be a sequence of independent, centered and p-fold integrable random variables for some $p \in [1, \infty)$. Then for each $m \in \mathbb{N}$ the process $\left(\left|\sum_{i=m}^{k} Z_i\right|^p\right)_{k \in \mathbb{N}_m}$ is a non-negative submartingale with respect to the filtration $(\mathcal{F}_k)_{k \in \mathbb{N}_m}$ with $\mathcal{F}_k := \sigma(Z_m, ..., Z_k)$.*

Proof. The proof can be found for instance in Albrecht [1, p. 15, Lemma 2.6] in the case $m = 1$. In the same manner we see the claim for each $m \in \mathbb{N}$. □

The next proposition helps us to find further moment estimates.

Lemma 2.15. *Let $(Z_i)_{i\in\mathbb{N}}$ be a sequence of independent, centered and p-fold integrable random variables for some $p \in [2,\infty)$. For each $n \in \mathbb{N}$ and $p \geq 2$ there exists a positive constant B_p depending only on p such that*

$$\mathbb{E}\left[\left|\sum_{i=1}^{n} Z_i\right|^p\right] \leq B_p n^{p/2-1} \sum_{i=1}^{n} \mathbb{E}\left[|Z_i|^p\right].$$

Proof. Fix $n \in \mathbb{N}$. By the Marcinkiewicz-Zygmund Inequality (see for instance Chow and Teicher [7, p. 386, Theorem 2]), there exists a positive constant B_p depending only on $p \in [1,\infty)$ such that

$$\mathbb{E}\left[\left|\sum_{i=1}^{n} Z_i\right|^p\right] \leq B_p \mathbb{E}\left[\left(\sum_{i=1}^{n} Z_i^2\right)^{p/2}\right]. \tag{2.3}$$

We see at once that the assertion of this lemma is true for $p = 2$. For $p > 2$ we apply the Hölder Inequality (see for instance Heuser [20, p. 347, Inequality 59.2]). For this purpose, we set $\tilde{p} := \frac{p}{2} > 1$. To hold $\frac{1}{\tilde{p}} + \frac{1}{\tilde{q}} = 1$, we obtain $\tilde{q} = \frac{p}{p-2}$. Hölder's Inequality yields

$$\begin{aligned}\left(\sum_{i=1}^{n} Z_i^2\right)^{p/2} = \left(\sum_{i=1}^{n} \left|Z_i^2 \cdot 1\right|\right)^{p/2} &\leq \left(\left(\sum_{i=1}^{n} |Z_i^2|^{\tilde{p}}\right)^{1/\tilde{p}} \left(\sum_{i=1}^{n} |1|^{\tilde{q}}\right)^{1/\tilde{q}}\right)^{p/2} \\ &= n^{p/2-1} \sum_{i=1}^{n} |Z_i|^p.\end{aligned}$$

By (2.3), we have

$$\mathbb{E}\left[\left|\sum_{i=1}^{n} Z_i\right|^p\right] \leq B_p n^{p/2-1} \sum_{i=1}^{n} \mathbb{E}\left[|Z_i|^p\right]. \qquad \square$$

Moreover, in this thesis we frequently use the following moment estimates for sums of centered observations (from our model).

Corollary 2.16. *Suppose there is some $p \in [2,\infty)$ such that $M_p < \infty$. Let $u, v \in \mathbb{N}_0$ with $u < v \leq n$, $n \in \mathbb{N}$. For each $p \geq 2$ there exists a positive constant B_p depending only on p such that*

$$\mathbb{E}\left[\left|\sum_{i=u+1}^{v} (X_i - \mathbb{E}[X_i])\right|^p\right] \leq 2^p B_p M_p (v-u)^{p/2}.$$

Proof. Fix $u, v \in \mathbb{N}_0$ and $n \in \mathbb{N}$ with $u < v \leq n$. By Lemma 2.15, there exists a positive constant B_p such that

$$\begin{aligned}\mathbb{E}\left[\left|\sum_{i=u+1}^{v}(X_i - \mathbb{E}[X_i])\right|^p\right] &= \mathbb{E}\left[\left|\sum_{i=1}^{v-u}(X_{u+i} - \mathbb{E}[X_{u+i}])\right|^p\right] \\ &\leq B_p(v-u)^{p/2-1}\sum_{i=1}^{v-u}\mathbb{E}\left[|X_{u+i} - \mathbb{E}[X_{u+i}]|^p\right]. \end{aligned} \tag{2.4}$$

Furthermore, we can conclude that

$$\begin{aligned}\mathbb{E}[|X_{u+i} - \mathbb{E}[X_{u+i}]|^p] &\leq 2^{p-1}(\mathbb{E}[|X_{u+i}|^p] + |\mathbb{E}[X_{u+i}]|^p) && \text{by } c_r\text{-In.} \\ &\leq 2^p\mathbb{E}[|X_{u+i}|^p] && \text{by Jensen In.} \\ &\leq 2^p M_p\end{aligned}$$

for all $i \in \{1, ..., v-u\}$. Combining this with (2.4) gives the claim. □

If the observations are not centered, we can at least state the following estimate.

Lemma 2.17. *Let $\kappa \in \mathbb{R}$ with $|\kappa| \leq M_p$, $p \in [1, \infty)$, and $u, v \in \mathbb{N}_0$ with $u < v \leq n$, $n \in \mathbb{N}$. Then there exists a positive constant C_p depending only on p such that*

$$\mathbb{E}\left[\left|\sum_{i=u+1}^{v}(X_i - \kappa)\right|^p\right] \leq C_p M_p (v-u)^p.$$

Proof. Fix $u, v \in \mathbb{N}_0$ with $u < v \leq n$, $n \in \mathbb{N}$. Let $\kappa \in \mathbb{R}$ with $|\kappa| \leq M_p$, $p \in [1, \infty)$. By the c_r-Inequality, we get

$$\mathbb{E}\left[\left|\sum_{i=u+1}^{v}(X_i - \kappa)\right|^p\right] = \mathbb{E}\left[\left|\sum_{i=1}^{v-u}(X_{u+i} - \kappa)\right|^p\right] \leq (v-u)^{p-1}\sum_{i=1}^{v-u}\mathbb{E}[|X_{u+i} - \kappa|^p]. \tag{2.5}$$

Another application of the c_r-Inequality leads to

$$\mathbb{E}[|X_{u+i} - \kappa|^p] \leq 2^{p-1}(\mathbb{E}[|X_{u+i}|^p] + |\kappa|^p) \leq 2^{p-1}(M_p + M_p^p) \leq C_p M_p$$

for all $i \in \{1, ..., v-u\}$, where $C_p > 0$ is a constant, which depends on p. Combining this with (2.5) completes the proof. □

Chapter 3

Known expectations

In this chapter we study the estimation of the multiple change-point by the least squares method and the asymptotic properties of such estimators if the expectations $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)$ are assumed to be known. This assumption is uncommon for practical applications, but the essential approach as well as the used methods to conclude results in our multiple change-point model can be well presented.
First we have a closer look at the estimator of the multiple change-point. The next section is concerned with the proof of weak and strong consistency. Finally, we investigate convergence in distribution to derive a confidence region for the moments of change.

3.1 Estimation of the multiple change-point

Our first purpose is to estimate the multiple change-point $\boldsymbol{\theta} = (\theta_1, \theta_2) \in \Theta$. To do this, we estimate the moments of change $\boldsymbol{\tau}_n = (\tau_n, \sigma_n) \in \Delta_n$ previously. By the least squares method, we are interested in finding all minimizers of the random criterion function $\bar{S}_n$

given by

$$\bar{S}_n(k,l) := \sum_{i=1}^{k}(X_i-\alpha)^2 + \sum_{i=k+1}^{l}(X_i-\beta)^2 + \sum_{i=l+1}^{n}(X_i-\gamma)^2, \quad (k,l)\in\Delta_n. \tag{3.1}$$

It is easily seen that $\bar{S}_n$ has at least one minimizer. To compute all minimizing points of $\bar{S}_n$ and get further results, we introduce another random criterion function $\bar{M}_n$, which has the same minimizers. Let

$$\bar{M}_n(k,l) := \sum_{i=1}^{k} a_1(X_i) + \sum_{i=1}^{l} a_2(X_i), \quad (k,l)\in\Delta_n,$$

where the mappings $a_1, a_2 : \mathbb{R} \longrightarrow \mathbb{R}$ are given by

$$a_1(x) := 2(\beta-\alpha)x + \alpha^2 - \beta^2 \quad \text{and} \quad a_2(x) := 2(\gamma-\beta)x + \beta^2 - \gamma^2. \tag{3.2}$$

Lemma 3.1. *Let $n \in \mathbb{N}$. Then*

$$\operatorname{Argmin}\left(\bar{S}_n\right) = \operatorname{Argmin}\left(\bar{M}_n\right).$$

Proof. Fix $n\in\mathbb{N}$ and $(k,l)\in\Delta_n$. An easy computation yields

$$\begin{aligned}
\bar{S}_n(k,l) &= \sum_{i=1}^{k}(X_i-\alpha)^2 + \sum_{i=k+1}^{l}(X_i-\beta)^2 + \sum_{i=l+1}^{n}(X_i-\gamma)^2 \\
&= \sum_{i=1}^{k}\left[(X_i-\alpha)^2 - (X_i-\beta)^2\right] + \sum_{i=1}^{l}\left[(X_i-\beta)^2 - (X_i-\gamma)^2\right] + \sum_{i=1}^{n}(X_i-\gamma)^2 \\
&= \sum_{i=1}^{k} a_1(X_i) + \sum_{i=1}^{l} a_2(X_i) + \sum_{i=1}^{n}(X_i-\gamma)^2 \\
&= \bar{M}_n(k,l) + \sum_{i=1}^{n}(X_i-\gamma)^2.
\end{aligned}$$

Since the last sum does not depend on $(k,l)\in\Delta_n$, it has no influence on the minimizing points of $\bar{S}_n$. □

To make sure that the estimator of $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ is well-defined, if more than one minimizer of $\bar{S}_n$ and $\bar{M}_n$ exists, it might be expedient to define a choice function

$$\bar{\phi} : \operatorname{Argmin}\left(\bar{M}_n\right) \longrightarrow \Delta_n,$$

which accurately chooses one minimizer. Hence it is meant that $\bar{\boldsymbol{\tau}}_n = \bar{\phi}\left(\text{Argmin}\left(\bar{M}_n\right)\right)$ when we write

$$\bar{\boldsymbol{\tau}}_n := (\bar{\tau}_n, \bar{\sigma}_n) := \underset{(k,l)\in\Delta_n}{\text{argmin}}\, \bar{M}_n(k,l) \tag{3.3}$$

hereafter.

Remark 3.2. The choice function can be selected arbitrarily. For instance, Seijo and Sen [25, p. 428, Definition 2.4] have suggested the smallest argmax functional for maximizing problems. The main idea is to choose the maximizer with the smallest first component. If there are several maximizing points with the smallest first element, then take this one with the smallest second component. In this work we assign the approach to minimizers.

The question arises under which condition has $\bar{M}_n$ an unique minimizer.

Lemma 3.3. *Let $n \in \mathbb{N}$ and let Q_1, Q_2, Q_3 be continuous distributions. Then*

$$\left|\text{Argmin}\left(\bar{M}_n\right)\right| = 1 \qquad \textit{almost surely.}$$

Proof. Fix $n \in \mathbb{N}$. Of course, $\bar{M}_n$ has at least one minimizer, because the domain Δ_n is finite. It follows that

$$\begin{aligned}
&\mathbb{P}\left[\left|\text{Argmin}\left(\bar{M}_n\right)\right| = 1\right] \\
&\qquad = 1 - \mathbb{P}\left[\left\{\left|\text{Argmin}\left(\bar{M}_n\right)\right| = 0\right\} \cup \left\{\left|\text{Argmin}\left(\bar{M}_n\right)\right| \geq 2\right\}\right] \\
&\qquad \geq 1 - \left(\mathbb{P}\left[\left|\text{Argmin}\left(\bar{M}_n\right)\right| = 0\right] + \mathbb{P}\left[\left|\text{Argmin}\left(\bar{M}_n\right)\right| \geq 2\right]\right) \\
&\qquad = 1 - \mathbb{P}\left[\left|\text{Argmin}\left(\bar{M}_n\right)\right| \geq 2\right] \\
&\qquad = 1 - \mathbb{P}\left[\bigcup_{(k_1,l_1)\neq(k_2,l_2)\in\Delta_n} \left\{\bar{M}_n(k_1,l_1) = \bar{M}_n(k_1,l_1)\right\}\right] \\
&\qquad \geq 1 - \sum_{(k_1,l_1)\neq(k_2,l_2)\in\Delta_n} \mathbb{P}\left[\bar{M}_n(k_1,l_1) - \bar{M}_n(k_2,l_2) = 0\right].
\end{aligned} \tag{3.4}$$

We distinguish several cases to compute $\bar{M}_n(k_1,l_1) - \bar{M}_n(k_2,l_2)$ by definition.

(i) (a) Let $1 \leq k_1 < k_2 < l_2 < l_1 \leq n-1$. Then

$$\bar{M}_n(k_1,l_1) - \bar{M}_n(k_2,l_2) = -\sum_{i=k_1+1}^{k_2} a_1(X_i) + \sum_{i=l_2+1}^{l_1} a_2(X_i).$$

(b) Let $1 \leq k_1 < k_2 \leq l_1 < l_2 \leq n-1$. Then

$$\bar{M}_n(k_1,l_1) - \bar{M}_n(k_2,l_2) = -\sum_{i=k_1+1}^{k_2} a_1(X_i) - \sum_{i=l_1+1}^{l_2} a_2(X_i).$$

(c) Let $1 \leq k_1 < l_1 < k_2 < l_2 \leq n-1$. Then

$$\bar{M}_n(k_1,l_1) - \bar{M}_n(k_2,l_2) = -\sum_{i=k_1+1}^{l_1} a_1(X_i) - \sum_{i=l_1+1}^{k_2} (a_1(X_i)+a_2(X_i)) - \sum_{i=k_2+1}^{l_2} a_2(X_i).$$

(ii) (a) Let $1 \leq k_2 < k_1 < l_1 < l_2 \leq n-1$. Then

$$\bar{M}_n(k_1,l_1) - \bar{M}_n(k_2,l_2) = \sum_{i=k_2+1}^{k_1} a_1(X_i) - \sum_{i=l_1+1}^{l_2} a_2(X_i).$$

(b) Let $1 \leq k_2 < k_1 \leq l_2 < l_1 \leq n-1$. Then

$$\bar{M}_n(k_1,l_1) - \bar{M}_n(k_2,l_2) = \sum_{i=k_2+1}^{k_1} a_1(X_i) + \sum_{i=l_2+1}^{l_1} a_2(X_i).$$

(c) Let $1 \leq k_2 < l_2 < k_1 < l_1 \leq n-1$. Then

$$\bar{M}_n(k_1,l_1) - \bar{M}_n(k_2,l_2) = \sum_{i=k_2+1}^{l_2} a_1(X_i) + \sum_{i=l_2+1}^{k_1} (a_1(X_i)+a_2(X_i)) + \sum_{i=k_1+1}^{l_1} a_2(X_i).$$

By the independence of $X_1, ..., X_n$ and the definitions of a_1 and a_2, we obtain sums of independent random variables in each case. Since Q_1, Q_2 and Q_3 are continuous distributions, we can conclude by convolution and the definitions of a_1 and a_2 that $\bar{M}_n(k_1,l_1) - \bar{M}_n(k_2,l_2)$ are continuous distributed random variables for each $(k_1,l_1),(k_2,l_2) \in \Delta_n$ with $(k_1,l_1) \neq (k_2,l_2)$. This gives $\mathbb{P}\left[\bar{M}_n(k_1,l_1) - \bar{M}_n(k_2,l_2) = 0\right] = 0$ for all $(k_1,l_1),(k_2,l_2) \in \Delta_n$ with $(k_1,l_1) \neq (k_2,l_2)$. By (3.4), we have $\mathbb{P}\left[\left|\mathrm{Argmin}\left(\bar{M}_n\right)\right| = 1\right] = 1$. □

Based on the estimator $\bar{\boldsymbol{\tau}}_n$ of moments of change $\boldsymbol{\tau}_n$, we are able to construct an estimator of the multiple change-point $\boldsymbol{\theta}$. By simple properties of the floor function (compare Lemma A.1 (i)), it is easy to check that $\frac{1}{n}\boldsymbol{\tau}_n = \frac{1}{n}(\lfloor n\theta_1 \rfloor, \lfloor n\theta_2 \rfloor) \xrightarrow[n\to\infty]{} (\theta_1, \theta_2) = \boldsymbol{\theta}$. Hence

$$\bar{\boldsymbol{\theta}}_n := \frac{1}{n}\bar{\boldsymbol{\tau}}_n \tag{3.5}$$

is a reasonable estimator of $\boldsymbol{\theta}$.

Though, the proof of consistency in the next section requires another form of $\bar{\boldsymbol{\theta}}_n$. Let us denote by $\bar{\rho}_n$ the random criterion function

$$\bar{\rho}_n(s,t) := \frac{1}{n}\bar{M}_n(\lfloor ns \rfloor, \lfloor nt \rfloor), \qquad (s,t) \in \Theta_n,$$

where

$$\Theta_n := \left\{ (s,t) \in \Theta \middle| s \geq \frac{1}{n}, t - s \geq \frac{1}{n}, 1 - t \geq \frac{1}{n} \right\}. \tag{3.6}$$

Lemma 3.4. *Let $n \in \mathbb{N}$. Then*

$$\bar{\boldsymbol{\theta}}_n = \underset{(s,t)\in\Theta_n}{\operatorname{argmin}}\, \bar{\rho}_n(s,t).$$

Proof. Fix $n \in \mathbb{N}$. By (3.3), we have $\bar{\boldsymbol{\tau}}_n = (\bar{\tau}_n, \bar{\sigma}_n) \in \Delta_n = \left\{(k,l) \in \mathbb{N}^2 \middle| 1 \leq k < l \leq n-1\right\}$. This gives $\bar{\boldsymbol{\theta}}_n = \frac{1}{n}(\bar{\tau}_n, \bar{\sigma}_n) \in \Theta_n$. Since $(\bar{\tau}_n, \bar{\sigma}_n)$ minimizes $\bar{M}_n$, we obtain for all $(s,t) \in \Theta_n$

$$\bar{\rho}_n\left(\bar{\boldsymbol{\theta}}_n\right) = \bar{\rho}_n\left(\frac{1}{n}\bar{\tau}_n, \frac{1}{n}\bar{\sigma}_n\right) = \frac{1}{n}\bar{M}_n(\bar{\tau}_n, \bar{\sigma}_n) \leq \frac{1}{n}\bar{M}_n(\lfloor ns \rfloor, \lfloor nt \rfloor) = \bar{\rho}_n(s,t).$$

The inequality is a consequence of $\Delta_n = \{(\lfloor ns \rfloor, \lfloor nt \rfloor) \in \mathbb{N}^2 | (s,t) \in \Theta_n\}$, which is shown in Lemma A.3. □

Remark 3.5. The factor n^{-1} in the definition of $\bar{\rho}_n$ does not influence the minimizing points of $\bar{\rho}_n$, but the proof of consistency of $\bar{\boldsymbol{\theta}}_n$ requires this factor.

Due to the following lemma, we get in the framework of Theorem 2.1 to prove consistency of $\bar{\boldsymbol{\theta}}_n$.

Lemma 3.6. *$\bar{\rho}_n$, $n \in \mathbb{N}$, is a stochastic process with trajectories in the multivariate Skorokhod space $D(\Theta_n)$.*

Proof. We outline the proof. The details are left to the reader. At first observe that a sequence of vectors converges to another vector if and only if the convergence occurs component by component. Notice that the floor function $x \longrightarrow \lfloor x \rfloor$ is an element of the Skorokhod space $D(\mathbb{R})$ and $\lfloor x \rfloor \in \mathbb{N}$ for any $x \in \mathbb{R}$. Moreover, from analysis it is well-known that a sequence of natural numbers converges to a natural number if and only if the numbers of the sequence are constant from an index. Hence, by definitions of $\bar{\rho}_n$ and $\bar{M}_n$, it is easy to check the claim. □

3.2 Consistency of the multiple change-point estimator

This section deals with the weak and strong consistency of $\bar{\boldsymbol{\theta}}_n$. To get consistency, we apply Theorem 2.1. The first part of this section is concerned with the uniform convergence of $\bar{\rho}_n$.

Proposition 3.7. *If $M_2 < \infty$, then there exists $C > 0$ such that for all $n \in \mathbb{N}$ and $\varepsilon > 0$*

$$\mathbb{P}\left[\sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \mathbb{E}[\bar{\rho}_n(s,t)]| > \varepsilon\right] \le C\varepsilon^{-2}n^{-1}.$$

Proof. Fix $n \in \mathbb{N}$. Let $C = C(\alpha, \beta, \gamma) > 0$ be a generic constant. We have

$$\begin{aligned}
&\sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \mathbb{E}[\bar{\rho}_n(s,t)]| \\
&= \frac{1}{n} \sup_{(s,t)\in\Theta_n} \left|\bar{M}_n(\lfloor ns\rfloor, \lfloor nt\rfloor) - \mathbb{E}\left[\bar{M}_n(\lfloor ns\rfloor, \lfloor nt\rfloor)\right]\right| && \text{by def. of } \bar{\rho}_n \\
&= \frac{1}{n} \max_{(k,l)\in\Delta_n} \left|\bar{M}_n(k,l) - \mathbb{E}\left[\bar{M}_n(k,l)\right]\right| && \text{by Lem. A.3} \\
&= \frac{1}{n} \max_{(k,l)\in\Delta_n} \left|\sum_{i=1}^{k}(a_1(X_i) - \mathbb{E}[a_1(X_i)]) + \sum_{i=1}^{l}(a_2(X_i) - \mathbb{E}[a_2(X_i)])\right| && \text{by def. of } \bar{M}_n \\
&\le \frac{1}{n} \max_{(k,l)\in\Delta_n} \left(\left|\sum_{i=1}^{k}(a_1(X_i) - \mathbb{E}[a_1(X_i)])\right| + \left|\sum_{i=1}^{l}(a_2(X_i) - \mathbb{E}[a_2(X_i)])\right|\right) && \text{by Tr. In.}
\end{aligned}$$

Throughout the proof, $Z_{1,i}$ and $Z_{2,i}$, $1 \le i \le n$, stand for

$$Z_{1,i} := a_1(X_i) - \mathbb{E}[a_1(X_i)] \quad \text{and} \quad Z_{2,i} := a_2(X_i) - \mathbb{E}[a_2(X_i)].$$

It is a simple matter to conclude that $Z_{1,1}, ..., Z_{1,n}$ as well as $Z_{2,1}, ..., Z_{2,n}$ are independent, centered and 2-fold integrable (note that $X_1, ..., X_n$ are independent and 2-fold integrable). We deduce that

$$\begin{aligned}
&\sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \mathbb{E}[\bar{\rho}_n(s,t)]| \\
&\qquad \le \frac{1}{n} \max_{(k,l)\in\{1,...,n\}^2} \left(\left|\sum_{i=1}^{k} Z_{1,i}\right| + \left|\sum_{i=1}^{l} Z_{2,i}\right|\right) && \text{by } \Delta_n \subseteq \{1,...,n\}^2 \\
&\qquad = \frac{1}{n}\left(\max_{1\le k\le n}\left|\sum_{i=1}^{k} Z_{1,i}\right| + \max_{1\le l\le n}\left|\sum_{i=1}^{l} Z_{2,i}\right|\right).
\end{aligned}$$

Consequently, wee see for all $\varepsilon > 0$ that

$$\begin{aligned}
&\mathbb{P}\left[\sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \mathbb{E}[\bar{\rho}_n(s,t)]| > \varepsilon\right] \\
&\quad \le \mathbb{P}\left[\max_{1\le k\le n}\left|\sum_{i=1}^k Z_{1,i}\right| + \max_{1\le l\le n}\left|\sum_{i=1}^l Z_{2,i}\right| > n\varepsilon\right] \\
&\quad \le \mathbb{P}\left[\max_{1\le k\le n}\left|\sum_{i=1}^k Z_{1,i}\right| > n\frac{\varepsilon}{2}\right] + \mathbb{P}\left[\max_{1\le l\le n}\left|\sum_{i=1}^l Z_{2,i}\right| > n\frac{\varepsilon}{2}\right] && \text{by Lem. A.4} \\
&\quad = \mathbb{P}\left[\max_{1\le k\le n}\left|\sum_{i=1}^k Z_{1,i}\right|^2 > \left(n\frac{\varepsilon}{2}\right)^2\right] + \mathbb{P}\left[\max_{1\le l\le n}\left|\sum_{i=1}^l Z_{2,i}\right|^2 > \left(n\frac{\varepsilon}{2}\right)^2\right].
\end{aligned}$$

Note that $\left(\left|\sum_{i=1}^k Z_{1,i}\right|^2\right)_{1\le k\le n}$ and $\left(\left|\sum_{i=1}^l Z_{2,i}\right|^2\right)_{1\le l\le n}$ are non-negative submartingales by Lemma 2.14. We thus apply Doob's Inequality, given in Lemma 2.10 (i), and obtain

$$\begin{aligned}
&\mathbb{P}\left[\sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \mathbb{E}[\bar{\rho}_n(s,t)]| > \varepsilon\right] \\
&\quad \le C\varepsilon^{-2}n^{-2}\left(\mathbb{E}\left[\left|\sum_{i=1}^n Z_{1,i}\right|^2\right] + \mathbb{E}\left[\left|\sum_{i=1}^n Z_{2,i}\right|^2\right]\right). \qquad (3.7)
\end{aligned}$$

In addition, we find some upper bounds for the second moments. We infer that

$$\begin{aligned}
\mathbb{E}\left[\left|\sum_{i=1}^n Z_{1,i}\right|^2\right] &= \mathbb{E}\left[\left|\sum_{i=1}^n (a_1(X_i) - \mathbb{E}[a_1(X_i)])\right|^2\right] \\
&\le C\sum_{i=1}^n \mathbb{E}\left[|a_1(X_i) - \mathbb{E}[a_1(X_i)]|^2\right] && \text{by Lem. 2.15} \\
&\le C\sum_{i=1}^n \left(\mathbb{E}\left[|a_1(X_i)|^2\right] + |\mathbb{E}[a_1(X_i)]|^2\right) && \text{by } c_r\text{-In.} \\
&\le C\sum_{i=1}^n \mathbb{E}\left[|a_1(X_i)|^2\right]. && \text{by Jensen In.} \qquad (3.8)
\end{aligned}$$

We infer for $i \in \{1, ..., n\}$ that

$$\begin{aligned}
\mathbb{E}\left[|a_1(X_i)|^2\right] &= \mathbb{E}\left[\left|2(\beta-\alpha)X_i + \alpha^2 - \beta^2\right|^2\right] && \text{by def. of } a_1 \\
&\leq 2\left(|2(\beta-\alpha)|^2 \mathbb{E}\left[|X_i|^2\right] + \left|\alpha^2 - \beta^2\right|^2\right) && \text{by } c_r\text{-In.} \\
&\leq 2\left(|2(\beta-\alpha)|^2 M_2 + \left|\alpha^2 - \beta^2\right|^2\right) \\
&\leq C && \text{by } M_2 < \infty.
\end{aligned}$$

By (3.8), we have

$$\mathbb{E}\left[\left|\sum_{i=1}^n Z_{1,i}\right|^2\right] \leq Cn. \tag{3.9}$$

In the same manner we can see that

$$\mathbb{E}\left[\left|\sum_{i=1}^n Z_{2,i}\right|^2\right] \leq Cn. \tag{3.10}$$

Combining (3.7) with (3.9) and (3.10) gives the claim. □

We now calculate $\mathbb{E}[\bar{\rho}_n(s,t)]$ for all $(s,t) \in \Theta$. To do this, we divide Θ into disjoint subsets (displayed in Figure 3.1) according to the position of $(s,t) \in \Theta$ relative to the multiple change-point $(\theta_1, \theta_2) \in \Theta$.

Let

$$\Theta = \bigcup_{i=1}^6 \Theta^i, \tag{3.11}$$

where

$$\begin{aligned}
\Theta^1 &:= \{(s,t) \in \Theta | s < t \leq \theta_1 < \theta_2\}, & \Theta^4 &:= \{(s,t) \in \Theta | \theta_1 < s < t \leq \theta_2\}, \\
\Theta^2 &:= \{(s,t) \in \Theta | s \leq \theta_1 < t \leq \theta_2\}, & \Theta^5 &:= \{(s,t) \in \Theta | \theta_1 < s \leq \theta_2 < t\}, \\
\Theta^3 &:= \{(s,t) \in \Theta | s \leq \theta_1 < \theta_2 < t\}, & \Theta^6 &:= \{(s,t) \in \Theta | \theta_1 < \theta_2 < s < t\}.
\end{aligned} \tag{3.12}$$

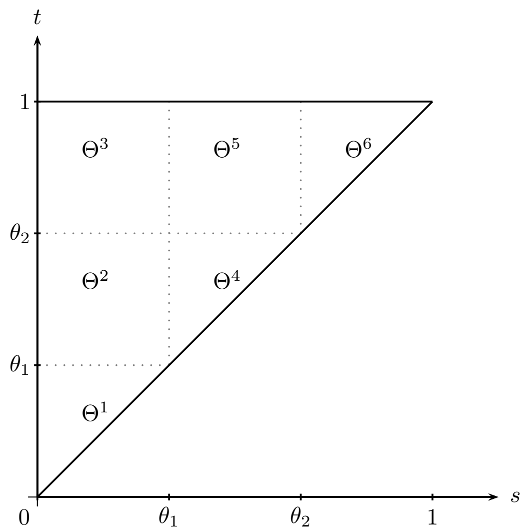

Figure 3.1: Partition of Θ into $\Theta^1, ..., \Theta^6$

Lemma 3.8. *Let $n \in \mathbb{N}$. Then*

$$\mathbb{E}[\bar{\rho}_n(s,t)] = \begin{cases} \left(\frac{\lfloor nt \rfloor}{n} - \frac{\lfloor ns \rfloor}{n}\right)(\alpha-\beta)^2 - \frac{\lfloor nt \rfloor}{n}(\alpha-\gamma), & (s,t) \in \Theta^1 \cap \Theta_n, \\ \left(\frac{\tau_n}{n} - \frac{\lfloor ns \rfloor}{n}\right)(\alpha-\beta)^2 - \frac{\tau_n}{n}(\alpha-\gamma)^2 & \\ \qquad + \left(\frac{\tau_n}{n} - \frac{\lfloor nt \rfloor}{n}\right)(\beta-\gamma)^2, & (s,t) \in \Theta^2 \cap \Theta_n, \\ \left(\frac{\tau_n}{n} - \frac{\lfloor ns \rfloor}{n}\right)(\alpha-\beta)^2 - \frac{\tau_n}{n}(\alpha-\gamma)^2 & \\ \qquad + \left(\frac{\lfloor nt \rfloor}{n} + \frac{\tau_n}{n} - 2\frac{\sigma_n}{n}\right)(\beta-\gamma)^2, & (s,t) \in \Theta^3 \cap \Theta_n, \\ \left(\frac{\lfloor ns \rfloor}{n} - \frac{\tau_n}{n}\right)(\alpha-\beta)^2 - \frac{\tau_n}{n}(\alpha-\gamma)^2 & \\ \qquad + \left(\frac{\tau_n}{n} - \frac{\lfloor nt \rfloor}{n}\right)(\beta-\gamma)^2, & (s,t) \in \Theta^4 \cap \Theta_n, \\ \left(\frac{\lfloor ns \rfloor}{n} - \frac{\tau_n}{n}\right)(\alpha-\beta)^2 - \frac{\tau_n}{n}(\alpha-\gamma)^2 & \\ \qquad + \left(\frac{\lfloor nt \rfloor}{n} + \frac{\tau_n}{n} - 2\frac{\sigma_n}{n}\right)(\beta-\gamma)^2, & (s,t) \in \Theta^5 \cap \Theta_n, \\ \left(\frac{\sigma_n}{n} - \frac{\tau_n}{n}\right)(\alpha-\beta)^2 & \\ \qquad + \left(\frac{\lfloor ns \rfloor}{n} - \frac{\tau_n}{n} - \frac{\sigma_n}{n}\right)(\alpha-\gamma)^2 & \\ \qquad + \left(\frac{\lfloor nt \rfloor}{n} - \frac{\lfloor ns \rfloor}{n} + \frac{\tau_n}{n} - \frac{\sigma_n}{n}\right)(\beta-\gamma)^2, & (s,t) \in \Theta^6 \cap \Theta_n. \end{cases}$$

Proof. Fix $n \in \mathbb{N}$. Recall that

$$\mathbb{E}[X_i] = \begin{cases} \alpha, & 1 \leq i \leq \tau_n, \\ \beta, & \tau_n + 1 \leq i \leq \sigma_n, \\ \gamma, & \sigma_n + 1 \leq i \leq n. \end{cases}$$

We leave it to the reader to verify that

$$\mathbb{E}[a_1(X_i)] = \begin{cases} -(\alpha - \beta)^2, & 1 \leq i \leq \tau_n, \\ (\alpha - \beta)^2, & \tau_n + 1 \leq i \leq \sigma_n, \\ (\alpha - \gamma)^2 - (\beta - \gamma)^2, & \sigma_n + 1 \leq i \leq n \end{cases}$$

and

$$\mathbb{E}[a_2(X_i)] = \begin{cases} (\alpha - \beta)^2 - (\alpha - \gamma)^2, & 1 \leq i \leq \tau_n, \\ -(\beta - \gamma)^2, & \tau_n + 1 \leq i \leq \sigma_n, \\ (\beta - \gamma)^2, & \sigma_n + 1 \leq i \leq n. \end{cases}$$

As an example, we compute the expectation of $\bar{\rho}_n(s,t)$ for $(s,t) \in \Theta^2 \cap \Theta_n$, since a similar procedure would bring the remaining cases. Lemma A.1 (ii) leads to

$$1 \leq \lfloor ns \rfloor \leq \tau_n < \lfloor nt \rfloor \leq \sigma_n < n.$$

By definitions of $\bar{\rho}_n$ and $\bar{M}_n$, the sums are split into segments according to above. Hence

$$\begin{aligned} \mathbb{E}[\bar{\rho}_n(s,t)] &= \frac{1}{n} \mathbb{E}\left[\bar{M}_n(\lfloor ns \rfloor, \lfloor nt \rfloor)\right] \\ &= \frac{1}{n} \left(\sum_{i=1}^{\lfloor ns \rfloor} \mathbb{E}[a_1(X_i)] + \sum_{i=1}^{\tau_n} \mathbb{E}[a_2(X_i)] + \sum_{i=\tau_n+1}^{\lfloor nt \rfloor} \mathbb{E}[a_2(X_i)] \right). \end{aligned}$$

A simple computation establishes the form as in the assertion. □

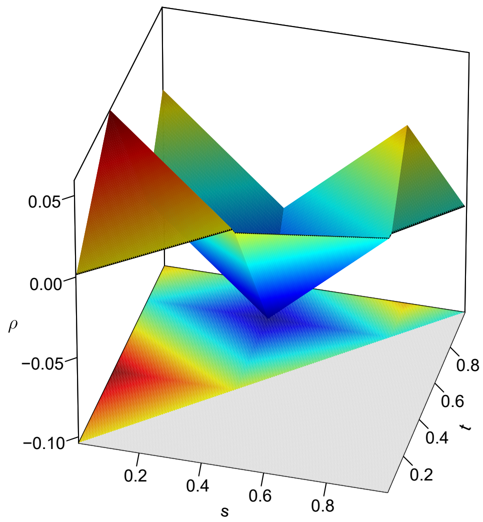

Figure 3.2: Plot of ρ for $\boldsymbol{\theta} = (\theta_1, \theta_2) = (0.4, 0.8)$ and $\boldsymbol{\alpha} = (\alpha, \beta, \gamma) = (0.6, 1, 0.5)$.

The function $\rho : \Theta \longrightarrow \mathbb{R}$ is defined by

$$\rho(s,t) := (\theta_1 - \theta_2)(\beta - \gamma)^2 - \theta_1(\alpha - \gamma)^2 \tag{3.13}$$
$$+ \begin{cases} (t-s)(\alpha-\beta)^2 + (\theta_1 - t)(\alpha-\gamma)^2 + (\theta_2 - \theta_1)(\beta-\gamma)^2, & (s,t) \in \Theta^1, \\ (\theta_1 - s)(\alpha-\beta)^2 + (\theta_2 - t)(\beta-\gamma)^2, & (s,t) \in \Theta^2, \\ (\theta_1 - s)(\alpha-\beta)^2 + (t - \theta_2)(\beta-\gamma)^2, & (s,t) \in \Theta^3, \\ (s - \theta_1)(\alpha-\beta)^2 + (\theta_2 - t)(\beta-\gamma)^2, & (s,t) \in \Theta^4, \\ (s - \theta_1)(\alpha-\beta)^2 + (t - \theta_2)(\beta-\gamma)^2, & (s,t) \in \Theta^5, \\ (s - \theta_2)(\alpha-\gamma)^2 + (t-s)(\beta-\gamma)^2 + (\theta_2 - \theta_1)(\alpha-\beta)^2, & (s,t) \in \Theta^6. \end{cases}$$

The function ρ (with domain Θ) is illustrated in Figure 3.2.

The following estimate states the uniform convergence of the expectation of $\bar{\rho}_n$ to ρ.

Proposition 3.9. *There exists $C > 0$ such that for all $n \in \mathbb{N}$*

$$\sup_{(s,t)\in\Theta_n} |\mathbb{E}[\bar{\rho}_n(s,t)] - \rho(s,t)| \leq Cn^{-1}.$$

Proof. Fix $n \in \mathbb{N}$. Our proof starts with the observation that the partition of Θ gives

$$\sup_{(s,t)\in\Theta_n} |\mathbb{E}[\bar{\rho}_n(s,t)] - \rho(s,t)| = \max_{i\in\{1,\dots,6\}} \sup_{(s,t)\in\Theta^i\cap\Theta_n} |\mathbb{E}[\bar{\rho}_n(s,t)] - \rho(s,t)|. \tag{3.14}$$

We show the way of our proceeding for $(s,t) \in \Theta^2 \cap \Theta_n$. We have

$$\begin{aligned}\rho(s,t) &= (\theta_1 - \theta_2)(\beta-\gamma)^2 - \theta_1(\alpha-\gamma)^2 + (\theta_1 - s)(\alpha-\beta)^2 + (\theta_2 - t)(\beta-\gamma)^2 \\ &= (\theta_1 - s)(\alpha-\beta)^2 - \theta_1(\alpha-\gamma)^2 + (\theta_1 - t)(\beta-\gamma)^2.\end{aligned}$$

Lemma 3.8 and the Triangle Inequality now imply

$$\begin{aligned}|\mathbb{E}[\bar{\rho}_n(s,t)] - \rho(s,t)| \le &\left|\frac{\lfloor ns\rfloor}{n} - s\right|(\alpha-\beta)^2 + \left|\frac{\lfloor nt\rfloor}{n} - t\right|(\beta-\gamma)^2 \\ &+ \left|\frac{\tau_n}{n} - \theta_1\right| \cdot \left|(\alpha-\beta)^2 - (\alpha-\gamma)^2 + (\beta-\gamma)^2\right|.\end{aligned}$$

Lemmas A.1 (iii) and A.1 (iv) lead to

$$\begin{aligned}&\sup_{(s,t)\in\Theta^2\cap\Theta_n} |\mathbb{E}[\bar{\rho}_n(s,t)] - \rho(s,t)| \\ &\le (\alpha-\beta)^2 \sup_{(s,t)\in\Theta^2\cap\Theta_n}\left|\frac{\lfloor ns\rfloor}{n} - s\right| + (\beta-\gamma)^2 \sup_{(s,t)\in\Theta^2\cap\Theta_n}\left|\frac{\lfloor nt\rfloor}{n} - t\right| \\ &\quad + \left|(\alpha-\beta)^2 - (\alpha-\gamma)^2 + (\beta-\gamma)^2\right| \cdot \left|\frac{\tau_n}{n} - \theta_1\right| \\ &\le (\alpha-\beta)^2 n^{-1} + (\beta-\gamma)^2 n^{-1} + \left|(\alpha-\beta)^2 - (\alpha-\gamma)^2 + (\beta-\gamma)^2\right| n^{-1} \\ &= Cn^{-1},\end{aligned}$$

where $C := C(\alpha,\beta,\gamma) := (\alpha-\beta)^2 + (\beta-\gamma)^2 + |(\alpha-\beta)^2 - (\alpha-\gamma)^2 + (\beta-\gamma)^2| > 0$. By an analogous estimate, an upper bound of the form Cn^{-1} can be found for the remaining partitions Θ^i of Θ, $i \in \{1,3,4,5,6\}$, such that we get the claim by (3.14). □

Propositions 3.7 and 3.9 help us to prove assumption (i) of Theorem 2.1. We now concern with assumption (ii) of Theorem 2.1, which says that $\boldsymbol{\theta} = (\theta_1, \theta_2) \in \Theta$ must be the well-separated minimizer of ρ.

Lemma 3.10. *The multiple change-point $\boldsymbol{\theta} = (\theta_1, \theta_2) \in \Theta$ is the unique minimizer of ρ.*

Proof. We have to show that $\rho(s,t) - \rho(\theta_1,\theta_2) > 0$ for all $(s,t) \in \Theta$ with $(s,t) \neq (\theta_1,\theta_2)$. Let us first observe that

$$\rho(\theta_1,\theta_2) = (\theta_1 - \theta_2)(\beta - \gamma)^2 - \theta_1(\alpha - \gamma)^2.$$

Recall the model assumptions $\alpha \neq \beta$ and $\beta \neq \gamma$. For $(s,t) \in \Theta^1$ and $(\theta_1,\theta_2) \in \Theta$ we notice $0 < s < t \leq \theta_1 < \theta_2 < 1$. Hence

$$\rho(s,t) - \rho(\theta_1,\theta_2) = (t-s)(\alpha - \beta)^2 + (\theta_1 - t)(\alpha - \gamma)^2 + (\theta_2 - \theta_1)(\beta - \gamma)^2 > 0.$$

We check at once that the same procedure leads to $\rho(s,t) - \rho(\theta_1,\theta_2) > 0$ for all $(s,t) \in \Theta^i$, $i \in \{3,4,5,6\}$. To complete the proof, we consider Θ^2. For $(s,t) \in \Theta^2$ and $(\theta_1,\theta_2) \in \Theta$ we note that $0 < s \leq \theta_1 < t \leq \theta_2 < 1$. The definition of ρ and $(s,t) \neq (\theta_1,\theta_2)$ yield

$$\rho(s,t) - \rho(\theta_1,\theta_2) = (\theta_1 - s)(\alpha - \beta)^2 + (\theta_2 - t)(\beta - \gamma)^2 > 0. \qquad \square$$

Throughout the entire work, let us denote by $\|\cdot\|$ the maximum norm.

Proposition 3.11. *The multiple change-point $\boldsymbol{\theta} \in \Theta$ is the well-separated minimizer of ρ.*

Proof. We first observe that $\boldsymbol{\theta} = (\theta_1,\theta_2) \in \Theta$ is a minimizer of ρ, which was proved in the previous lemma. By (2.1), it is sufficient to show that

$$\inf\{\rho(s,t) : \|(\theta_1,\theta_2) - (s,t)\| \geq \varepsilon,\ (s,t) \in \Theta\} - \rho(\theta_1,\theta_2) > 0$$

for all $\varepsilon > 0$. Fix $\varepsilon > 0$. By decomposition of Θ, we get

$$\begin{aligned}
&\inf\{\rho(s,t) : \|(\theta_1,\theta_2) - (s,t)\| \geq \varepsilon,\ (s,t) \in \Theta\} - \rho(\theta_1,\theta_2) \\
&= \inf\{\rho(s,t) - \rho(\theta_1,\theta_2) : \|(\theta_1,\theta_2) - (s,t)\| \geq \varepsilon,\ (s,t) \in \Theta\} \\
&= \min_{i \in \{1,\ldots,6\}} \inf\left\{\rho(s,t) - \rho(\theta_1,\theta_2) : \|(\theta_1,\theta_2) - (s,t)\| \geq \varepsilon,\ (s,t) \in \Theta^i\right\} \qquad \text{by (3.11)} \\
&=: \min_{i \in \{1,\ldots,6\}} \varrho_i. \qquad (3.15)
\end{aligned}$$

Write $C := C(\alpha,\beta,\gamma) := \min\{(\alpha-\beta)^2, (\beta-\gamma)^2\}$ and note that $C > 0$ by model assumptions $\alpha \neq \beta$ and $\beta \neq \gamma$. By definition of ρ, we obtain

$$\begin{aligned}
\varrho_2 &= \inf\left\{\rho(s,t) - \rho(\theta_1,\theta_2) : \|(\theta_1,\theta_2) - (s,t)\| \geq \varepsilon,\ (s,t) \in \Theta^2\right\} \\
&= \inf\left\{(\theta_1 - s)(\alpha-\beta)^2 + (\theta_2 - t)(\beta-\gamma)^2 : \max\{\theta_1 - s, \theta_2 - t\} \geq \varepsilon,\ (s,t) \in \Theta^2\right\} \\
&\geq \inf\left\{C((\theta_1 - s) + (\theta_2 - t)) : \max\{\theta_1 - s, \theta_2 - t\} \geq \varepsilon,\ (s,t) \in \Theta^2\right\} \\
&\geq \begin{cases} \inf\left\{C(\varepsilon + (\theta_2 - t)) : (s,t) \in \Theta^2\right\}, & \theta_1 - s \geq \theta_2 - t, \\ \inf\left\{C((\theta_1 - s) + \varepsilon) : (s,t) \in \Theta^2\right\}, & \theta_1 - s < \theta_2 - t, \end{cases} \\
&= C\varepsilon \\
&> 0.
\end{aligned}$$

We can now proceed analogously to conclude that $\varrho_3 > 0$, $\varrho_4 > 0$ and $\varrho_5 > 0$. Furthermore, by definition of ρ, we have

$$\begin{aligned}
\varrho_1 &= \inf\left\{\rho(s,t) - \rho(\theta_1,\theta_2) : \|(\theta_1,\theta_2) - (s,t)\| \geq \varepsilon,\ (s,t) \in \Theta^1\right\} \\
&= \inf\Big\{(t - s)(\alpha-\beta)^2 + (\theta_1 - t)(\alpha-\gamma)^2 \\
&\qquad\quad + (\theta_2 - \theta_1)(\beta-\gamma)^2 : \|(\theta_1,\theta_2) - (s,t)\| \geq \varepsilon,\ (s,t) \in \Theta^1\Big\} \\
&= (\theta_2 - \theta_1)(\beta-\gamma)^2 \\
&\geq C(\theta_2 - \theta_1) \\
&> 0,
\end{aligned}$$

the last inequality being a consequence of $(\theta_1,\theta_2) \in \Theta$. It is a simple matter to check $\varrho_6 > 0$. By (3.15), the proof is complete. □

We can now formulate and prove the weak consistency of $\bar{\boldsymbol{\theta}}_n$.

Theorem 3.12. *If $M_2 < \infty$, then*

$$\bar{\boldsymbol{\theta}}_n \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\theta}.$$

Proof. We apply Theorem 2.1. $\bar{\rho}_n$, $n \in \mathbb{N}$, is a stochastic process with trajectories in the multivariate Skorokhod space $D(\Theta_n)$ by Lemma 3.6. ρ has trajectories in the multivariate Skorokhod space $D(\Theta)$, since ρ is continuous, as is easy to check. Moreover, $(\Theta_n)_{n\in\mathbb{N}} \subseteq \Theta$ is a sequence of sets such that $\Theta_n \subseteq \Theta_{n+1}$ for every $n \in \mathbb{N}$ with $\bigcup_{n\in\mathbb{N}} \Theta_n = \Theta$. By Lemma 3.4, $\bar{\boldsymbol{\theta}}_n$ is a minimizer of $\bar{\rho}_n$ for any $n \in \mathbb{N}$. We infer that

$$\begin{aligned}
&\sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \rho(s,t)| \\
&\quad = \sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \mathbb{E}[\bar{\rho}_n(s,t)] + \mathbb{E}[\bar{\rho}_n(s,t)] - \rho(s,t)| \\
&\quad \leq \sup_{(s,t)\in\Theta_n} (|\bar{\rho}_n(s,t) - \mathbb{E}[\bar{\rho}_n(s,t)]| + |\mathbb{E}[\bar{\rho}_n(s,t)] - \rho(s,t)|) \qquad \text{by Tr. In.} \\
&\quad \leq \sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \mathbb{E}[\bar{\rho}_n(s,t)]| + \sup_{(s,t)\in\Theta_n} |\mathbb{E}[\bar{\rho}_n(s,t)] - \rho(s,t)| \qquad (3.16)
\end{aligned}$$

for each $n \in \mathbb{N}$. Letting $n \to \infty$, Propositions 3.7 and 3.9 lead to

$$\sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \rho(s,t)| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

In addition, $\boldsymbol{\theta} \in \Theta$ is the well-separated minimizer of ρ, see Proposition 3.11. An application of Theorem 2.1 finishes the proof. □

We can even prove strong consistency of $\bar{\boldsymbol{\theta}}_n$.

Theorem 3.13. *Suppose there is some $p \in (2,\infty)$ such that $M_p < \infty$. Then*

$$\bar{\boldsymbol{\theta}}_n \xrightarrow[n\to\infty]{a.s.} \boldsymbol{\theta}.$$

Proof. We apply Theorem 2.1 again. The basic framework is the same as in proof of Theorem 3.12. Furthermore, $\boldsymbol{\theta} \in \Theta$ is the well-separated minimizer of ρ by propsition 3.11. By (3.16), we observe that

$$\begin{aligned}
&\sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \rho(s,t)| \\
&\quad \leq \sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \mathbb{E}[\bar{\rho}_n(s,t)]| + \sup_{(s,t)\in\Theta_n} |\mathbb{E}[\bar{\rho}_n(s,t)] - \rho(s,t)|.
\end{aligned}$$

Assumption (i) of Theorem 2.1 is fulfilled if

$$\sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \rho(s,t)| \xrightarrow[n\to\infty]{a.s.} 0.$$

By Proposition 3.9, the proof is completed by showing that

$$\sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \mathbb{E}[\bar{\rho}_n(s,t)]| \xrightarrow[n\to\infty]{a.s.} 0. \tag{3.17}$$

For this purpose, we set

$$A_n(\varepsilon) := \left\{ \sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \mathbb{E}[\bar{\rho}_n(s,t)]| > \varepsilon \right\}$$

for each $n \in \mathbb{N}$ and $\varepsilon > 0$. By a similar estimate as in the proof of Proposition 3.7, there exists a constant $C > 0$ such that $\mathbb{P}[A_n(\varepsilon)] \leq C\varepsilon^{-p} n^{-p/2}$ for all $\varepsilon > 0$. Hence

$$\sum_{n=1}^{\infty} \mathbb{P}[A_n(\varepsilon)] \leq C\varepsilon^{-p} \sum_{n=1}^{\infty} n^{-p/2} < \infty.$$

The finiteness holds, because the series converges for $p > 2$. The first Borel-Cantelli Lemma (see for instance Schmidt [24, p. 227, Lemma 11.1.12]) leads to $\mathbb{P}\left[\limsup_{n\to\infty} A_n(\varepsilon)\right] = 0$ for all $\varepsilon > 0$. Hence

$$\begin{aligned}
\mathbb{P}\left[\sup_{(s,t)\in\Theta_n} |\bar{\rho}_n(s,t) - \mathbb{E}[\bar{\rho}_n(s,t)]| \xrightarrow[n\to\infty]{} 0 \right] &= \mathbb{P}\left[\bigcap_{\varepsilon\in\mathbb{Q}_{>0}} \bigcup_{m\in\mathbb{N}} \bigcap_{n\in\mathbb{N}_m} A_n(\varepsilon)^{\complement} \right] \\
&= 1 - \mathbb{P}\left[\bigcup_{\varepsilon\in\mathbb{Q}_{>0}} \bigcap_{m\in\mathbb{N}} \bigcup_{n\in\mathbb{N}_m} A_n(\varepsilon) \right] \\
&= 1 - \mathbb{P}\left[\bigcup_{\varepsilon\in\mathbb{Q}_{>0}} \limsup_{n\to\infty} A_n(\varepsilon) \right] \\
&= 1.
\end{aligned}$$

The last equality holds, because the countable union of null sets is also a null set. We have shown (3.17), which is our desired conclusion. □

3.3 Convergence in distribution

We proceed with the study of convergence in distribution of $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$. To do this, we apply Theorem 2.3. To stay in the framework of Theorem 2.3, the main idea is to introduce another process $\bar{\Gamma}_n$, which is minimized by $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$ for each $n \in \mathbb{N}$. The so-called *rescaled process* $\bar{\Gamma}_n$ is defined by

$$\bar{\Gamma}_n(k,l) := \bar{M}_n(\tau_n + k, \sigma_n + l) - \bar{M}_n(\tau_n, \sigma_n), \qquad (k,l) \in H_n,$$

where

$$H_n := \left\{ (k,l) \in \mathbb{Z}^2 \middle| k \geq 1 - \tau_n,\ l - k \geq 1 - (\sigma_n - \tau_n),\ n - l \geq \sigma_n + 1 \right\}.$$

Lemma 3.14. *Let $n \in \mathbb{N}$. Then*

$$\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n \in \operatorname{Argmin}\left(\bar{\Gamma}_n\right).$$

Proof. Fix $n \in \mathbb{N}$. Note that $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n = (\bar{\tau}_n - \tau_n, \bar{\sigma}_n - \sigma_n)$ lies in H_n, which is clear from $\bar{\boldsymbol{\tau}}_n \in \Delta_n$ and $\boldsymbol{\tau}_n \in \mathbb{N}^2$. Moreover, from $(k,l) \in H_n$ it follows that $(\tau_n + k, \sigma_n + l) \in \Delta_n$. Since $\bar{\boldsymbol{\tau}}_n$ is a minimizer of $\bar{M}_n$, we obtain

$$\bar{\Gamma}_n(\bar{\tau}_n - \tau_n, \bar{\sigma}_n - \sigma_n) = \bar{M}_n(\bar{\tau}_n, \bar{\sigma}_n) - \bar{M}_n(\tau_n, \sigma_n) \leq \bar{M}_n(\tau_n + k, \sigma_n + l) - \bar{M}_n(\tau_n, \sigma_n) = \bar{\Gamma}_n(k,l)$$

for all $(k,l) \in H_n$. □

$\bar{\Gamma}_n$ has the following form.

Lemma 3.15. *Let $n \in \mathbb{N}$ and $(k,l) \in H_n$. Then*

$$\bar{\Gamma}_n(k,l) = \bar{\Gamma}_{n,1}(k) + \bar{\Gamma}_{n,2}(l)$$

with

$$\bar{\Gamma}_{n,1}(k) := \begin{cases} \sum_{i=1}^{k} a_1(X_{\tau_n + i}), & k \geq 0, \\ -\sum_{i=1}^{-k} a_1(X_{\tau_n - i + 1}), & k < 0 \end{cases} \quad \text{and} \quad \bar{\Gamma}_{n,2}(l) := \begin{cases} \sum_{i=1}^{l} a_2(X_{\sigma_n + i}), & l \geq 0, \\ -\sum_{i=1}^{-l} a_2(X_{\sigma_n - i + 1}), & l < 0, \end{cases}$$

where a_1 and a_2 are given by (3.2).

Proof. Fix $n \in \mathbb{N}$ and let $(k,l) \in H_n$. By definitions of $\bar{\Gamma}_n$ and $\bar{M}_n$, we have

$$\begin{aligned}\bar{\Gamma}_n(k,l) &= \bar{M}_n(\tau_n + k, \sigma_n + l) - \bar{M}_n(\tau_n, \sigma_n)\\ &= \left(\sum_{i=1}^{\tau_n+k} a_1(X_i) - \sum_{i=1}^{\tau_n} a_1(X_i)\right) + \left(\sum_{i=1}^{\sigma_n+l} a_2(X_i) - \sum_{i=1}^{\sigma_n} a_2(X_i)\right)\\ &=: \tilde{\Gamma}_{n,1}(k) + \tilde{\Gamma}_{n,2}(l).\end{aligned}$$

For $k \geq 0$ we have

$$\tilde{\Gamma}_{n,1}(k) = \sum_{i=\tau_n+1}^{\tau_n+k} a_1(X_i) = \sum_{i=1}^{k} a_1(X_{\tau_n+i}) = \bar{\Gamma}_{n,1}(k)$$

and for $k < 0$

$$\tilde{\Gamma}_{n,1}(k) = -\sum_{i=\tau_n+k+1}^{\tau_n} a_1(X_i) = -\sum_{i=1}^{-k} a_1(X_{\tau_n-i+1}) = \bar{\Gamma}_{n,1}(k).$$

The cases $l \geq 0$ and $l < 0$ to obtain the form of $\bar{\Gamma}_{n,2}$ are left to the reader. □

To show assumption (i) of Theorem 2.3, we establish convergence in distribution of all finite-dimensional distributions of $\bar{\Gamma}_n$. To this end, let $(\xi_{i,r})_{i\in\mathbb{N}}$, $r \in \{1,2,3\}$, be three independent sequences, which for each r consist of independent and identically distributed random variables with common distribution Q_r. Set

$$\Gamma(k,l) := \Gamma_1(k) + \Gamma_2(l), \qquad (k,l) \in \mathbb{Z}^2, \tag{3.18}$$

where

$$\Gamma_1(k) := \begin{cases} \sum\limits_{i=1}^{k} a_1(\xi_{i,2}), & k \geq 0,\\ -\sum\limits_{i=1}^{-k} a_1(\xi_{i,1}), & k < 0 \end{cases} \quad \text{and} \quad \Gamma_2(l) := \begin{cases} \sum\limits_{i=1}^{l} a_2(\xi_{i,3}), & l \geq 0,\\ -\sum\limits_{i=1}^{-l} a_2(\xi_{i,2}), & l < 0. \end{cases}$$

Remark 3.16. A trivial verification shows that Γ is a sum of two random walks with

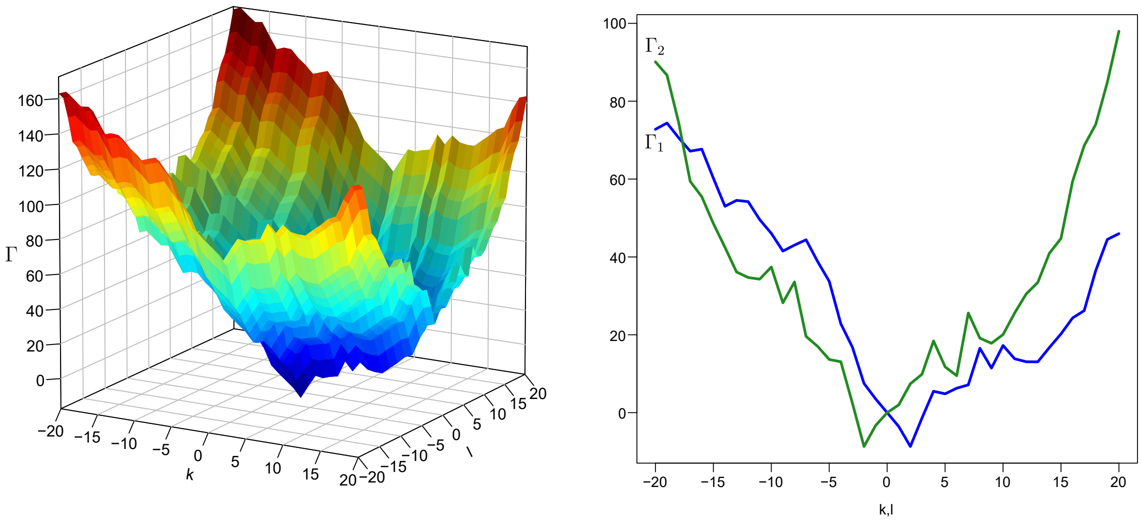

Figure 3.3: Plot of Γ, Γ_1 and Γ_2 for $Q_1 = N(3.2, 1)$, $Q_2 = N(5, 1)$, $Q_3 = N(2.6, 1)$.

positive drift. More precisely, we have $\Gamma(k, l) = \Gamma_1(k) + \Gamma_2(l)$ with

$$\Gamma_1(k) = \begin{cases} 2(\beta - \alpha) \sum_{i=1}^{k} (\xi_{i,2} - \beta) + k(\alpha - \beta)^2, & k \geq 0, \\ -2(\beta - \alpha) \sum_{i=1}^{-k} (\xi_{i,1} - \alpha) - k(\alpha - \beta)^2, & k < 0 \end{cases} \quad \text{and}$$

$$\Gamma_2(l) = \begin{cases} 2(\gamma - \beta) \sum_{i=1}^{l} (\xi_{i,3} - \gamma) + l(\beta - \gamma)^2, & l \geq 0, \\ -2(\gamma - \beta) \sum_{i=1}^{-l} (\xi_{i,2} - \beta) - l(\beta - \gamma)^2, & l < 0 \end{cases}$$

for each $(k, l) \in \mathbb{Z}^2$. Furthermore, the drift functions of Γ_1 and Γ_2 are given by

$$\mathbb{E}[\Gamma_1(k)] = (\alpha - \beta)^2 |k| \qquad \text{and} \qquad \mathbb{E}[\Gamma_2(l)] = (\beta - \gamma)^2 |l|.$$

The processes Γ, Γ_1 and Γ_2 are displayed in Figure 3.3. Note that these processes are only defined on integer numbers, but for clarity the processes are illustrated on real numbers (single points are connected).

Lemma 3.17. *Let $m \in \mathbb{N}$. Then for each collection $(k_1, l_1), ..., (k_m, l_m) \in \mathbb{Z}^2$ there exists $n_0 = n_0(k_1, l_1, ..., k_m, l_m) \in \mathbb{N}$ such that for all $n \geq n_0$*

$$\left(\bar{\Gamma}_n(k_1, l_1), ..., \bar{\Gamma}_n(k_m, l_m)\right) \stackrel{\mathcal{L}}{=} (\Gamma(k_1, l_1), ..., \Gamma(k_m, l_m)).$$

Proof. Fix $m \in \mathbb{N}$ and $(k_1, l_1), ..., (k_m, l_m) \in \mathbb{Z}^2$. Since it holds

$$\tau_n \xrightarrow[n\to\infty]{} \infty, \quad \sigma_n - \tau_n \xrightarrow[n\to\infty]{} \infty \quad \text{and} \quad n - \sigma_n \xrightarrow[n\to\infty]{} \infty \tag{3.19}$$

by Lemma A.2, we have $\bigcup_{n\in\mathbb{N}} H_n = \mathbb{Z}^2$. From this we can conclude that there exists $n_1 = n_1(k_1, l_1, ..., k_m, l_m) \in \mathbb{N}$ such that for all $n \geq n_1$

$$(k_1, l_1), ..., (k_m, l_m) \in H_n. \tag{3.20}$$

Furthermore, by (3.19), there exists $n_2 = n_2(k_1, l_1, ..., k_m, l_m) \in \mathbb{N}$ such that for all $n \geq n_2$

$$\max\{|k_1|, |l_1|, ..., |k_m|, |l_m|\} \leq \sigma_n - \tau_n. \tag{3.21}$$

From now on, let $n \geq n_0 := \max\{n_1, n_2\}$. We consider several cases. Fix $r \in \{1, ..., m\}$.

(i) We distinguish two cases for k_r.

(a) Let $k_r > 0$. By (3.21), for all $i \in \{1, ..., k_r\}$ we see that

$$\tau_n + 1 \leq \tau_n + i \leq \tau_n + k_r \leq \tau_n + (\sigma_n - \tau_n) = \sigma_n,$$

and consequently $X_{\tau_n+i} \sim Q_2$.

(b) Let $k_r < 0$. By (3.20), for all $i \in \{1, ..., -k_r\}$ we find that

$$\tau_n \geq \tau_n - i + 1 \geq \tau_n + k_r + 1 \geq \tau_n + (1 - \tau_n) + 1 = 2,$$

and so $X_{\tau_n-i+1} \sim Q_1$.

(ii) We distinguish two cases for l_r.

(a) Let $l_r > 0$. By (3.20), for all $i \in \{1, ..., l_r\}$ we obtain

$$\sigma_n + 1 \leq \sigma_n + i \leq \sigma_n + l_r \leq \sigma_n + (n - \sigma_n - 1) = n - 1,$$

and hence $X_{\sigma_n+i} \sim Q_3$.

(b) Let $l_r < 0$. By (3.21), for all $i \in \{1, ..., -l_r\}$ we have

$$\sigma_n \geq \sigma_n - i + 1 \geq \sigma_n - (-l_r) + 1 \geq \sigma_n - (\sigma_n - \tau_n) + 1 = \tau_n + 1,$$

which gives $X_{\sigma_n - i+1} \sim Q_2$.

Write $\mathbf{X}_n := (X_1, ..., X_n)$ and $\boldsymbol{\xi}_n := (\xi_{1,1}, ..., \xi_{\tau_n,1}, \xi_{1,2}, ..., \xi_{\sigma_n-\tau_n,2}, \xi_{1,3}, ..., \xi_{n-\sigma_n,3})$, where $(\xi_{i,1})_{i\in\mathbb{N}}$, $(\xi_{i,2})_{i\in\mathbb{N}}$ and $(\xi_{i,3})_{i\in\mathbb{N}}$ are the sequences defined in (3.18). The independence of observations, distinction of cases and independence assumptions to the sequences imply

$$\begin{aligned}
\mathbb{P}_{\mathbf{X}_n} = \bigotimes_{i=1}^{n} \mathbb{P}_{X_i} &= \bigotimes_{i=1}^{\tau_n} \mathbb{P}_{X_i} \otimes \bigotimes_{i=\tau_n+1}^{\sigma_n} \mathbb{P}_{X_i} \otimes \bigotimes_{i=\sigma_n+1}^{n} \mathbb{P}_{X_i} \\
&= \bigotimes_{i=1}^{\tau_n} \mathbb{P}_{\xi_{i,1}} \otimes \bigotimes_{i=1}^{\sigma_n-\tau_n} \mathbb{P}_{\xi_{i,2}} \otimes \bigotimes_{i=1}^{n-\sigma_n} \mathbb{P}_{\xi_{i,3}} \\
&= \mathbb{P}_{\boldsymbol{\xi}_n}.
\end{aligned}$$

Therefore $\mathbf{X}_n \overset{\mathcal{L}}{=} \boldsymbol{\xi}_n$. The crucial fact is that the processes $\bar{\Gamma}_n$ and Γ depend on random variables. Thus, $\bar{\Gamma}_n$ and Γ can be considered as measurable transformations of $\mathbf{X}_n$ and $\boldsymbol{\xi}_n$, which lead to

$$\begin{aligned}
\left(\bar{\Gamma}_n(k_1, l_1), ..., \bar{\Gamma}_n(k_m, l_m)\right) &= \left(\bar{\Gamma}_n(k_1, l_1; \mathbf{X}_n), ..., \bar{\Gamma}_n(k_m, l_m; \mathbf{X}_n)\right) \\
&\overset{\mathcal{L}}{=} \left(\bar{\Gamma}_n(k_1, l_1; \boldsymbol{\xi}_n), ..., \bar{\Gamma}_n(k_m, l_m; \boldsymbol{\xi}_n)\right) \\
&= (\Gamma(k_1, l_1; \boldsymbol{\xi}_n), ..., \Gamma(k_m, l_m; \boldsymbol{\xi}_n)) \\
&= (\Gamma(k_1, l_1), ..., \Gamma(k_m, l_m)).
\end{aligned}$$

□

We get convergence in distribution of all finite-dimensional distributions of $\bar{\Gamma}_n$ to Γ.

Proposition 3.18. *Let $m \in \mathbb{N}$ and $(k_1, l_1), ..., (k_m, l_m) \in \mathbb{Z}^2$. Then*

$$\left(\bar{\Gamma}_n(k_1, l_1), ..., \bar{\Gamma}_n(k_m, l_m)\right) \xrightarrow[n\to\infty]{\mathcal{L}} (\Gamma(k_1, l_1), ..., \Gamma(k_m, l_m)).$$

Proof. The assertion follows directly from Lemma 3.17. □

The task is now to prove stochastic boundedness of $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$, see assumption (ii) of Theorem 2.3. The following two technical lemmas are useful to estimate the error probability previously.

Lemma 3.19. *There exist $\delta > 0$ and $n_0 \in \mathbb{N}$ such that*

(i) $1 + n\delta < \tau_n - n\delta$,

(ii) $\tau_n + n\delta < \sigma_n - n\delta$ *and*

(iii) $\sigma_n + n\delta < n - n\delta$

for all $n \geq n_0$.

Proof. The procedure is to find a condition to $\delta > 0$ such that all inequalities are satisfied for a sufficiently large $n \in \mathbb{N}$. For example, we consider Inequality (ii), which is equivalent to

$$2n\delta < \sigma_n - \tau_n. \tag{3.22}$$

First observe that the properties of the floor function (Lemma A.1 (i)) lead to

$$\sigma_n - \tau_n = \lfloor n\theta_2 \rfloor - \lfloor n\theta_1 \rfloor > n\theta_2 - 1 - n\theta_1 = n(\theta_2 - \theta_1) - 1.$$

Consequently, the inequality holds in (3.22) if $\delta < \frac{1}{2}(\theta_2 - \theta_1) - \frac{1}{2n}$. Suppose for a moment that $n > \frac{5}{\theta_2 - \theta_1}$. Then we can choose δ with $\delta < \frac{2}{5}(\theta_2 - \theta_1)$. Set $n_2(\theta_1, \theta_2) := \left\lfloor \frac{5}{\theta_2 - \theta_1} \right\rfloor + 1$ and $\delta_2(\theta_1, \theta_2) := \frac{1}{3}(\theta_2 - \theta_1)$. Now, (3.22) holds for $\delta_2(\theta_1, \theta_2)$ and every $n \geq n_2(\theta_1, \theta_2)$. By similar arguments, we get $\delta_1(\theta_1) := \frac{1}{3}\theta_1$, $n_1(\theta_1) := \left\lfloor \frac{10}{\theta_1} \right\rfloor + 1$ and $\delta_3(\theta_2) := \frac{1}{3}(1 - \theta_2)$ such that Inequality (i) holds for $\delta_1(\theta_1)$ and all $n \geq n_1(\theta_1)$, and Inequality (iii) is fulfilled for $\delta_3(\theta_2)$ and all $n \in \mathbb{N}$. If we choose $\delta := \delta(\theta_1, \theta_2) := \min\{\delta_1(\theta_1), \delta_2(\theta_1, \theta_2), \delta_3(\theta_2)\}$ and $n_0 := n_0(\theta_1, \theta_2) := \max\{n_1(\theta_1), n_2(\theta_1, \theta_2)\}$, then the lemma follows. From $(\theta_1, \theta_2) \in \Theta$ we deduce that $\delta > 0$. □

Recall that $\|\cdot\|$ stands for the maximum norm. Let us denote by $H_{n,x,\delta}$ the set

$$H_{n,x,\delta} := \{(k,l) \in H_n | x \leq \|(k,l)\| \leq n\delta\}$$

for $n \in \mathbb{N}$, $x > 0$ and $\delta > 0$.

Lemma 3.20. *Let $x > 0$, $\delta > 0$ and $n \in \mathbb{N}$. Then*

$$\{x \leq \|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta\} \subseteq \bigcup_{(k,l)\in H_{n,x,\delta}} \left\{-\bar{\Gamma}_n(k,l) \geq 0\right\}.$$

Proof. Suppose, contrary to our claim, that there exists $\omega \in \{x \leq \|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta\}$, but

$$\omega \notin \bigcup_{(k,l)\in H_{n,x,\delta}} \left\{-\bar{\Gamma}_n(k,l) \geq 0\right\}.$$

So, we get $-\bar{\Gamma}_n(k,l) < 0$ for all $(k,l) \in H_{n,x,\delta}$. By Lemma 3.14, we have $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n \in H_n$, and, in consequence, $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n \in H_{n,x,\delta}$ by assumption. The definition of $\bar{\Gamma}_n$ gives

$$0 > -\bar{\Gamma}_n(\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n) = \bar{M}_n(\boldsymbol{\tau}_n) - \bar{M}_n(\bar{\boldsymbol{\tau}}_n),$$

which contradicts the fact that $\bar{\boldsymbol{\tau}}_n$ minimizes $\bar{M}_n$. □

We now derive an error estimate.

Lemma 3.21. *Suppose that $M_2 < \infty$. Then there exist $n_0 \in \mathbb{N}$, $\delta > 0$ and a constant $C > 0$ such that for all $n \geq n_0$ we have*

$$\mathbb{P}[x \leq \|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta] \leq Cx^{-1}$$

for all $x \geq 2$.

Proof. Let $x \geq 2$. By Lemma 3.19, there exist $\delta > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ the conditions (i)-(iii) in Lemma 3.19 hold. Consider $n \geq n_0$ and $\delta > 0$ and let

$C = C(\alpha, \beta, \gamma) > 0$ be a generic constant. By Lemma 3.20, we first observe that

$$\begin{aligned}\{x \leq \|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta\} &\subseteq \bigcup_{(k,l)\in H_{n,x,\delta}} \{-\bar{\Gamma}_n(k,l) \geq 0\} \\ &\subseteq \bigcup_{\substack{x\leq|k|\leq n\delta \\ |l|\leq n\delta}} \{-\bar{\Gamma}_n(k,l) \geq 0\} \cup \bigcup_{\substack{|k|\leq n\delta \\ x\leq|l|\leq n\delta}} \{-\bar{\Gamma}_n(k,l) \geq 0\} \\ &=: E \cup F. \end{aligned} \tag{3.23}$$

To simplify notation, the fact that some in this proof defined sets, random variables and probabilities depend on n, x or δ is omitted. We give the proof only for the estimate of the probability of E; the other case follows the same pattern. We find that

$$\begin{aligned} E \subseteq& \bigcup_{\substack{x\leq k\leq n\delta \\ 0\leq l\leq n\delta}} \{-\bar{\Gamma}_n(k,l) \geq 0\} \cup \bigcup_{\substack{x\leq k\leq n\delta \\ -n\delta\leq l<0}} \{-\bar{\Gamma}_n(k,l) \geq 0\} \\ &\cup \bigcup_{\substack{-n\delta\leq k\leq -x \\ 0\leq l\leq n\delta}} \{-\bar{\Gamma}_n(k,l) \geq 0\} \cup \bigcup_{\substack{-n\delta\leq k\leq -x \\ -n\delta\leq l<0}} \{-\bar{\Gamma}_n(k,l) \geq 0\} \\ =:& E^{(++)} \cup E^{(+-)} \cup E^{(-+)} \cup E^{(--)}. \end{aligned} \tag{3.24}$$

We describe our proceeding only for the estimate of the probability of $E^{(++)}$ in detail. It holds

$$\begin{aligned} E^{(++)} &= \bigcup_{x\leq k\leq n\delta} \bigcup_{0\leq l\leq n\delta} \{-\left(\bar{\Gamma}_{n,1}(k) + \bar{\Gamma}_{n,2}(l)\right) \geq 0\} \\ &\subseteq \left\{ \max_{x\leq k\leq n\delta} \sum_{i=1}^{k} -a_1(X_{\tau_n+i}) + \max_{0\leq l\leq n\delta} \sum_{i=1}^{l} -a_2(X_{\sigma_n+i}) \geq 0 \right\} \qquad \text{by Lem. 3.15} \\ &=: \left\{ Y_1^{(+)} + Y_2^{(+)} \geq 0 \right\}. \end{aligned} \tag{3.25}$$

By Lemma 3.19 (ii), we see that $\left(X_{\tau_n+1}, ..., X_{\tau_n+\lfloor n\delta\rfloor}\right)$ and $\left(X_{\sigma_n+1}, ..., X_{\sigma_n+\lfloor n\delta\rfloor}\right)$ are independent vectors. Thus,

$$Y_1^{(+)} = Y_1^{(+)}\left(X_{\tau_n+1}, ..., X_{\tau_n+\lfloor n\delta\rfloor}\right) \quad \text{and} \quad Y_2^{(+)} = Y_2^{(+)}\left(X_{\sigma_n+1}, ..., X_{\sigma_n+\lfloor n\delta\rfloor}\right)$$

as two measurable transformations of independent vectors are also independent. By (3.25) and Lemma A.8, we get

$$\mathbb{P}\left[E^{(++)}\right] = \mathbb{P}\left[Y_1^{(+)} + Y_2^{(+)} \geq 0\right] = \int_{\mathbb{R}} \mathbb{P}\left[Y_1^{(+)} \geq -y\right] \ \mathbb{P}_{Y_2^{(+)}}(\mathrm{d}y). \tag{3.26}$$

For abbreviation, we write $Z_{1,i} := -a_1(X_{\tau_n+i}) + \mathbb{E}[a_1(X_{\tau_n+i})]$, $1 \leq i \leq \lfloor n\delta \rfloor$. We next consider the integrand. For all $y \in \mathbb{R}$ we obtain

$$\begin{aligned}\mathbb{P}\left[Y_1^{(+)} \geq -y\right] &= \mathbb{P}\left[\max_{x \leq k \leq n\delta} \sum_{i=1}^{k} -a_1(X_{\tau_n+i}) \geq -y\right] \\ &= \mathbb{P}\left[\bigcup_{x \leq k \leq n\delta} \left\{\sum_{i=1}^{k} Z_{1,i} \geq \sum_{i=1}^{k} \mathbb{E}[a_1(X_{\tau_n+i})] - y\right\}\right].\end{aligned}$$

By Lemma 3.19 (ii), we conclude that $\tau_n + 1 \leq \tau_n + i < \sigma_n$ for $1 \leq i \leq k$ with $x \leq k \leq n\delta$. In the proof of Lemma 3.8 we have seen that

$$\mathbb{E}[a_1(X_{\tau_n+i})] = (\alpha - \beta)^2 \quad \text{for} \quad 1 \leq i \leq k \quad \text{with} \quad x \leq k \leq n\delta.$$

By model assumptions $\alpha \neq \beta$ and $\beta \neq \gamma$, it holds $\mu := \min\left\{(\alpha-\beta)^2, (\beta-\gamma)^2\right\} > 0$. It follows for all $y \in \mathbb{R}$ that

$$\begin{aligned}\mathbb{P}\left[Y_1^{(+)} \geq -y\right] &= \mathbb{P}\left[\bigcup_{x \leq k \leq n\delta} \left\{\sum_{i=1}^{k} Z_{1,i} \geq k(\alpha-\beta)^2 - y\right\}\right] \\ &\leq \mathbb{P}\left[\bigcup_{x \leq k \leq n\delta} \left\{\sum_{i=1}^{k} Z_{1,i} \geq k\mu - y\right\}\right] \\ &=: \mathbb{P}(y).\end{aligned} \tag{3.27}$$

We distinguish several cases for y to get an estimate for $\mathbb{P}(y)$.

(i) In the case $y \leq 0$ we have $-y \geq 0$. The independence of $X_1, ..., X_n$ leads to the independence of $Z_{1,1}, ..., Z_{1,k}$, $x \leq k \leq n\delta$. Furthermore, $(k^{-1})_{\lfloor x \rfloor + 1 \leq k \leq \lfloor n\delta \rfloor}$ is a non-increasing sequence of positive numbers. The Hájek-Rényi Inequality (Lemma 2.8) implies

$$
\begin{aligned}
\mathbb{P}(y) &\leq \mathbb{P}\left[\bigcup_{x\leq k\leq n\delta}\left\{\sum_{i=1}^{k} Z_{1,i} \geq k\mu\right\}\right] \\
&\leq \mathbb{P}\left[\max_{x\leq k\leq n\delta} k^{-1}\left|\sum_{i=1}^{k} Z_{1,i}\right| \geq \mu\right] \\
&\leq \mathbb{P}\left[\max_{\lfloor x\rfloor\leq k\leq \lfloor n\delta\rfloor} k^{-1}\left|\sum_{i=1}^{k} Z_{1,i}\right| \geq \mu\right] \\
&\leq \mu^{-2}\left(\lfloor x\rfloor^{-2}\sum_{k=1}^{\lfloor x\rfloor}\mathbb{V}[Z_{1,k}] + \sum_{k=\lfloor x\rfloor+1}^{\lfloor n\delta\rfloor} k^{-2}\mathbb{V}[Z_{1,k}]\right).
\end{aligned}
$$

By definition of $Z_{1,k}$ and Equation (3.2), it is evident that

$$
\mathbb{V}[Z_{1,k}] = 4(\alpha-\beta)^2\mathbb{V}[X_{\tau_n+k}] \leq 4(\alpha-\beta)^2 M_2
$$

for all $k \in \{1, ..., \lfloor n\delta\rfloor\}$. From $M_2 < \infty$ we conclude that

$$
\mathbb{P}(y) \leq C\left(\lfloor x\rfloor^{-1} + \sum_{i=\lfloor x\rfloor+1}^{\lfloor n\delta\rfloor} k^{-2}\right).
$$

Note that the properties of the floor function give $\lfloor x\rfloor \geq x-1 \geq \frac{1}{2}x$ for $x \geq 2$. Lemmas A.5 and A.1 (i) now yield

$$
\mathbb{P}(y) \leq C\lfloor x\rfloor^{-1} \leq Cx^{-1}.
$$

(ii) Let $y > 0$. By $k \geq x$, we have

$$
k\mu - y = k\left(\mu - \frac{y}{k}\right) \geq k\left(\mu - \frac{y}{x}\right).
$$

(a) Let $0 < y < \frac{1}{2}\mu x$. It follows that $k\mu - y \geq \frac{1}{2}k\mu$. As in (i), we obtain

$$
\begin{aligned}
\mathbb{P}(y) &\leq \mathbb{P}\left[\bigcup_{x\leq k\leq n\delta}\left\{\sum_{i=1}^{k} Z_{1,i} \geq \frac{1}{2}k\mu\right\}\right] \\
&\leq \mathbb{P}\left[\max_{\lfloor x\rfloor\leq k\leq \lfloor n\delta\rfloor} k^{-1}\left|\sum_{i=1}^{k} Z_{1,i}\right| \geq \frac{1}{2}\mu\right] \\
&\leq Cx^{-1}.
\end{aligned}
$$

(b) In the case $y \geq \frac{1}{2}\mu x$ we estimate $\mathbb{P}(y) \leq 1$.

Applying (3.26) and (3.27) with regard to the previous distinction of cases gives

$$\begin{aligned}
&\mathbb{P}\left[E^{(++)}\right] \\
&\leq \int_{(-\infty,0]} \mathbb{P}(y)\ \mathbb{P}_{Y_2^{(+)}}(\mathrm{d}y) + \int_{\left(0,\frac{1}{2}\mu x\right)} \mathbb{P}(y)\ \mathbb{P}_{Y_2^{(+)}}(\mathrm{d}y) + \int_{\left[\frac{1}{2}\mu x,\infty\right)} \mathbb{P}(y)\ \mathbb{P}_{Y_2^{(+)}}(\mathrm{d}y) \\
&\leq Cx^{-1}\mathbb{P}\left[Y_2^{(+)} \leq 0\right] + Cx^{-1}\mathbb{P}\left[0 < Y_2^{(+)} < \frac{1}{2}\mu x\right] + \mathbb{P}\left[Y_2^{(+)} \geq \frac{1}{2}\mu x\right] \\
&\leq Cx^{-1} + \mathbb{P}\left[Y_2^{(+)} \geq \frac{1}{2}\mu x\right]. \qquad (3.28)
\end{aligned}$$

For abbreviation, we write $Z_{2,i} := -a_2(X_{\sigma_n+i}) + \mathbb{E}[a_2(X_{\sigma_n+i})]$, $1 \leq i \leq \lfloor n\delta \rfloor$. We now handle the probability in the last estimate. By definition, we have

$$\begin{aligned}
\left\{Y_2^{(+)} \geq \frac{1}{2}\mu x\right\} &= \left\{\max_{0\leq l\leq n\delta} \sum_{i=1}^{l} -a_2(X_{\sigma_n+i}) \geq \frac{1}{2}\mu x\right\} \\
&= \bigcup_{0\leq l\leq n\delta} \left\{\sum_{i=1}^{l} Z_{2,i} \geq \sum_{i=1}^{l} \mathbb{E}[a_2(X_{\sigma_n+i})] + \frac{1}{2}\mu x\right\}.
\end{aligned}$$

Note that $\left\{\sum_{i=1}^{l} Z_{2,i} \geq \sum_{i=1}^{l} \mathbb{E}[a_2(X_{\sigma_n+i})] + \frac{1}{2}\mu x\right\} = \emptyset$ for $l = 0$, because $\frac{1}{2}\mu x > 0$. From Lemma 3.19 (iii) we deduce that $\sigma_n + 1 \leq \sigma_n + i < n$ for $1 \leq i \leq l$ with $1 \leq l \leq n\delta$. The proof of Lemma 3.8 provides

$$\mathbb{E}[a_2(X_{\sigma_n+i})] = (\beta - \gamma)^2 \quad \text{for} \quad 1 \leq i \leq l \quad \text{with} \quad 1 \leq l \leq n\delta.$$

It follows that

$$\begin{aligned}
\left\{Y_2^{(+)} \geq \frac{1}{2}\mu x\right\} &= \bigcup_{1\leq l\leq n\delta} \left\{\sum_{i=1}^{l} Z_{2,i} \geq l(\beta-\gamma)^2 + \frac{1}{2}\mu x\right\} \\
&\subseteq \bigcup_{1\leq l\leq n\delta} \left\{\sum_{i=1}^{l} Z_{2,i} \geq l\mu + \frac{1}{2}\mu x\right\} \\
&\subseteq \bigcup_{1\leq l\leq n\delta} \left\{(2l+x)^{-1}\left|\sum_{i=1}^{l} Z_{2,i}\right| \geq \frac{1}{2}\mu\right\}.
\end{aligned}$$

Observe that $Z_{2,1}, ..., Z_{2,l}$, $1 \leq l \leq n\delta$, are independent and $(2l+x)^{-1}_{1\leq l\leq \lfloor n\delta \rfloor}$ is a non-increasing sequence of positive numbers for any $x > 0$. Another application of the Hájek-Rényi Inequality (Lemma 2.8) yields

$$\begin{aligned}\mathbb{P}\left[Y_2^{(+)} \geq \frac{1}{2}\mu x\right] &\leq \mathbb{P}\left[\bigcup_{1\leq l\leq n\delta}\left\{(2l+x)^{-1}\left|\sum_{i=1}^{l} Z_{2,i}\right| \geq \frac{1}{2}\mu\right\}\right] \\ &= \mathbb{P}\left[\max_{1\leq l\leq \lfloor n\delta\rfloor}(2l+x)^{-1}\left|\sum_{i=1}^{l} Z_{2,i}\right| \geq \frac{1}{2}\mu\right] \\ &\leq 4\mu^{-2}\sum_{l=1}^{\lfloor n\delta\rfloor}(2l+x)^{-2}\mathbb{V}[Z_{2,l}].\end{aligned}$$

By definition of $Z_{l,2}$ and Equation (3.2), it is clear that

$$\mathbb{V}[Z_{2,l}] = 4(\beta-\gamma)^2\mathbb{V}[X_{\sigma_n+l}] \leq 4(\beta-\gamma)^2 M_2$$

for all $l \in \{1, ..., \lfloor n\delta\rfloor\}$. Note that we have $2l + x \geq l + 1 + \lfloor x\rfloor$ for all $l \in \{1, ..., \lfloor n\delta\rfloor\}$. We thus get

$$\begin{aligned}\mathbb{P}\left[Y_2^{(+)} \geq \frac{1}{2}\mu x\right] &\leq C\sum_{l=1}^{\lfloor n\delta\rfloor}(l+1+\lfloor x\rfloor)^{-2} && \text{by } 2l+x \geq l+1+\lfloor x\rfloor,\ M_2<\infty \\ &= C\sum_{m=\lfloor x\rfloor+2}^{\lfloor n\delta\rfloor+\lfloor x\rfloor+1} m^{-2} \\ &\leq C(\lfloor x\rfloor+1)^{-1} && \text{by Lem. A.5} \\ &\leq Cx^{-1}. && \text{by Lem. A.1 (i)}\end{aligned}$$

Summarizing, by (3.28), we have

$$\mathbb{P}\left[E^{(++)}\right] \leq Cx^{-1}.$$

The rest of the proof runs as before. We outline the proof for $E^{(+-)}$, $E^{(-+)}$ and $E^{(--)}$. Set

$$Y_1^{(-)} := \max_{-n\delta\leq k\leq -x}\sum_{i=1}^{-k} a_1(X_{\tau_n-i+1}) \quad \text{and} \quad Y_2^{(-)} := \max_{-n\delta\leq l<0}\sum_{i=1}^{-l} a_2(X_{\sigma_n-i+1}).$$

By Equation (3.24) and Lemma 3.15, it holds

$$E^{(+-)} = \left\{Y_1^{(+)} + Y_2^{(-)} \geq 0\right\}, \quad E^{(-+)} = \left\{Y_1^{(-)} + Y_2^{(+)} \geq 0\right\} \quad \text{and}$$
$$E^{(--)} = \left\{Y_1^{(-)} + Y_2^{(-)} \geq 0\right\}.$$

The pairwise independence of the measurable transformations

$$Y_1^{(+)} = Y_1^{(+)}\left(X_{\tau_n+1}, ..., X_{\tau_n+\lfloor n\delta \rfloor}\right) \quad \text{and} \quad Y_2^{(-)} = Y_2^{(-)}\left(X_{\sigma_n-\lfloor n\delta \rfloor+1}, ..., X_{\sigma_n}\right),$$
$$Y_1^{(-)} = Y_1^{(-)}\left(X_{\tau_n-\lfloor n\delta \rfloor+1}, ..., X_{\tau_n}\right) \quad \text{and} \quad Y_2^{(+)} = Y_2^{(+)}\left(X_{\sigma_n+1}, ..., X_{\sigma_n+\lfloor n\delta \rfloor}\right),$$
$$Y_1^{(-)} = Y_1^{(-)}\left(X_{\tau_n-\lfloor n\delta \rfloor+1}, ..., X_{\tau_n}\right) \quad \text{and} \quad Y_2^{(-)} = Y_2^{(-)}\left(X_{\sigma_n-\lfloor n\delta \rfloor+1}, ..., X_{\sigma_n}\right)$$

follow from Lemma 3.19 and the independence of the observations $X_1, ..., X_n$. Lemma 3.19 shows that $1 \leq \tau_n - i + 1 \leq \tau_n$ for $1 \leq i \leq -k$ with $x \leq -k \leq n\delta$ and $\tau_n + 1 \leq \sigma_n - i + 1 \leq \sigma_n$ for $1 \leq i \leq -l$ with $1 \leq -l \leq n\delta$. The proof of Lemma 3.8 establishes

$$\mathbb{E}[a_1(X_{\tau_n-i+1})] = -(\alpha - \beta)^2 \quad \text{for} \quad 1 \leq i \leq -k \quad \text{with} \quad x \leq -k \leq n\delta \quad \text{and}$$
$$\mathbb{E}[a_2(X_{\sigma_n-i+1})] = -(\beta - \gamma)^2 \quad \text{for} \quad 1 \leq i \leq -l \quad \text{with} \quad 1 \leq -l \leq n\delta.$$

Similar arguments used in the estimate of the probability of $E^{(++)}$ lead to

$$\mathbb{P}\left[E^{(+-)}\right] \leq Cx^{-1}, \quad \mathbb{P}\left[E^{(-+)}\right] \leq Cx^{-1} \quad \text{and} \quad \mathbb{P}\left[E^{(--)}\right] \leq Cx^{-1}.$$

Applying (3.24) yields

$$\mathbb{P}[E] \leq \mathbb{P}\left[E^{(++)}\right] + \mathbb{P}\left[E^{(+-)}\right] + \mathbb{P}\left[E^{(-+)}\right] + \mathbb{P}\left[E^{(--)}\right] \leq Cx^{-1}.$$

In the same manner we can see that

$$\mathbb{P}[F] \leq Cx^{-1}.$$

Altogether, by (3.23), we have

$$\mathbb{P}[x \leq \|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta] \leq \mathbb{P}[E] + \mathbb{P}[F] \leq Cx^{-1},$$

which is our claim. □

We obtain the stochastic boundedness of $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$.

Proposition 3.22. *If $M_2 < \infty$, then*

$$\lim_{x\to\infty} \limsup_{n\to\infty} \mathbb{P}[\|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \geq x] = 0.$$

Proof. By Lemma 3.21, there exist $n_0 \in \mathbb{N}$, $\delta > 0$ and a constant $C > 0$ such that

$$\begin{aligned}\mathbb{P}[\|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \geq x] &\leq \mathbb{P}[x \leq \|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta] + \mathbb{P}[\|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| > n\delta] \\ &\leq Cx^{-1} + \mathbb{P}[\|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| > n\delta] \qquad (3.29)\end{aligned}$$

for all $x \geq 2$ and $n \geq n_0$. Fix for a moment $n \geq n_0$. Furthermore, we conclude that

$$\begin{aligned}&\mathbb{P}[\|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| > n\delta] \\ &\quad = \mathbb{P}\left[\|\bar{\boldsymbol{\tau}}_n - n\boldsymbol{\theta} + n\boldsymbol{\theta} - \boldsymbol{\tau}_n\| > n\delta\right] \\ &\quad \leq \mathbb{P}\left[\|\bar{\boldsymbol{\tau}}_n - n\boldsymbol{\theta}\| + \|n\boldsymbol{\theta} - \boldsymbol{\tau}_n\| > n\delta\right] && \text{by Tr. In.} \\ &\quad \leq \mathbb{P}\left[\|\bar{\boldsymbol{\tau}}_n - n\boldsymbol{\theta}\| > \frac{1}{2}n\delta\right] + \mathbb{P}\left[\|n\boldsymbol{\theta} - \boldsymbol{\tau}_n\| > \frac{1}{2}n\delta\right] && \text{by Lem. A.4.} \qquad (3.30)\end{aligned}$$

By Lemma A.1 (i), we get

$$\|n\boldsymbol{\theta} - \boldsymbol{\tau}_n\| = \max\{n\theta_1 - \lfloor n\theta_1 \rfloor, n\theta_2 - \lfloor n\theta_2 \rfloor\} \leq 1.$$

Accordingly, by definition of $\bar{\boldsymbol{\theta}}_n$ (see (3.5)), it follows that

$$\mathbb{P}[\|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| > n\delta] \leq \mathbb{P}\left[\left\|\bar{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right\| > \frac{1}{2}\delta\right] + \mathbb{P}\left[1 > \frac{1}{2}n\delta\right].$$

By (3.29), we infer that

$$\mathbb{P}[\|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \geq x] \leq Cx^{-1} + \mathbb{P}\left[\left\|\bar{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right\| > \frac{1}{2}\delta\right] + \mathbb{P}\left[1 > \frac{1}{2}n\delta\right].$$

The weak consistency of $\bar{\boldsymbol{\theta}}_n$, given in Theorem 3.12, implies

$$\limsup_{n\to\infty} \mathbb{P}[\|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \geq x] \leq Cx^{-1}.$$

Letting $x \to \infty$ completes the proof. □

To ensure convergence in distribution of $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$, we look closer at the limit process Γ. The aim is to find conditions such that Γ has almost surely an unique minimizer (assumption (iii) of Theorem 2.3). We begin by considering the minimizers of Γ_1 and Γ_2. After this we establish the relation to the minimizers of Γ.

Lemma 3.23. *There exist a minimizer of Γ_1 and a minimizer of Γ_2 almost surely.*

Proof. We only show the assertion for Γ_1. The same approach can be applied to Γ_2. Set

$$Y_i := Y_i(k) := \begin{cases} 2(\beta-\alpha)\,(\xi_{i,2}-\beta)+(\alpha-\beta)^2, & k \geq 0, \\ -2(\beta-\alpha)\,(\xi_{i,1}-\alpha)+(\alpha-\beta)^2, & k < 0 \end{cases}$$

for $i \in \mathbb{N}$. By Remark 3.16, for each $k \in \mathbb{Z}$, Γ_1 can be written as

$$\Gamma_1(k) = \sum_{i=1}^{|k|} Y_i.$$

Note that Y_i, $i \in \mathbb{N}$, are independent and identically distributed, which follow from the assumptions to the sequences $(\xi_{i,1})_{i\in\mathbb{N}}$ and $(\xi_{i,2})_{i\in\mathbb{N}}$. The Strong Law of Large Numbers (see for instance Schmidt [24, p. 347, Theorem 15.2.7]) yields

$$\tfrac{1}{|k|}\Gamma_1(k) = \tfrac{1}{|k|}\sum_{i=1}^{|k|} Y_i \xrightarrow[|k|\to\infty]{a.s.} \mathbb{E}[Y_1] = (\alpha-\beta)^2.$$

By model assumption $\alpha \neq \beta$, we have $(\alpha-\beta)^2 > 0$. By $\frac{1}{|k|} \xrightarrow[|k|\to\infty]{} 0$, it follows that

$$\Gamma_1(k) \xrightarrow[|k|\to\infty]{a.s.} \infty.$$

We thus get $\operatorname{Argmin}(\Gamma_1) \neq \emptyset$ almost surely. □

Lemma 3.24. *Let Q_1, Q_2 and Q_3 be continuous distributions. Then each process Γ_1 and Γ_2 has an unique minimizer almost surely.*

Proof. We only show the claim for Γ_1 again. A similar approach can be applied to Γ_2. Lemma 3.23 and similar arguments used in the proof of Lemma 3.3 lead to

$$\mathbb{P}[|\operatorname{Argmin}(\Gamma_1)| = 1] \geq 1 - \sum_{k_1 \neq k_2 \in \mathbb{Z}} \mathbb{P}[\Gamma_1(k_1) - \Gamma_1(k_2) = 0]. \tag{3.31}$$

We discuss the cases $k_1 > k_2$ and $k_2 > k_1$ to compute $\Gamma_1(k_1) - \Gamma_1(k_2)$ by Equation (3.18).

(i) (a) Let $k_1 > k_2 > 0$. Then

$$\Gamma_1(k_1) - \Gamma_1(k_2) = \sum_{i=1}^{k_1} a_1(\xi_{i,2}) - \sum_{i=1}^{k_2} a_1(\xi_{i,2}) = \sum_{i=k_2+1}^{k_1} a_1(\xi_{i,2}).$$

(b) Let $k_1 > 0 > k_2$. Then

$$\Gamma_1(k_1) - \Gamma_1(k_2) = \sum_{i=1}^{k_1} a_1(\xi_{i,2}) + \sum_{i=1}^{-k_2} a_1(\xi_{i,1}).$$

(c) Let $0 > k_1 > k_2$. Then

$$\Gamma_1(k_1) - \Gamma_1(k_2) = -\sum_{i=1}^{-k_1} a_1(\xi_{i,1}) + \sum_{i=1}^{-k_2} a_1(\xi_{i,1}) = \sum_{i=-k_1+1}^{-k_2} a_1(\xi_{i,1}).$$

(ii) (a) Let $k_2 > k_1 > 0$. Then

$$\Gamma_1(k_1) - \Gamma_1(k_2) = \sum_{i=1}^{k_1} a_1(\xi_{i,2}) - \sum_{i=1}^{k_2} a_1(\xi_{i,2}) = -\sum_{i=k_1+1}^{k_2} a_1(\xi_{i,2}).$$

(b) Let $k_2 > 0 > k_1$. Then

$$\Gamma_1(k_1) - \Gamma_1(k_2) = -\sum_{i=1}^{-k_1} a_1(\xi_{i,1}) - \sum_{i=1}^{k_2} a_1(\xi_{i,2}).$$

(c) Let $0 > k_2 > k_1$. Then

$$\Gamma_1(k_1) - \Gamma_1(k_2) = -\sum_{i=1}^{-k_1} a_1(\xi_{i,1}) + \sum_{i=1}^{-k_2} a_1(\xi_{i,1}) = -\sum_{i=-k_2+1}^{-k_1} a_1(\xi_{i,1}).$$

By the independence assumptions to the sequences $(\xi_{i,1})_{i\in\mathbb{N}}$ and $(\xi_{i,2})_{i\in\mathbb{N}}$, we obtain sums of independent random variables in each case. Since Q_1, Q_2 and Q_3 are continuous distributions, we can conclude by convolution and definition of a_1 (see (3.2)) that $\Gamma_1(k_1)-\Gamma_1(k_2)$ are continuous distributed random variables for each $k_1, k_2 \in \mathbb{Z}$ with $k_1 \neq k_2$, and consequently $\mathbb{P}[\Gamma_1(k_1) - \Gamma_1(k_2) = 0] = 0$ for all $k_1, k_2 \in \mathbb{Z}$ with $k_1 \neq k_2$. By (3.31), we have $\mathbb{P}[|\operatorname{Argmin}(\Gamma_1)| = 1] = 1$. □

Lemma 3.25. *It holds*

$$\operatorname{Argmin}(\Gamma) = \operatorname{Argmin}(\Gamma_1) \times \operatorname{Argmin}(\Gamma_2).$$

Proof. The proof is straightforward.

(i) We first prove that $(m_1, m_2) \in \operatorname{Argmin}(\Gamma)$ implies $m_1 \in \operatorname{Argmin}(\Gamma_1)$ and $m_2 \in \operatorname{Argmin}(\Gamma_2)$. Fix $(m_1, m_2) \in \operatorname{Argmin}(\Gamma)$. As defined in (3.18), it follows that

$$\Gamma_1(m_1) + \Gamma_2(m_2) = \Gamma(m_1, m_2) \leq \Gamma(k, l) = \Gamma_1(k) + \Gamma_2(l)$$

for all $(k, l) \in \mathbb{Z}^2$. We choose $l := m_2$ to get $\Gamma_1(m_1) \leq \Gamma_1(k)$ for all $k \in \mathbb{Z}$ or $k := m_1$ to see $\Gamma_2(m_2) \leq \Gamma_2(l)$ for all $l \in \mathbb{Z}$. Hence $m_1 \in \operatorname{Argmin}(\Gamma_1)$ and $m_2 \in \operatorname{Argmin}(\Gamma_2)$.

(ii) Fix $m_1 \in \operatorname{Argmin}(\Gamma_1)$ and $m_2 \in \operatorname{Argmin}(\Gamma_2)$. To deduce $(m_1, m_2) \in \operatorname{Argmin}(\Gamma)$, observe by Equation (3.18) that

$$\Gamma(m_1, m_2) = \Gamma_1(m_1) + \Gamma_2(m_2) \leq \Gamma_1(k) + \Gamma_2(l) = \Gamma(k, l)$$

for all $(k, l) \in \mathbb{Z}^2$. □

Proposition 3.26. *Let Q_1, Q_2 and Q_3 be continuous distributions. Then Γ has an unique minimizer almost surely.*

Proof. Combining Lemma 3.24 with Lemma 3.25 gives the assertion. □

We are now in a position to show that $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$ converges in distribution to the minimizer of a sum of random walks if the underlying distributions are continuous.

Theorem 3.27. *If $M_2 < \infty$, then*

$$\limsup_{n\to\infty} \mathbb{P}[\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n \in F] \leq \mathbb{P}[\mathrm{Argmin}(\Gamma) \cap F \neq \emptyset] \qquad \textit{for all } F \subseteq \mathbb{Z}^2.$$

In addition, if Q_1, Q_2 and Q_3 are continuous, then $\mathrm{Argmin}(\Gamma) = \{\boldsymbol{T}\}$ *almost surely and*

$$\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n \xrightarrow[n\to\infty]{\mathcal{L}} \boldsymbol{T} \quad \textit{in} \quad \mathbb{Z}^2.$$

Proof. We apply Theorem 2.3. By Lemmas 3.14, 3.23 and 3.25, we first observe that $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$ is a minimizer of $\bar{\Gamma}_n$ and Γ has at least one minimizer. Assumptions (i) and (ii) of Theorem 2.3 are fulfilled by Propositions 3.18 and 3.22, which give the first claim. The second claim is obtained by applying Proposition 3.26. □

Corollary 3.28. *Suppose that $M_2 < \infty$. Let Q_1, Q_2 and Q_3 be continuous distributions. Then*

$$\bar{\tau}_n - \tau_n \xrightarrow[n\to\infty]{\mathcal{L}} \operatorname*{argmin}_{k\in\mathbb{Z}} \Gamma_1(k) \quad \textit{in} \quad \mathbb{Z} \qquad \textit{and} \qquad \bar{\sigma}_n - \sigma_n \xrightarrow[n\to\infty]{\mathcal{L}} \operatorname*{argmin}_{l\in\mathbb{Z}} \Gamma_2(l) \quad \textit{in} \quad \mathbb{Z}.$$

Proof. By Theorem 3.27 and Lemma 3.25, we get

$$(\bar{\tau}_n - \tau_n, \bar{\sigma}_n - \sigma_n) = \bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n \xrightarrow[n\to\infty]{\mathcal{L}} \operatorname*{argmin}_{(k,l)\in\mathbb{Z}^2} \Gamma(k,l) = \left(\operatorname*{argmin}_{k\in\mathbb{Z}} \Gamma_1(k), \operatorname*{argmin}_{l\in\mathbb{Z}} \Gamma_2(l)\right) \quad \text{in} \quad \mathbb{Z}^2.$$

Since the projections are continuous, the assertion follows from the Continuous Mapping Theorem (see for instance Van der Vaart [26, p. 7, Theorem 2.3]). □

3.4 Asymptotic confidence region

As an application of Theorem 3.27, this section is intended to present an asymptotic confidence region to estimate the moments of change $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$.

The statistician is interested in finding a preferably small (asymptotic) confidence region. For this purpose, let $F^{-1}_{\|\boldsymbol{T}\|}(\vartheta)$, $\vartheta \in (0,1)$, stand for the ϑ-quantile of the distribution function $F_{\|\boldsymbol{T}\|}$ of $\|\boldsymbol{T}\|$, where $\boldsymbol{T}$ is the almost surely unique minimizer of Γ (see Theorem 3.27).

Based on Theorem 3.27 and the Continuous Mapping Theorem, we derive an asymptotic confidence region.

Theorem 3.29. *Suppose that $M_2 < \infty$. Let Q_1, Q_2 and Q_3 be continuous distributions and $\vartheta \in (0,1)$. For each $n \in \mathbb{N}$, the random interval*

$$I_n(\vartheta) := \left[\bar{\tau}_n - F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta), \bar{\tau}_n + F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta)\right] \times \left[\bar{\sigma}_n - F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta), \bar{\sigma}_n + F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta)\right]$$

is an asymptotic confidence region for $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ at level $1-\vartheta$.

Proof. Fix $\vartheta \in (0,1)$. Since the maximum norm is non-negative and continuous on $\mathbb{Z}^2$, by Theorem 3.27 and the Continuous Mapping Theorem (see for instance Van der Vaart [26, p. 7, Theorem 2.3]), we conclude that

$$\|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \xrightarrow[n\to\infty]{\mathcal{L}} \|\boldsymbol{T}\| \quad \text{in} \quad \mathbb{N}_0.$$

Since $\|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\|$ and $\|\boldsymbol{T}\|$ are discrete random variables, we obtain

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}[\boldsymbol{\tau}_n \in I_n(\vartheta)] &= \lim_{n\to\infty} \mathbb{P}\left[\max\{|\bar{\tau}_n - \tau_n|, |\bar{\sigma}_n - \sigma_n|\} \le F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta)\right] \\
&= \lim_{n\to\infty} \mathbb{P}\left[\|\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \le F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta)\right] \\
&= \mathbb{P}\left[\|\boldsymbol{T}\| \le F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta)\right] \\
&\ge 1-\vartheta.
\end{aligned}$$

□

Observe that the quantile $F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta)$, $\vartheta \in (0,1)$, which is used in the theorem above, is unknown. Though, it can be approximated by a Monte-Carlo simulation:

(i) Generate $N \in \mathbb{N}$ processes $\Gamma^{(1)}, ..., \Gamma^{(N)}$ as defined in (3.18):

(1) Determine $\xi_{i,r}$, $i \in \mathbb{N}$, $1 \le r \le 3$, based on bootstrap method:

(a) Generate $m \in \mathbb{N}$ independent random variables $U_1, ..., U_m \sim U(0,1)$.

(b) Let $G_{m,1}$ and $G_{m,2}$ and $G_{m,3}$ be the empirical distribution functions pertaining to $X_1, ..., X_{\bar{\tau}_n}$ and $X_{\bar{\tau}_n+1}, ..., X_{\bar{\sigma}_n}$ and $X_{\bar{\sigma}_n+1}, ..., X_n$. For each $1 \le i \le m$ put $\xi_{i,r} := G^{-1}_{m,r}(U_i)$, $1 \le r \le 3$.

(2) Use the known expectations α, β and γ to compute $a_1(\xi_{i,r})$, $1 \le i \le m$, $1 \le r \le 2$, and $a_2(\xi_{i,r})$, $1 \le i \le m$, $2 \le r \le 3$, as defined in (3.2).

(ii) For each $1 \le i \le N$ compute $\boldsymbol{T}^{(i)} := \operatorname{argmin}_{(k,l)\in\{-m,...,m\}^2} \Gamma^{(i)}(k,l)$ by Lemma 3.25.

(iii) Let H_N be the empirical distribution function pertaining to $\left\|\boldsymbol{T}^{(1)}\right\|, ..., \left\|\boldsymbol{T}^{(N)}\right\|$. Then $H_N^{-1}(1-\vartheta)$ is a reasonable estimate for $F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta)$.

For further investigation of the asymptotic confidence region based on a simulation study, we refer the reader to Chapter 5. Here, numerous asymptotic confidence regions are implemented to determine the approximated coverage probability.

Chapter 4

Unknown expectations

In this chapter we proceed with the estimation of the multiple change-point in a more general setting. From now on, the expectations $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)$ are assumed to be unknown. At first we simultaneously estimate the multiple change-point and the expectations by the least squares method. Furthermore, weak consistency of the resulting estimators is proved. The next section is devoted to the introduction of another estimator of the multiple change-point. We state and prove consistency and convergence in distribution, which are the main results of this work. Finally, an asymptotic confidence region for the moments of change $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ is derived.

4.1 Parameter estimation

Here and subsequently, we let $\bar{X}_{u,v}$, $u, v \in \mathbb{N}_0$ with $u < v \leq n$, stand for

$$\bar{X}_{u,v} := \frac{1}{v-u} \sum_{i=u+1}^{v} X_i.$$

4.1.1 Estimation approach

At first our focus lies on the simultaneous estimation of the moments of change $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ and the expectations $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)$ by the least squares method. To do this, we are interested in finding all minimizing points of the random criterion function

$$S_n(k,l,a,b,c) := \sum_{i=1}^{k}(X_i - a)^2 + \sum_{i=k+1}^{l}(X_i - b)^2 + \sum_{i=l+1}^{n}(X_i - c)^2, \quad (k,l) \in \Delta_n, \quad (a,b,c) \in \mathbb{R}^3. \tag{4.1}$$

To solve this problem on a simple way, let us introduce the random criterion function

$$\hat{M}_n(k,l) := k\bar{X}_{0,k}^2 + (l-k)\bar{X}_{k,l}^2 + (n-l)\bar{X}_{l,n}^2, \quad (k,l) \in \Delta_n.$$

Roughly speaking, S_n can be minimized by maximizers of $\hat{M}_n$ and means of segments of $X_1, ..., X_n$, where the borders of each segment are obtained by the maximizers of $\hat{M}_n$.

Theorem 4.1. *Let $n \in \mathbb{N}$. Then*

$$\operatorname{Argmin}(S_n) = \left\{ \left(\hat{k}_n, \hat{l}_n, \bar{X}_{0,\hat{k}_n}, \bar{X}_{\hat{k}_n,\hat{l}_n}, \bar{X}_{\hat{l}_n,n}\right) \in \Delta_n \times \mathbb{R}^3 \,\middle|\, \left(\hat{k}_n, \hat{l}_n\right) \in \operatorname{Argmax}\left(\hat{M}_n\right) \right\}.$$

The following lemma is essential for the proof.

Lemma 4.2 (Interchange of order of minimization)**.** *Let $p, q \in \mathbb{N}$. For any sets $A \subseteq \mathbb{R}^p$ and $B \subseteq \mathbb{R}^q$ let $f : A \times B \to \bar{\mathbb{R}}$ be a mapping. Set*

$$\tilde{f}_B(\mathbf{a}) := \inf_{\mathbf{b}\in B} f(\mathbf{a},\mathbf{b}) \quad \text{and} \quad \tilde{f}_A(\mathbf{b}) := \inf_{\mathbf{a}\in A} f(\mathbf{a},\mathbf{b}).$$

Then

$$\inf_{(\mathbf{a},\mathbf{b})\in A\times B} f(\mathbf{a},\mathbf{b}) = \inf_{\mathbf{a}\in A} \tilde{f}_B(\mathbf{a}) = \inf_{\mathbf{b}\in B} \tilde{f}_A(\mathbf{b})$$

and

$$\begin{aligned}\operatorname{Arginf}(f) &= \left\{ \left(\bar{\mathbf{a}}, \bar{\mathbf{b}}\right) \in A \times B \,\middle|\, \bar{\mathbf{a}} \in \operatorname*{Arginf}_{\mathbf{a}\in A} \tilde{f}_B(\mathbf{a}),\ \bar{\mathbf{b}} \in \operatorname*{Arginf}_{\mathbf{b}\in B} f\left(\bar{\mathbf{a}}, \mathbf{b}\right) \right\} \\ &= \left\{ \left(\bar{\mathbf{a}}, \bar{\mathbf{b}}\right) \in A \times B \,\middle|\, \bar{\mathbf{b}} \in \operatorname*{Arginf}_{\mathbf{b}\in B} \tilde{f}_A(\mathbf{b}),\ \bar{\mathbf{a}} \in \operatorname*{Arginf}_{\mathbf{a}\in A} f\left(\mathbf{a}, \bar{\mathbf{b}}\right) \right\}.\end{aligned}$$

Proof. An appropriate assertion for supremizing problems was shown in Albrecht [1, p. 59, A.1]. A similar proof can be used for infimizing problems. □

Proof of Theorem 4.1. The proof is divided into two steps. Fix $n \in \mathbb{N}$ and write

$$\hat{S}_n(k,l) := \sum_{i=1}^{k} \left(X_i - \bar{X}_{0,k}\right)^2 + \sum_{i=k+1}^{l} \left(X_i - \bar{X}_{k,l}\right)^2 + \sum_{i=l+1}^{n} \left(X_i - \bar{X}_{l,n}\right)^2, \quad (k,l) \in \Delta_n.$$

(i) We begin by proving

$$\begin{aligned} &\operatorname{Argmin}(S_n) \\ &= \left\{ \left(\hat{k}_n, \hat{l}_n, \bar{X}_{0,\hat{k}_n}, \bar{X}_{\hat{k}_n,\hat{l}_n}, \bar{X}_{\hat{l}_n,n}\right) \in \Delta_n \times \mathbb{R}^3 \,\middle|\, \left(\hat{k}_n, \hat{l}_n\right) \in \operatorname{Argmin}\left(\hat{S}_n\right) \right\} \end{aligned} \tag{4.2}$$

based on Lemma 4.2. Fix $(k,l) \in \Delta_n$ and consider $S_n(k,l,\cdot)$ as a function on $\mathbb{R}^3$ at first. An easy computation of the gradient and the Hessian matrix shows that

$$\nabla S_n(k,l,a,b,c) = \begin{pmatrix} -2\sum_{i=1}^{k} X_i + 2ka \\ -2\sum_{i=k+1}^{l} X_i + 2(l-k)b \\ -2\sum_{i=l+1}^{n} X_i + 2(n-l)c \end{pmatrix} \quad \text{and}$$

$$\nabla^2 S_n(k,l,a,b,c) = \begin{pmatrix} 2k & 0 & 0 \\ 0 & 2(l-k) & 0 \\ 0 & 0 & 2(n-l) \end{pmatrix}.$$

For all $\mathbf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3 \backslash \{0\}$ we have

$$\mathbf{y}^T \nabla^2 S_n(k,l,a,b,c)\mathbf{y} = 2\left(ky_1^2 + (l-k)y_2^2 + (n-l)y_3^2\right) > 0$$

by $(k,l) \in \Delta_n$. Accordingly, the Hessian matrix is positive-definit. Therefore, $S_n(k,l,\cdot)$ is strictly convex. We are able to conclude that $S_n(k,l,\cdot)$ has an unique minimizer. Using the necessary condition $\nabla S_n(k,l,a,b,c) = 0$ we get

$$\left(\bar{X}_{0,k}, \bar{X}_{k,l}, \bar{X}_{l,n}\right) = \underset{(a,b,c)\in\mathbb{R}^3}{\operatorname{argmin}}\, S_n(k,l,\cdot). \tag{4.3}$$

Altogether, by definitions, we have

$$\min_{(a,b,c)\in\mathbb{R}^3} S_n(k,l,a,b,c) = S_n\left(k,l,\bar{X}_{0,k},\bar{X}_{k,l},\bar{X}_{l,n}\right) = \hat{S}_n(k,l)$$

for all $(k,l)\in\Delta_n$. We now apply Lemma 4.2. By definition of $\hat{S}_n$, it is clear that we can find an element $\left(\hat{k}_n,\hat{l}_n\right)\in \operatorname{Argmin}\left(\hat{S}_n\right)$. By (4.3), we have

$$\left(\bar{X}_{0,\hat{k}_n},\bar{X}_{\hat{k}_n,\hat{l}_n},\bar{X}_{\hat{l}_n,n}\right) = \operatorname*{argmin}_{(a,b,c)\in\mathbb{R}^3} S_n\left(\hat{k}_n,\hat{l}_n,\cdot\right),$$

which proves the claim (4.2) .

(ii) In the second part we show the theorem. By the Binomial Formula, a simple calculation yields

$$\hat{S}_n(k,l) = \sum_{i=1}^{n} X_i^2 - \hat{M}_n(k,l)$$

for all $(k,l)\in\Delta_n$. Since the sum does not depend on $(k,l)\in\Delta_n$, we have

$$\operatorname{Argmin}\left(\hat{S}_n\right) = \operatorname{Argmax}\left(\hat{M}_n\right).$$

The proof is completed by combining this with (4.2). □

Similarly to Equation (3.3), we use a choice function $\hat{\phi}:\operatorname{Argmin}\left(\hat{M}_n\right)\longrightarrow\Delta_n$ if more than one minimizing point of $\hat{M}_n$ exists. Write

$$\hat{\boldsymbol{\tau}}_n := (\hat{\tau}_n,\hat{\sigma}_n) := \operatorname*{argmax}_{(k,l)\in\Delta_n} \hat{M}_n(k,l) \tag{4.4}$$

for $\hat{\boldsymbol{\tau}}_n = \hat{\phi}\left(\operatorname{Argmin}\left(\hat{M}_n\right)\right)$. Furthermore, let

$$\hat{\boldsymbol{\alpha}}_n := \left(\hat{\alpha}_n,\hat{\beta}_n,\hat{\gamma}_n\right) := \left(\bar{X}_{0,\hat{\tau}_n},\bar{X}_{\hat{\tau}_n,\hat{\sigma}_n},\bar{X}_{\hat{\sigma}_n,n}\right). \tag{4.5}$$

According to Theorem 4.1, the parameter vector $(\boldsymbol{\tau}_n,\boldsymbol{\alpha})\in\Delta_n\times\mathbb{R}^3$ can be estimated by the least squares estimator

$$(\hat{\boldsymbol{\tau}}_n,\hat{\boldsymbol{\alpha}}_n)\in\operatorname{Argmin}(S_n). \tag{4.6}$$

Remark 4.3. Albrecht [1] has investigated the estimation of multiple change-points in normal distribution models with changes in mean (variance is constant). The maximum-likelihood method was applied to estimate the moments of change $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ and the unknown expectations $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)$ simultaneously. Now, it turns out that the maximum-likelihood estimator in this parametric model and the least squares estimator $(\hat{\boldsymbol{\tau}}_n, \hat{\boldsymbol{\alpha}}_n)$ in our non-parametric model are identical.

A simulation study (see Chapter 5 for more details) gives the conjecture that $\hat{M}_n$ has an unique minimizer if all distributions are continuous.

Conjecture 4.4. Let $n \in \mathbb{N}$ and let Q_1, Q_2, Q_3 be continuous distributions. Then

$$\left|\text{Argmax}\left(\hat{M}_n\right)\right| = 1 \quad \text{almost surely.}$$

The further approach to estimate the multiple change-point $\boldsymbol{\theta} = (\theta_1, \theta_2)$ is analogous to the last part of Section 3.1. We define the estimator

$$\hat{\boldsymbol{\theta}}_n := \frac{1}{n}\hat{\boldsymbol{\tau}}_n \tag{4.7}$$

for the multiple change-point and set

$$\hat{\rho}_n(s,t) := \frac{1}{n}\hat{M}_n(\lfloor ns \rfloor, \lfloor nt \rfloor), \quad (s,t) \in \Theta_n,$$

where Θ_n is given by (3.6).

Lemma 4.5. *Let $n \in \mathbb{N}$. Then*

$$\hat{\boldsymbol{\theta}}_n = \underset{(s,t)\in\Theta_n}{\text{argmax}}\, \hat{\rho}_n(s,t).$$

Proof. By similar arguments used in the proof of Lemma 3.4, we obtain the claim. □

Remark 4.6. The factor n^{-1} in the definition of $\hat{\rho}_n$ does not influence the maximizing points of $\hat{\rho}_n$, but the proof of consistency of $\hat{\boldsymbol{\theta}}_n$ requires this factor.

Lemma 4.7. *$\hat{\rho}_n$, $n \in \mathbb{N}$, is a stochastic process with trajectories in the multivariate Skorokhod space $D(\Theta_n)$.*

Proof. The proof of Lemma 3.6 remains valid for $\hat{\rho}_n$ and $\hat{M}_n$ instead of $\bar{\rho}_n$ and $\bar{M}_n$. □

4.1.2 Consistency of the multiple change-point estimator

In this section we discuss the weak and strong consistency of $\hat{\boldsymbol{\theta}}_n$. To apply Theorem 2.1 again, some results are adapted from Albrecht [1].

Let us introduce the function $\hat{\rho} : \Theta \longrightarrow \mathbb{R}$ defined by

$$\hat{\rho}(s,t) = \begin{cases} t\alpha^2 + (1-t)\left(\frac{\theta_1 - t}{1-t}\alpha + \frac{\theta_2-\theta_1}{1-t}\beta + \frac{1-\theta_2}{1-t}\gamma\right)^2, & (s,t) \in \Theta^1, \\ s\alpha^2 + (t-s)\left(\frac{\theta_1-s}{t-s}\alpha + \frac{t-\theta_1}{t-s}\beta\right)^2 + (1-t)\left(\frac{\theta_2-t}{1-t}\beta + \frac{1-\theta_2}{1-t}\gamma\right)^2, & (s,t) \in \Theta^2, \\ s\alpha^2 + (t-s)\left(\frac{\theta_1-s}{t-s}\alpha + \frac{\theta_2-\theta_1}{t-s}\beta + \frac{t-\theta_2}{t-s}\gamma\right)^2 + (1-t)\gamma^2, & (s,t) \in \Theta^3, \\ s\left(\frac{\theta_1}{s}\alpha + \frac{s-\theta_1}{s}\beta\right)^2 + (t-s)\beta^2 + (1-t)\left(\frac{\theta_2-t}{1-t}\beta + \frac{1-\theta_2}{1-t}\gamma\right)^2, & (s,t) \in \Theta^4, \\ s\left(\frac{\theta_1}{s}\alpha + \frac{s-\theta_1}{s}\beta\right)^2 + (t-s)\left(\frac{\theta_2-s}{t-s}\beta + \frac{t-\theta_2}{t-s}\gamma\right)^2 + (1-t)\gamma^2, & (s,t) \in \Theta^5, \\ s\left(\frac{\theta_1}{s}\alpha + \frac{\theta_2-\theta_1}{s}\beta + \frac{s-\theta_2}{s}\gamma\right)^2 + (1-s)\gamma^2, & (s,t) \in \Theta^6, \end{cases} \tag{4.8}$$

where $\Theta^1, ..., \Theta^6$ are given by (3.12). The function $\hat{\rho}$ is illustrated in Figure 4.1.

We show uniform convergence in probability of $\hat{\rho}_n$ to $\hat{\rho}$ (assumption (i) of Theorem 2.1).

Proposition 4.8. *Suppose there is some $p \in (2, \infty)$ such that $M_p < \infty$. Then*

$$\sup_{(s,t)\in\Theta_n} |\hat{\rho}_n(s,t) - \hat{\rho}(s,t)| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Proof. Fix $n \in \mathbb{N}$ for a moment. By Lemma 3.2 in Albrecht [1, p. 24], we get the decomposition $\hat{\rho}_n(s,t) = \hat{\delta}_n(s,t) + \hat{\varrho}_n(s,t)$ for all $(s,t) \in \Theta_n$, where $\hat{\delta}_n$ is a certain stochastic

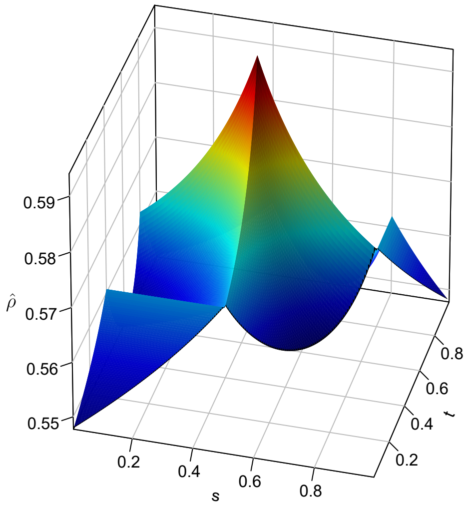

Figure 4.1: Plot of $\hat{\rho}$ for $\boldsymbol{\theta} = (\theta_1, \theta_2) = (0.4, 0.8)$ and $\boldsymbol{\alpha} = (\alpha, \beta, \gamma) = (0.6, 1, 0.5)$.

process and $\hat{\varrho}_n$ a deterministic function. The concrete forms are specified in Lemma B.1. We have

$$\begin{aligned}
\sup_{(s,t)\in\Theta_n} |\hat{\rho}_n(s,t) - \hat{\rho}(s,t)| &= \sup_{(s,t)\in\Theta_n} \left|\hat{\hat{\delta}}_n(s,t) + \hat{\varrho}_n(s,t) - \hat{\rho}(s,t)\right| \\
&\leq \sup_{(s,t)\in\Theta_n} \left(\left|\hat{\hat{\delta}}_n(s,t)\right| + |\hat{\varrho}_n(s,t) - \hat{\rho}(s,t)|\right) \qquad \text{by Tr. In.} \\
&\leq \sup_{(s,t)\in\Theta_n} \left|\hat{\hat{\delta}}_n(s,t)\right| + \sup_{(s,t)\in\Theta_n} |\hat{\varrho}_n(s,t) - \hat{\rho}(s,t)|\,. \qquad (4.9)
\end{aligned}$$

Furthermore, we find in the proof of Lemma 3.6 in Albrecht [1, p. 42-54] that for each $\varepsilon > 0$ there exists a constant $C_p > 0$, which depends only on p, such that

$$\mathbb{P}\left[\sup_{(s,t)\in\Theta_n} \left|\hat{\hat{\delta}}_n(s,t)\right| > \varepsilon\right] \leq C_p \varepsilon^{-p} n^{-(p/2-1)} \ln n.$$

Since $p > 2$, we have $\frac{p}{2} - 1 > 0$. L'Hôpital's rule yields for all $\varepsilon > 0$

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}\left[\sup_{(s,t)\in\Theta_n} \left|\hat{\hat{\delta}}_n(s,t)\right| > \varepsilon\right] &\leq C_p \varepsilon^{-p} \lim_{n\to\infty} n^{-(p/2-1)} \ln n \\
&= C_p \varepsilon^{-p} \left(\frac{p}{2} - 1\right)^{-1} \lim_{n\to\infty} n^{-(p/2-1)} \\
&= 0,
\end{aligned}$$

which leads to

$$\sup_{(s,t)\in\Theta_n}\left|\hat{\delta}_n(s,t)\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Moreover, by Lemma 3.5 in Albrecht [1, p. 38], it holds

$$\sup_{(s,t)\in\Theta_n}\left|\hat{\varrho}_n(s,t)-\hat{\rho}(s,t)\right| \xrightarrow[n\to\infty]{} 0. \tag{4.10}$$

The assertion follows by applying (4.9). □

We obtain assumption (ii) of Theorem 2.1 by the following proposition.

Proposition 4.9. *The multiple change-point $\boldsymbol{\theta}\in\Theta$ is the well-separated maximizer of $\hat{\rho}$.*

Proof. This was proved by Albrecht [1, p. 32, Lemma 3.4]. □

We can now prove weak consistency of $\hat{\boldsymbol{\theta}}_n$.

Theorem 4.10. *Suppose there is some $p\in(2,\infty)$ such that $M_p<\infty$. Then*

$$\hat{\boldsymbol{\theta}}_n \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\theta}.$$

Proof. We apply Theorem 2.1. $\hat{\rho}_n$, $n\in\mathbb{N}$, is a stochastic process with trajectories in the multivariate Skorokhod space by Lemma 4.7. $\hat{\rho}$ has trajectories in the multivariate Skorokhod space $D(\Theta)$, since $\hat{\rho}$ is continuous, which was shown in Albrecht [1, p. 29, Lemma 3.3]. Moreover, $(\Theta_n)_{n\in\mathbb{N}}\subseteq\Theta$ is a sequence of sets such that $\Theta_n\subseteq\Theta_{n+1}$ for every $n\in\mathbb{N}$ with $\bigcup_{n\in\mathbb{N}}\Theta_n=\Theta$. By Lemma 4.5, $\hat{\boldsymbol{\theta}}_n$ is a maximizer of $\hat{\rho}_n$ for any $n\in\mathbb{N}$. Assumption (i) of Theorem 2.1 and (2.2) are satisfied by Propositions 4.8 and 4.9. Applying Theorem 2.1 and Remark 2.2 gives the claim. □

Albrecht [1] even showed the strong consistency of $\hat{\boldsymbol{\theta}}_n$ if there exists a larger moment.

Theorem 4.11. *Suppose there is some* $p \in (4, \infty)$ *such that* $M_p < \infty$. *Then*

$$\hat{\boldsymbol{\theta}}_n \xrightarrow[n\to\infty]{a.s.} \boldsymbol{\theta}.$$

Proof. The proof can be found in Albrecht [1, p. 52, Theorem 3.7]. □

4.1.3 Stochastic boundedness

We now treat stochastic boundedness of $\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$, which is required to prove consistency of $\hat{\boldsymbol{\alpha}}_n$ in the next section.

We begin with the observation that the function $\hat{\rho}$ has a local peak at the multiple change-point $\boldsymbol{\theta} = (\theta_1, \theta_2)$.

Lemma 4.12. *There exist* $\delta > 0$ *and a constant* $L = L(\delta) > 0$ *such that*

$$\hat{\rho}(\theta_1, \theta_2) - \hat{\rho}(s, t) \geq L\|(s, t) - (\theta_1, \theta_2)\| \tag{4.11}$$

for all $(s, t) \in B_\delta(\theta_1, \theta_2)$.

To show this lemma, we characterize the property (4.11) by directional derivatives. For this purpose, let us recall the definition of directional derivatives. Let $U \subseteq \mathbb{R}^q$, $q \in \mathbb{N}$, be an open set, $f : U \longrightarrow \mathbb{R}$ a mapping and $\mathbf{v} \in \mathbb{R}^q$ a vector with $\|\mathbf{v}\| = 1$. The limit

$$\partial_{\mathbf{v}} f(\mathbf{t}) := \lim_{\lambda \downarrow 0} \frac{f(\mathbf{t} + \lambda \mathbf{v}) - f(\mathbf{t})}{\lambda} \tag{4.12}$$

is said to be *directional derivative* of f in $\mathbf{t} \in U$ if the limit exists.

Lemma 4.13. *The following conditions are equivalent:*

(i) $\hat{\rho}$ *sastisfies* (4.11).

(ii) It holds $\max\left\{\partial_{(1,0)}\hat{\rho}(\theta_1, \theta_2), \partial_{(-1,0)}\hat{\rho}(\theta_1, \theta_2), \partial_{(0,1)}\hat{\rho}(\theta_1, \theta_2), \partial_{(0,-1)}\hat{\rho}(\theta_1, \theta_2)\right\} < 0.$

Proof. Döring [9, p. 52, Lemma 3.11] showed the assertion for another function. The same proof remains valid for $\hat{\rho}$. □

Proof of Lemma 4.12. The proof is straightforward. It is sufficient to compute the directional derivatives of $\hat{\rho}$, which are given in Lemma 4.13. To this end, we have to consider $\hat{\rho}$ on the domains Θ^2, Θ^3 and Θ^4 (compare Equation (3.12) and Figure 3.1). As defined in (4.8), the representation

$$\hat{\rho}(s,t) = \begin{cases} s\alpha^2 + (t-s)\left(\frac{\theta_1 - s}{t-s}(\alpha-\beta)+\beta\right)^2 + (1-t)\left(\frac{\theta_2-t}{1-t}(\beta-\gamma)+\gamma\right)^2, & (s,t)\in\Theta^2,\\ s\alpha^2 + (t-s)\left(\frac{\theta_1 - s}{t-s}(\alpha-\beta)+\beta+\frac{t-\theta_2}{t-s}(\gamma-\beta)\right)^2 + (1-t)\gamma^2, & (s,t)\in\Theta^3,\\ s\left(\frac{\theta_1-s}{s}(\alpha-\beta)+\alpha\right)^2 + (t-s)\beta^2 + (1-t)\left(\frac{\theta_2-t}{1-t}(\beta-\gamma)+\gamma\right)^2, & (s,t)\in\Theta^4 \end{cases}$$

simplifies the computation of the directional derivatives. We observe that

$$\hat{\rho}(\theta_1,\theta_2) = \theta_1\alpha^2 + (\theta_2-\theta_1)\beta^2 + (1-\theta_2)\gamma^2.$$

We first look at

$$\partial_{(-1,0)}\hat{\rho}(\theta_1,\theta_2) = \lim_{\lambda\downarrow 0}\frac{\hat{\rho}(\theta_1-\lambda,\theta_2)-\hat{\rho}(\theta_1,\theta_2)}{\lambda}.$$

For $\lambda > 0$ we have $(\theta_1-\lambda,\theta_2)\in\Theta^2$. A trivial verification shows that

$$\hat{\rho}(\theta_1-\lambda,\theta_2)-\hat{\rho}(\theta_1,\theta_2) = \frac{\lambda^2(\alpha-\beta)^2}{\theta_2-\theta_1+\lambda} - \lambda(\alpha-\beta)^2.$$

Hence

$$\partial_{(-1,0)}\hat{\rho}(\theta_1,\theta_2) = \lim_{\lambda\downarrow 0}\left(\frac{\lambda(\alpha-\beta)^2}{\theta_2-\theta_1+\lambda} - (\alpha-\beta)^2\right) = -(\alpha-\beta)^2.$$

In the same manner we can see that

$$\begin{aligned} \partial_{(1,0)}\hat{\rho}(\theta_1,\theta_2) &= \lim_{\lambda\downarrow 0}\left(\frac{\lambda(\alpha-\beta)^2}{\theta_1+\lambda} - (\alpha-\beta)^2\right) = -(\alpha-\beta)^2,\\ \partial_{(0,1)}\hat{\rho}(\theta_1,\theta_2) &= \lim_{\lambda\downarrow 0}\left(\frac{\lambda(\beta-\gamma)^2}{\theta_2-\theta_1+\lambda} - (\beta-\gamma)^2\right) = -(\beta-\gamma)^2,\\ \partial_{(0,-1)}\hat{\rho}(\theta_1,\theta_2) &= \lim_{\lambda\downarrow 0}\left(\frac{\lambda(\beta-\gamma)^2}{1-\theta_2+\lambda} - (\beta-\gamma)^2\right) = -(\beta-\gamma)^2. \end{aligned}$$

By model assumptions $\alpha \neq \beta$ and $\beta \neq \gamma$, we conclude that

$$\max\left\{-(\alpha-\beta)^2, -(\beta-\gamma)^2\right\} < 0,$$

which establishes condition (ii) in Lemma 4.13. An application of Lemma 4.13 completes our proof. □

The following both lemmas are useful to get an error estimate.

Lemma 4.14. *Let $\hat{\varrho}_n$ be the deterministic function from Lemma B.1. Then there exist an arbitrary small $\delta > 0$, a constant $L = L(\delta) > 0$ and $n_0 \in \mathbb{N}$ such that*

$$\hat{\varrho}_n\left(\frac{\tau_n}{n}, \frac{\sigma_n}{n}\right) - \hat{\varrho}_n(s,t) \geq \frac{1}{2}L\frac{\|(\lfloor ns\rfloor, \lfloor nt\rfloor) - (\tau_n, \sigma_n)\|}{n} \tag{4.13}$$

for all $n \geq n_0$ and $(s,t) \in B_\delta(\theta_1, \theta_2)$.

Proof. We first observe that $\left(\frac{\tau_n}{n}, \frac{\sigma_n}{n}\right) \in \Theta_n$ by model assumption (1.1). Furthermore, by Lemma 4.12, there exist an arbitrary small $\delta > 0$ and a constant $L = L(\delta) > 0$ such that $\hat{\rho}(\theta_1, \theta_2) - \hat{\rho}(s,t) \geq L\|(s,t) - (\theta_1, \theta_2)\|$ for all $(s,t) \in B_\delta(\theta_1, \theta_2)$. Moreover, for all $n \in \mathbb{N}$ and $(s,t) \in B_\delta(\theta_1, \theta_2)$ we have

$$\begin{aligned}
&\hat{\varrho}_n\left(\frac{\tau_n}{n}, \frac{\sigma_n}{n}\right) - \hat{\varrho}_n(s,t) \\
&\quad = \left(\hat{\varrho}_n\left(\frac{\tau_n}{n}, \frac{\sigma_n}{n}\right) - \hat{\varrho}_n(s,t)\right)\frac{n}{\|(\lfloor ns\rfloor, \lfloor nt\rfloor) - (\tau_n, \sigma_n)\|}\frac{\|(\lfloor ns\rfloor, \lfloor nt\rfloor) - (\tau_n, \sigma_n)\|}{n} \\
&\quad =: b_n(s,t)\frac{\|(\lfloor ns\rfloor, \lfloor nt\rfloor) - (\tau_n, \sigma_n)\|}{n}.
\end{aligned} \tag{4.14}$$

The uniform convergence of $\hat{\varrho}_n$, given in (4.10), and the properties of the floor function (uniform convergence, see Lemma A.1 (iii)) lead to

$$\lim_{n\to\infty} b_n(s,t) = \lim_{n\to\infty} \frac{\hat{\varrho}_n\left(\frac{\tau_n}{n}, \frac{\sigma_n}{n}\right) - \hat{\varrho}_n(s,t)}{\left\|\left(\frac{\lfloor ns\rfloor}{n}, \frac{\lfloor nt\rfloor}{n}\right) - \left(\frac{\sigma_n}{n}, \frac{\tau_n}{n}\right)\right\|} = \frac{\hat{\rho}(\theta_1, \theta_2) - \hat{\rho}(s,t)}{\|(s,t) - (\theta_1, \theta_2)\|} \geq L$$

uniformly for all $(s,t) \in B_\delta(\theta_1, \theta_2)$. Combining this with (4.14) ensures the existence of $n_0 \in \mathbb{N}$ such that (4.13) holds for all $n \geq n_0$ and $(s,t) \in B_\delta(\theta_1, \theta_2)$. □

Recall that $\|\cdot\|$ stands for the maximum norm. Let $G_{n,x,\delta}$ denote the set

$$G_{n,x,\delta} := \{(k,l) \in \Delta_n | x \leq \|(k,l) - (\tau_n, \sigma_n)\| \leq n\delta\}$$

for $n \in \mathbb{N}$, $x > 0$ and $\delta > 0$.

Lemma 4.15. *Let $x > 0$, $\delta > 0$ and $n \in \mathbb{N}$. Then*

$$\{x \leq \|\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta\} \subseteq \bigcup_{(k,l)\in G_{n,x,\delta}} \left\{\hat{\rho}_n\left(\frac{k}{n}, \frac{l}{n}\right) - \hat{\rho}_n\left(\frac{\tau_n}{n}, \frac{\sigma_n}{n}\right) \geq 0\right\}.$$

Proof. It is easily seen that $\left(\frac{k}{n}, \frac{l}{n}\right) \in \Theta_n$ for $(k,l) \in G_{n,x,\delta}$ and $\left(\frac{\tau_n}{n}, \frac{\sigma_n}{n}\right) \in \Theta_n$ by model assumption (1.1). Conversely, suppose that there exists $\omega \in \{x \leq \|\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta\}$, but

$$\omega \notin \bigcup_{(k,l)\in G_{n,x,\delta}} \left\{\hat{\rho}_n\left(\frac{k}{n}, \frac{l}{n}\right) - \hat{\rho}_n\left(\frac{\tau_n}{n}, \frac{\sigma_n}{n}\right) \geq 0\right\}.$$

It follows that $\hat{\rho}_n\left(\frac{k}{n}, \frac{l}{n}\right) - \hat{\rho}_n\left(\frac{\tau_n}{n}, \frac{\sigma_n}{n}\right) < 0$ for all $(k,l) \in G_{n,x,\delta}$. By definition of $\hat{\boldsymbol{\tau}}_n$, we see that $\hat{\boldsymbol{\tau}}_n = (\hat{\tau}_n, \hat{\sigma}_n) \in \Delta_n$, and so $\hat{\boldsymbol{\tau}}_n \in G_{n,x,\delta}$ by assumption. The definition of $\hat{\boldsymbol{\theta}}_n$ gives

$$0 > \hat{\rho}_n\left(\frac{\hat{\tau}_n}{n}, \frac{\hat{\sigma}_n}{n}\right) - \hat{\rho}_n\left(\frac{\tau_n}{n}, \frac{\sigma_n}{n}\right) = \hat{\rho}_n\left(\frac{1}{n}\hat{\boldsymbol{\tau}}_n\right) - \hat{\rho}_n\left(\frac{1}{n}\boldsymbol{\tau}_n\right) = \hat{\rho}_n\left(\hat{\boldsymbol{\theta}}_n\right) - \hat{\rho}_n\left(\frac{1}{n}\boldsymbol{\tau}_n\right),$$

which contradicts the fact that $\hat{\boldsymbol{\theta}}_n$ maximizes $\hat{\rho}_n$ by Lemma 4.5. □

The following error estimate provides the basis for the proof of stochastic boundedness.

Lemma 4.16. *Suppose there is some $p \in (2, \infty)$ such that $M_p < \infty$. Then there exist $n_0 \in \mathbb{N}$, $\delta > 0$, $\kappa \in \left(0, \frac{1}{2}\right)$ and a constant $C > 0$ such that for all $n \geq n_0$ we have*

$$\mathbb{P}[x \leq \|\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta] \leq C\left(n^{-(p/2-1)} + x^{-(1/2-\kappa)p}\right)$$

for all $x \geq 4$.

Proof. Fix $x \geq 4$. By Lemma 3.19, there exist $\delta > 0$ and $\tilde{n}_0 \in \mathbb{N}$ such that for all $n \geq \tilde{n}_0$ the conditions hold in Lemma 3.19. Let us regard $n \geq \tilde{n}_0$ and $\delta > 0$ as fixed. By Lemma 4.15 and the decomposition of $\hat{\rho}_n$ (Lemma B.1), we see that

$$\begin{aligned}
&\{x \leq \|\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta\} \\
&\subseteq \bigcup_{(k,l)\in G_{n,x,\delta}} \left\{\hat{\rho}_n\left(\frac{k}{n},\frac{l}{n}\right) - \hat{\rho}_n\left(\frac{\tau_n}{n},\frac{\sigma_n}{n}\right) \geq 0\right\} \\
&= \bigcup_{(k,l)\in G_{n,x,\delta}} \left\{\hat{\delta}_n\left(\frac{k}{n},\frac{l}{n}\right) - \hat{\delta}_n\left(\frac{\tau_n}{n},\frac{\sigma_n}{n}\right) + \hat{\varrho}_n\left(\frac{k}{n},\frac{l}{n}\right) - \hat{\varrho}_n\left(\frac{\tau_n}{n},\frac{\sigma_n}{n}\right) \geq 0\right\}.
\end{aligned}$$

By Lemma 4.14, there exist a constant $L = L(\delta) > 0$ and $\hat{n}_0 \in \mathbb{N}$ such that for all $n \geq \hat{n}_0$

$$\begin{aligned}
&\{x \leq \|\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta\} \\
&\subseteq \bigcup_{(k,l)\in G_{n,x,\delta}} \left\{\hat{\delta}_n\left(\frac{k}{n},\frac{l}{n}\right) - \hat{\delta}_n\left(\frac{\tau_n}{n},\frac{\sigma_n}{n}\right) - \frac{1}{2}L\frac{\|(k,l)-(\tau_n,\sigma_n)\|}{n} \geq 0\right\} \\
&= \bigcup_{(k,l)\in G_{n,x,\delta}} \left\{\frac{n}{\|(k,l)-(\tau_n,\sigma_n)\|}\left(\hat{\delta}_n\left(\frac{k}{n},\frac{l}{n}\right) - \hat{\delta}_n\left(\frac{\tau_n}{n},\frac{\sigma_n}{n}\right)\right) \geq \frac{1}{2}L\right\} \\
&\subseteq \bigcup_{\substack{x\leq|k-\tau_n|\leq n\delta \\ |k-\tau_n|\geq|l-\sigma_n|}} \left\{\frac{n}{|k-\tau_n|}\left(\hat{\delta}_n\left(\frac{k}{n},\frac{l}{n}\right) - \hat{\delta}_n\left(\frac{\tau_n}{n},\frac{\sigma_n}{n}\right)\right) \geq \frac{1}{2}L\right\} \\
&\qquad \cup \bigcup_{\substack{x\leq|l-\sigma_n|\leq n\delta \\ |l-\sigma_n|\geq|k-\tau_n|}} \left\{\frac{n}{|l-\sigma_n|}\left(\hat{\delta}_n\left(\frac{k}{n},\frac{l}{n}\right) - \hat{\delta}_n\left(\frac{\tau_n}{n},\frac{\sigma_n}{n}\right)\right) \geq \frac{1}{2}L\right\} \\
&=: E \cup F. \qquad (4.15)
\end{aligned}$$

To simplify notation, the fact that some in this proof defined sets and random variables depend on n, x or δ is omitted. From now on, let $n \geq \max\{\tilde{n}_0, \hat{n}_0\}$. We give the proof only for the estimate of the probability of E; the other case follows the same pattern.

Computing the absolute values give

$$\begin{aligned}E \subseteq & \bigcup_{\substack{x\leq k-\tau_n\leq n\delta\\ 0\leq l-\sigma_n\leq k-\tau_n}} \left\{\frac{n}{k-\tau_n}\left(\hat{\delta}_n\left(\frac{k}{n},\frac{l}{n}\right)-\hat{\delta}_n\left(\frac{\tau_n}{n},\frac{\sigma_n}{n}\right)\right)\geq \frac{1}{2}L\right\}\\ & \cup \bigcup_{\substack{x\leq k-\tau_n\leq n\delta\\ 0\leq \sigma_n-l\leq k-\tau_n}} \left\{\frac{n}{k-\tau_n}\left(\hat{\delta}_n\left(\frac{k}{n},\frac{l}{n}\right)-\hat{\delta}_n\left(\frac{\tau_n}{n},\frac{\sigma_n}{n}\right)\right)\geq \frac{1}{2}L\right\}\\ & \cup \bigcup_{\substack{x\leq \tau_n-k\leq n\delta\\ 0\leq l-\sigma_n\leq \tau_n-k}} \left\{\frac{n}{\tau_n-k}\left(\hat{\delta}_n\left(\frac{k}{n},\frac{l}{n}\right)-\hat{\delta}_n\left(\frac{\tau_n}{n},\frac{\sigma_n}{n}\right)\right)\geq \frac{1}{2}L\right\}\\ & \cup \bigcup_{\substack{x\leq \tau_n-k\leq n\delta\\ 0\leq \sigma_n-l\leq \tau_n-k}} \left\{\frac{n}{\tau_n-k}\left(\hat{\delta}_n\left(\frac{k}{n},\frac{l}{n}\right)-\hat{\delta}_n\left(\frac{\tau_n}{n},\frac{\sigma_n}{n}\right)\right)\geq \frac{1}{2}L\right\}\\ =: & \bigcup_{i=1}^{4} E_i. \end{aligned} \tag{4.16}$$

The technique of the proof is presented for E_1 and E_2. Throughout the proof, we use the abbreviation $S_{u,v} := \sum_{i=u+1}^{v}(X_i - \mathbb{E}[X_i])$ for $u, v \in \mathbb{N}_0$ with $u < v$. Let us consider E_1. We first observe that

$$\tau_n + x \leq k \leq \tau_n + n\delta < \sigma_n \leq l \leq \sigma_n + n\delta < n \tag{4.17}$$

by Lemma 3.19. Hence $\left(\frac{k}{n}, \frac{l}{n}\right) \in \Theta^5 \cap \Theta_n$. By Lemma B.1 and the Binomial Formula, an easy computation yields

$$
\begin{aligned}
&\hat{\delta}_n\left(\frac{k}{n},\frac{l}{n}\right)-\hat{\delta}_n\left(\frac{\tau_n}{n},\frac{\sigma_n}{n}\right)\\
&=\frac{1}{nk}(S_{0,\tau_n}+S_{\tau_n,k})^2+\frac{2(\tau_n\alpha+(k-\tau_n)\beta)}{nk}(S_{0,\tau_n}+S_{\tau_n,k})+\frac{1}{n(l-k)}(S_{k,\sigma_n}+S_{\sigma_n,l})^2\\
&\quad+\frac{2((\sigma_n-k)\beta+(l-\sigma_n)\gamma)}{n(l-k)}(S_{k,\sigma_n}+S_{\sigma_n,l})+\frac{1}{n(n-l)}S_{l,n}^2+\frac{2\gamma}{n}S_{l,n}\\
&\quad-\left[\frac{1}{n\tau_n}S_{0,\tau_n}^2+\frac{2\alpha}{n}S_{0,\tau_n}+\frac{1}{n(\sigma_n-\tau_n)}(S_{\tau_n,k}+S_{k,\sigma_n})^2+\frac{2\beta}{n}(S_{\tau_n,k}+S_{k,\sigma_n})\right.\\
&\qquad\left.+\frac{1}{n(n-\sigma_n)}(S_{\sigma_n,l}+S_{l,n})^2+\frac{2\gamma}{n}(S_{\sigma_n,l}+S_{l,n})\right]\\
&=\frac{1}{n}\left(\frac{1}{k}-\frac{1}{\tau_n}\right)S_{0,\tau_n}^2+\frac{2(k-\tau_n)(\beta-\alpha)}{nk}S_{0,\tau_n}+\frac{2}{nk}S_{0,\tau_n}S_{\tau_n,k}\\
&\quad+\frac{1}{n}\left(\frac{1}{k}-\frac{1}{\sigma_n-\tau_n}\right)S_{\tau_n,k}^2+\frac{2\tau_n(\alpha-\beta)}{nk}S_{\tau_n,k}-\frac{2}{n(\sigma_n-\tau_n)}S_{\tau_n,k}S_{k,\sigma_n}\\
&\quad+\frac{1}{n}\left(\frac{1}{l-k}-\frac{1}{\sigma_n-\tau_n}\right)S_{k,\sigma_n}^2+\frac{2(l-\sigma_n)(\gamma-\beta)}{n(l-k)}S_{k,\sigma_n}+\frac{2}{n(l-k)}S_{k,\sigma_n}S_{\sigma_n,l}\\
&\quad+\frac{1}{n}\left(\frac{1}{l-k}-\frac{1}{n-\sigma_n}\right)S_{\sigma_n,l}^2+\frac{2(\sigma_n-k)(\beta-\gamma)}{n(l-k)}S_{\sigma_n,l}-\frac{2}{n(n-\sigma_n)}S_{\sigma_n,l}S_{l,n}\\
&\quad+\frac{1}{n}\left(\frac{1}{n-l}-\frac{1}{n-\sigma_n}\right)S_{l,n}^2\\
&=:\sum_{i=1}^{13}A_{1,i}(k,l).
\end{aligned}
$$

Observe that for $l = \sigma_n$ we have $A_{1,i}(k,l) = 0$, $i \in \{9, ..., 13\}$. We can conclude that

$$\begin{aligned}
E_1 &= \bigcup_{\substack{x \le k-\tau_n \le n\delta \\ 1 \le l-\sigma_n \le k-\tau_n}} \left\{ \sum_{i=1}^{13} \frac{n}{k-\tau_n} A_{1,i}(k,l) \ge \frac{1}{2} L \right\} && \text{by def. of } E_1 \\
&\subseteq \bigcup_{i=1}^{13} \left(\bigcup_{\substack{x \le k-\tau_n \le n\delta \\ 1 \le l-\sigma_n \le k-\tau_n}} \left\{ \frac{n}{k-\tau_n} A_{1,i}(k,l) \ge \frac{1}{26} L \right\} \right) && \text{by Lem. A.4} \\
&= \bigcup_{i=1}^{13} \left\{ \max_{\substack{x \le k-\tau_n \le n\delta \\ 1 \le l-\sigma_n \le k-\tau_n}} \frac{n}{k-\tau_n} A_{1,i}(k,l) \ge \frac{1}{26} L \right\} \\
&\subseteq \bigcup_{i=1}^{13} \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le l-\sigma_n \le k-\tau_n}} \frac{n}{k-\tau_n} A_{1,i}(k,l) \ge \frac{1}{26} L \right\} \\
&=: \bigcup_{i=1}^{13} E_{1,i}.
\end{aligned} \tag{4.18}$$

We now estimate the probabilities of the events $E_{1,i}$, $i \in \{1, ..., 13\}$, successively. To do this, note that there exist positive constants $c_1 = c_1(\theta_1)$, $c_2 = c_2(\theta_1, \theta_2)$, $c_3 = c_3(\theta_2)$, $c_4 = c_4(\delta)$, $c_5 = c_5(\theta_1, \theta_2, \delta)$ and natural numbers $n_1 = n_1(\theta_1)$, $n_2 = n_2(\theta_1, \theta_2)$, $n_3 = n_3(\theta_2)$, $n_4 = n_4(\delta)$, $n_5 = n_5(\theta_1, \theta_2, \delta)$ such that

$$\begin{aligned}
\tau_n &\ge c_1 n \quad \text{for all} \quad n \ge n_1, \\
\sigma_n - \tau_n &\ge c_2 n \quad \text{for all} \quad n \ge n_2, \\
n - \sigma_n &\ge c_3 n \quad \text{for all} \quad n \ge n_3, \\
\lfloor n\delta \rfloor &\ge c_4 n \quad \text{for all} \quad n \ge n_4, \\
\sigma_n - \tau_n - \lfloor n\delta \rfloor &\ge c_5 n \quad \text{for all} \quad n \ge n_5.
\end{aligned} \tag{4.19}$$

To see this, consider for example the last assertion. By properties of the floor function (see Lemma A.1 (i)), we obtain

$$\sigma_n - \tau_n - \lfloor n\delta \rfloor = \lfloor n\theta_2 \rfloor - \lfloor n\theta_1 \rfloor - \lfloor n\delta \rfloor \ge n(\theta_2 - \theta_1 - \delta) - 1$$

for each $n \in \mathbb{N}$. It is easily seen that there exists $n_5 = n_5(\theta_1, \theta_2, \delta) \in \mathbb{N}$ such that $1 \le \frac{1}{2} n(\theta_2 - \theta_1 - \delta)$ for all $n \ge n_5$. Accordingly, we get $\sigma_n - \tau_n - \lfloor n\delta \rfloor \ge c_5 n$ with

$c_5 := c_5(\theta_1, \theta_2, \delta) := \frac{1}{2}(\theta_2 - \theta_1 - \delta)$ for each $n \geq n_5$.

From now on, fix $n \geq n_0 := \max\{\tilde{n}_0, \hat{n}_0, n_1, n_2, n_3, n_4, n_5\}$ and let $C = C(p, \kappa, \delta, \theta_1, \theta_2, \alpha, \beta, \gamma) > 0$ be a generic constant and $p > 2$. We begin by estimating the probability of $E_{1,1}$. It holds

$$E_{1,1} = \left\{ \max_{\substack{\lfloor x \rfloor \leq k - \tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq l - \sigma_n \leq k - \tau_n}} \frac{n}{k - \tau_n} A_{1,1}(k,l) \geq \frac{1}{26} L \right\}$$
$$\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \leq k - \tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq l - \sigma_n \leq k - \tau_n}} \left| \frac{n}{k - \tau_n} A_{1,1}(k,l) \right|^{p/2} \geq \left(\frac{1}{26} L \right)^{p/2} \right\}.$$

We have

$$\begin{aligned}
\left| \frac{n}{k - \tau_n} A_{1,1}(k,l) \right|^{p/2} &= \left| \frac{n}{(k - \tau_n)n} \left(\frac{1}{k} - \frac{1}{\tau_n} \right) S_{0,\tau_n}^2 \right|^{p/2} && \text{by def. of } A_{1,1} \\
&= (k\tau_n)^{-p/2} \left| S_{0,\tau_n} \right|^p \\
&< \tau_n^{-p} \left| S_{0,\tau_n} \right|^p && \text{by } k > \tau_n \\
&\leq C n^{-p} \left| S_{0,\tau_n} \right|^p && \text{by (4.19).}
\end{aligned}$$

By Markov's Inequality (see Lemma 2.4), Corollary 2.16, $M_p < \infty$ and $\tau_n \leq n$, it follows that

$$\mathbb{P}\left[E_{1,1}\right] \leq \mathbb{P}\left[\left| S_{0,\tau_n} \right|^p \geq C n^p \right] \leq C n^{-p} \mathbb{E}\left[\left| S_{0,\tau_n} \right|^p \right] \leq C M_p n^{-p} \tau_n^{p/2} \leq C n^{-p/2}.$$

We next consider $E_{1,2}$. It holds

$$E_{1,2} = \left\{ \max_{\substack{\lfloor x \rfloor \leq k - \tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq l - \sigma_n \leq k - \tau_n}} \frac{n}{k - \tau_n} A_{1,2}(k,l) \geq \frac{1}{26} L \right\}$$
$$\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \leq k - \tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq l - \sigma_n \leq k - \tau_n}} \left| \frac{n}{k - \tau_n} A_{1,2}(k,l) \right|^{p} \geq \left(\frac{1}{26} L \right)^{p} \right\}.$$

We find that

$$\begin{aligned}
\left|\frac{n}{k-\tau_n}A_{1,2}(k,l)\right|^p &= \left|\frac{2n(k-\tau_n)(\beta-\alpha)}{(k-\tau_n)nk}S_{0,\tau_n}\right|^p && \text{by def. of } A_{1,2}\\
&= 2^p|\alpha-\beta|^p k^{-p}\left|S_{0,\tau_n}\right|^p\\
&< 2^p|\alpha-\beta|^p \tau_n^{-p}\left|S_{0,\tau_n}\right|^p && \text{by } k>\tau_n\\
&\le Cn^{-p}\left|S_{0,\tau_n}\right|^p && \text{by (4.19).}
\end{aligned}$$

Similar arguments used in the estimate of the probability of $E_{1,1}$ give

$$\mathbb{P}\left[E_{1,2}\right] \le \mathbb{P}\left[\left|S_{0,\tau_n}\right|^p \ge Cn^p\right] \le Cn^{-p/2}.$$

We next consider $E_{1,3}$. It holds

$$\begin{aligned}
E_{1,3} &= \left\{\max_{\substack{\lfloor x\rfloor \le k-\tau_n\le\lfloor n\delta\rfloor\\ 1\le l-\sigma_n\le k-\tau_n}} \frac{n}{k-\tau_n}A_{1,3}(k,l) \ge \frac{1}{26}L\right\}\\
&\subseteq \left\{\max_{\substack{\lfloor x\rfloor \le k-\tau_n\le\lfloor n\delta\rfloor\\ 1\le l-\sigma_n\le k-\tau_n}} \left|\frac{n}{k-\tau_n}A_{1,3}(k,l)\right|^p \ge \left(\frac{1}{26}L\right)^p\right\}.
\end{aligned}$$

We see that

$$\begin{aligned}
\left|\frac{n}{k-\tau_n}A_{1,3}(k,l)\right|^p &= \left|\frac{2n}{(k-\tau_n)nk}S_{0,\tau_n}S_{\tau_n,k}\right|^p && \text{by def. of } A_{1,3}\\
&= 2^p k^{-p}(k-\tau_n)^{-p}\left|S_{0,\tau_n}S_{\tau_n,k}\right|^p\\
&< 2^p \tau_n^{-p}(k-\tau_n)^{-p}\left|S_{0,\tau_n}S_{\tau_n,k}\right|^p && \text{by } k>\tau_n\\
&\le Cn^{-p}(k-\tau_n)^{-p}\left|S_{0,\tau_n}S_{\tau_n,k}\right|^p && \text{by (4.19).}
\end{aligned}$$

Hence

$$\begin{aligned}
\mathbb{P}\left[E_{1,3}\right] &\le \mathbb{P}\left[\max_{\lfloor x\rfloor\le k-\tau_n\le\lfloor n\delta\rfloor}(k-\tau_n)^{-p}\left|S_{0,\tau_n}S_{\tau_n,k}\right|^p \ge Cn^p\right]\\
&= \mathbb{P}\left[\bigcup_{k=\tau_n+\lfloor x\rfloor}^{\tau_n+\lfloor n\delta\rfloor}\left\{(k-\tau_n)^{-p}\left|S_{0,\tau_n}S_{\tau_n,k}\right|^p \ge Cn^p\right\}\right]\\
&\le \sum_{k=\tau_n+\lfloor x\rfloor}^{\tau_n+\lfloor n\delta\rfloor}\mathbb{P}\left[\left|S_{0,\tau_n}S_{\tau_n,k}\right|^p \ge Cn^p(k-\tau_n)^p\right]. \qquad (4.20)
\end{aligned}$$

The independence of $X_1, ..., X_n$ establishes the independence of S_{0,τ_n} and $S_{\tau_n,k}$ for each $k \in \{\tau_n + \lfloor x \rfloor, ..., \tau_n + \lfloor n\delta \rfloor\}$. By Markov's Inequality (see Lemma 2.4), we conclude that

$$\begin{aligned}\mathbb{P}\left[E_{1,3}\right] &\leq Cn^{-p} \sum_{k=\tau_n+\lfloor x \rfloor}^{\tau_n+\lfloor n\delta \rfloor} (k-\tau_n)^{-p} \mathbb{E}\left[|S_{0,\tau_n} S_{\tau_n,k}|^p\right] \\ &= Cn^{-p} \sum_{k=\tau_n+\lfloor x \rfloor}^{\tau_n+\lfloor n\delta \rfloor} (k-\tau_n)^{-p} \mathbb{E}\left[|S_{0,\tau_n}|^p\right] \mathbb{E}\left[|S_{\tau_n,k}|^p\right]. \qquad (4.21)\end{aligned}$$

We can deduce that

$$\begin{aligned}\mathbb{P}\left[E_{1,3}\right] &\leq CM_p^2 n^{-p} \tau_n^{p/2} \sum_{k=\tau_n+\lfloor x \rfloor}^{\tau_n+\lfloor n\delta \rfloor} (k-\tau_n)^{-p/2} && \text{by Cor. 2.16} \\ &\leq Cn^{-p/2} \sum_{m=\lfloor x \rfloor}^{\lfloor n\delta \rfloor} m^{-p/2} && \text{by } \tau_n \leq n,\ M_p < \infty \\ &\leq Cn^{-p/2} \sum_{m=1}^{\infty} m^{-p/2}.\end{aligned}$$

Since the series converges for $p > 2$, we obtain

$$\mathbb{P}\left[E_{1,3}\right] \leq Cn^{-p/2}.$$

We next consider $E_{1,4}$. Since $(k-\tau_n)^{-1} > 0$ and $S^2_{\tau_n,k} \geq 0$, we have

$$\begin{aligned}\frac{n}{k-\tau_n} A_{1,4}(k,l) &= \frac{n}{(k-\tau_n)n}\left(\frac{1}{k} - \frac{1}{\sigma_n - \tau_n}\right) S^2_{\tau_n,k} && \text{by def. of } A_{1,4} \\ &\leq k^{-1}(k-\tau_n)^{-1} S^2_{\tau_n,k} && \text{by } k^{-1} - (\sigma_n - \tau_n)^{-1} \leq k^{-1} \\ &< \tau_n^{-1}(k-\tau_n)^{-1} S^2_{\tau_n,k} && \text{by } k > \tau_n \\ &\leq Cn^{-1}(k-\tau_n)^{-1} S^2_{\tau_n,k} && \text{by (4.19).}\end{aligned}$$

By definition of $E_{1,4}$, it holds

$$\begin{aligned}E_{1,4} &= \left\{ \max_{\substack{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq l-\sigma_n \leq k-\tau_n}} \frac{n}{k-\tau_n} A_{1,4}(k,l) \geq \frac{1}{26} L \right\} \\ &\subseteq \left\{ \max_{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor} (k-\tau_n)^{-1} S^2_{\tau_n,k} \geq Cn \right\} \\ &\subseteq \left\{ \max_{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor} (k-\tau_n)^{-p/2} |S_{\tau_n,k}|^p \geq Cn^{p/2} \right\}.\end{aligned}$$

Let $\tilde{S}_{\tau_n,m}$ denote the process $\tilde{S}_{\tau_n,m} := \sum_{i=1}^{m}(X_{\tau_n+i} - \mathbb{E}[X_{\tau_n+i}])$ for $\lfloor x \rfloor \leq m \leq \lfloor n\delta \rfloor$. Notice that $\left(\left|\tilde{S}_{\tau_n,m}\right|^p\right)_{\lfloor x \rfloor \leq m \leq \lfloor n\delta \rfloor}$ is a non-negative submartingale by Lemma 2.14 and $\left(m^{-p/2}\right)_{\lfloor x \rfloor \leq m \leq \lfloor n\delta \rfloor}$ is a non-increasing sequence. By an index transformation and Chow's Inequality (Lemma 2.9), we get

$$\begin{aligned}
&\mathbb{P}[E_{1,4}] \\
&\leq \mathbb{P}\left[\max_{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor} (k-\tau_n)^{-p/2} |S_{\tau_n,k}|^p \geq C n^{p/2}\right] \\
&= \mathbb{P}\left[\max_{\lfloor x \rfloor \leq m \leq \lfloor n\delta \rfloor} m^{-p/2} \left|\tilde{S}_{\tau_n,m}\right|^p \geq C n^{p/2}\right] \\
&\leq C n^{-p/2}\left[\lfloor n\delta \rfloor^{-p/2} \mathbb{E}\left[\left|\tilde{S}_{\tau_n,\lfloor n\delta \rfloor}\right|^p\right] + \sum_{m=\lfloor x \rfloor}^{\lfloor n\delta \rfloor - 1}\left(m^{-p/2} - (m+1)^{-p/2}\right)\mathbb{E}\left[\left|\tilde{S}_{\tau_n,m}\right|^p\right]\right] \\
&\leq C n^{-p/2}\left[\lfloor n\delta \rfloor^{-p/2} \mathbb{E}\left[\left|\tilde{S}_{\tau_n,\lfloor n\delta \rfloor}\right|^p\right] + \frac{p}{2}\sum_{m=\lfloor x \rfloor}^{\lfloor n\delta \rfloor - 1} m^{-(p/2+1)}\mathbb{E}\left[\left|\tilde{S}_{\tau_n,m}\right|^p\right]\right]. \qquad (4.22)
\end{aligned}$$

The last inequality follows from Lemma A.6. We see that

$$\begin{aligned}
\mathbb{P}[E_{1,4}] &\leq C M_p n^{-p/2}\left[1 + \frac{p}{2}\sum_{m=\lfloor x \rfloor}^{\lfloor n\delta \rfloor - 1} m^{-1}\right] && \text{by Cor. 2.16} \\
&\leq C n^{-p/2}\left[1 + \frac{p}{2}(1 + \ln(\lfloor n\delta \rfloor - 1))\right] && \text{by Lem. A.5, } M_p < \infty \\
&\leq C n^{-p/2}\ln(n) && \text{by } \lfloor n\delta \rfloor - 1 \leq n.
\end{aligned}$$

We next consider $E_{1,5}$. It holds

$$\begin{aligned}
E_{1,5} &= \left\{\max_{\substack{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq l-\sigma_n \leq k-\tau_n}} \frac{n}{k-\tau_n} A_{1,5}(k,l) \geq \frac{1}{26}L\right\} \\
&\subseteq \left\{\max_{\substack{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq l-\sigma_n \leq k-\tau_n}} \left|\frac{n}{k-\tau_n} A_{1,5}(k,l)\right|^p \geq \left(\frac{1}{26}L\right)^p\right\}.
\end{aligned}$$

We find that

$$\begin{aligned}
\left|\frac{n}{k-\tau_n}A_{1,5}(k,l)\right|^p &= \left|\frac{2n\tau_n(\alpha-\beta)}{(k-\tau_n)nk}S_{\tau_n,k}\right|^p && \text{by def. of } A_{1,5}\\
&= 2^p|\alpha-\beta|^p\tau_n^p k^{-p}(k-\tau_n)^{-p}\left|S_{\tau_n,k}\right|^p\\
&< 2^p|\alpha-\beta|^p(k-\tau_n)^{-p}\left|S_{\tau_n,k}\right|^p && \text{by } k>\tau_n\\
&\leq C(k-\tau_n)^{-p}\left|S_{\tau_n,k}\right|^p.
\end{aligned}$$

Applying Chow's Inequality (Lemma 2.9) and Lemma A.6 similarly to (4.22) yields

$$\begin{aligned}
\mathbb{P}\left[E_{1,5}\right] &\leq \mathbb{P}\left[\max_{\lfloor x\rfloor\leq k-\tau_n\leq\lfloor n\delta\rfloor}(k-\tau_n)^{-p}\left|S_{\tau_n,k}\right|^p\geq C\right]\\
&\leq C\left[\lfloor n\delta\rfloor^{-p}\mathbb{E}\left[\left|\tilde{S}_{\tau_n,\lfloor n\delta\rfloor}\right|^p\right]+p\sum_{m=\lfloor x\rfloor}^{\lfloor n\delta\rfloor-1}m^{-(p+1)}E\left[\left|\tilde{S}_{\tau_n,m}\right|^p\right]\right]. \qquad (4.23)
\end{aligned}$$

We get

$$\begin{aligned}
\mathbb{P}\left[E_{1,5}\right] &\leq CM_p\left[\lfloor n\delta\rfloor^{-p/2}+p\sum_{m=\lfloor x\rfloor}^{\lfloor n\delta\rfloor-1}m^{-(p/2+1)}\right] && \text{by Cor. 2.16} \qquad (4.24)\\
&\leq C\left(\lfloor n\delta\rfloor^{-p/2}+2(\lfloor x\rfloor-1)^{-p/2}\right) && \text{by Lem. A.5, } M_p<\infty.
\end{aligned}$$

Note that the properties of the floor function give $\lfloor x\rfloor-1\geq x-2\geq\frac{1}{2}x$ for $x\geq 4$. By (4.19), we infer that

$$\mathbb{P}\left[E_{1,5}\right]\leq C\left(n^{-p/2}+x^{-p/2}\right).$$

We next consider $E_{1,6}$. It holds

$$\begin{aligned}
E_{1,6} &= \left\{\max_{\substack{\lfloor x\rfloor\leq k-\tau_n\leq\lfloor n\delta\rfloor\\ 1\leq l-\sigma_n\leq k-\tau_n}}\frac{n}{k-\tau_n}A_{1,6}(k,l)\geq\frac{1}{26}L\right\}\\
&\subseteq\left\{\max_{\substack{\lfloor x\rfloor\leq k-\tau_n\leq\lfloor n\delta\rfloor\\ 1\leq l-\sigma_n\leq k-\tau_n}}\left|\frac{n}{k-\tau_n}A_{1,6}(k,l)\right|^p\geq\left(\frac{1}{26}L\right)^p\right\}.
\end{aligned}$$

We obtain

$$\begin{aligned}\left|\frac{n}{k-\tau_n}A_{1,6}(k,l)\right|^p &= \left|\frac{-2n}{(k-\tau_n)n(\sigma_n-\tau_n)}S_{\tau_n,k}S_{k,\sigma_n}\right|^p && \text{by def. of } A_{1,6}\\ &= 2^p(\sigma_n-\tau_n)^{-p}(k-\tau_n)^{-p}\left|S_{\tau_n,k}S_{k,\sigma_n}\right|^p\\ &\le Cn^{-p}(k-\tau_n)^{-p}\left|S_{\tau_n,k}S_{k,\sigma_n}\right|^p && \text{by (4.19).}\end{aligned}$$

The independence of the observations $X_1, ..., X_n$ and (4.17) lead to the independence of $S_{\tau_n,k}$ and S_{k,σ_n} for each $k \in \{\tau_n + \lfloor x\rfloor, ..., \tau_n + \lfloor n\delta\rfloor\}$. Applying subadditivity of $\mathbb{P}$ and Markov's Inequality similarly to (4.20) and (4.21) yields

$$\begin{aligned}\mathbb{P}\left[E_{1,6}\right] &\le \mathbb{P}\left[\max_{\lfloor x\rfloor\le k-\tau_n\le\lfloor n\delta\rfloor}(k-\tau_n)^{-p}\left|S_{\tau_n,k}S_{k,\sigma_n}\right|^p \ge Cn^p\right]\\ &\le Cn^{-p}\sum_{k=\tau_n+\lfloor x\rfloor}^{\tau_n+\lfloor n\delta\rfloor}(k-\tau_n)^{-p}\mathbb{E}\left[|S_{\tau_n,k}|^p\right]\mathbb{E}\left[|S_{k,\sigma_n}|^p\right].\end{aligned}$$

Since $k > \tau_n$, we have $\sigma_n - k \le \sigma_n - \tau_n \le n$. It follows that

$$\begin{aligned}\mathbb{P}\left[E_{1,6}\right] &\le CM_p^2n^{-p}\sum_{k=\tau_n+\lfloor x\rfloor}^{\tau_n+\lfloor n\delta\rfloor}(k-\tau_n)^{-p/2}(\sigma_n-k)^{p/2} && \text{by Cor. 2.16}\\ &\le Cn^{-p/2}\sum_{k=\tau_n+\lfloor x\rfloor}^{\tau_n+\lfloor n\delta\rfloor}(k-\tau_n)^{-p/2} && \text{by } \sigma_n-k\le n,\ M_p<\infty\\ &\le Cn^{-p/2},\end{aligned}$$

since $\sum_{m=1}^{\infty} m^{-p/2} < \infty$. We next consider $E_{1,7}$. It holds

$$\begin{aligned}E_{1,7} &= \left\{\max_{\substack{\lfloor x\rfloor\le k-\tau_n\le\lfloor n\delta\rfloor\\ 1\le l-\sigma_n\le k-\tau_n}}\frac{n}{k-\tau_n}A_{1,7}(k,l) \ge \frac{1}{26}L\right\}\\ &\subseteq \left\{\max_{\substack{\lfloor x\rfloor\le k-\tau_n\le\lfloor n\delta\rfloor\\ 1\le l-\sigma_n\le k-\tau_n}}\left|\frac{n}{k-\tau_n}A_{1,7}(k,l)\right|^{p/2} \ge \left(\frac{1}{26}L\right)^{p/2}\right\}.\end{aligned}$$

Since $\sigma_n < l$, we have $\sigma_n - \tau_n - (l-k) < k - \tau_n$. We thus get

$$\begin{aligned}
&\left|\frac{n}{k-\tau_n} A_{1,7}(k,l)\right|^{p/2} \\
&= \left|\frac{n}{(k-\tau_n)n}\left(\frac{1}{l-k} - \frac{1}{\sigma_n - \tau_n}\right) S_{k,\sigma_n}^2\right|^{p/2} && \text{by def. of } A_{1,7} \\
&= (\sigma_n - \tau_n)^{-p/2}(l-k)^{-p/2}(\sigma_n - \tau_n - (l-k))^{p/2}(k-\tau_n)^{-p/2}\left|S_{k,\sigma_n}\right|^p \\
&< (\sigma_n - \tau_n)^{-p/2}(\sigma_n - k)^{-p/2}\left|S_{k,\sigma_n}\right|^p \\
&\le Cn^{-p/2}(\sigma_n - k)^{-p/2}\left|S_{k,\sigma_n}\right|^p && \text{by (4.19).}
\end{aligned}$$

Put $\tilde{S}_{m,\sigma_n} := \sum_{i=1}^{m}(X_{\sigma_n - i+1} - \mathbb{E}\left[X_{\sigma_n-i+1}\right])$ for $\sigma_n - \tau_n - \lfloor n\delta \rfloor \le m \le \sigma_n - \tau_n - \lfloor x \rfloor$. Notice that $\left(\left|\tilde{S}_{m,\sigma_n}\right|^p\right)_{\sigma_n - \tau_n - \lfloor n\delta \rfloor \le m \le \sigma_n - \tau_n - \lfloor x \rfloor}$ is a non-negative submartingale by Lemma 2.14 and $(m^{-p})_{\sigma_n - \tau_n - \lfloor n\delta \rfloor \le m \le \sigma_n - \tau_n - \lfloor x \rfloor}$ is a non-increasing sequence. By an index transformation and similar arguments used in the case $E_{1,4}$, we see that

$$\begin{aligned}
\mathbb{P}\left[E_{1,7}\right] &\le \mathbb{P}\left[\max_{\lfloor x \rfloor \le k - \tau_n \le \lfloor n\delta \rfloor} (\sigma_n - k)^{-p/2}\left|S_{k,\sigma_n}\right|^p \ge Cn^{p/2}\right] \\
&= \mathbb{P}\left[\max_{\sigma_n - \tau_n - \lfloor n\delta \rfloor \le m \le \sigma_n - \tau_n - \lfloor x \rfloor} m^{-p/2}\left|\tilde{S}_{m,\sigma_n}\right|^p \ge Cn^{p/2}\right] \\
&\le Cn^{-p/2}\ln(n).
\end{aligned}$$

We next consider $E_{1,8}$. It holds

$$\begin{aligned}
E_{1,8} &= \left\{\max_{\substack{\lfloor x \rfloor \le k - \tau_n \le \lfloor n\delta \rfloor \\ 1 \le l - \sigma_n \le k - \tau_n}} \frac{n}{k-\tau_n} A_{1,8}(k,l) \ge \frac{1}{26} L\right\} \\
&\subseteq \left\{\max_{\substack{\lfloor x \rfloor \le k - \tau_n \le \lfloor n\delta \rfloor \\ 1 \le l - \sigma_n \le k - \tau_n}} \left|\frac{n}{k-\tau_n} A_{1,8}(k,l)\right|^p \ge \left(\frac{1}{26} L\right)^p\right\}.
\end{aligned}$$

Since $l > \sigma_n$ and $k \le \tau_n + \lfloor n\delta \rfloor$, we have $l - k > \sigma_n - \tau_n - \lfloor n\delta \rfloor$. We obtain

$$\begin{aligned}
&\left| \frac{n}{k-\tau_n} A_{1,8}(k,l) \right|^p \\
&= \left| \frac{2n(l-\sigma_n)(\gamma-\beta)}{(k-\tau_n)n(l-k)} S_{k,\sigma_n} \right|^p && \text{by def. of } A_{1,8} \\
&= 2^p |\beta-\gamma|^p (l-k)^{-p} (l-\sigma_n)^p (k-\tau_n)^{-p} \left| S_{k,\sigma_n} \right|^p \\
&< 2^p |\beta-\gamma|^p (\sigma_n - \tau_n - \lfloor n\delta \rfloor)^{-p} \left| S_{k,\sigma_n} \right|^p && \text{by } l-k > \sigma_n - \tau_n - \lfloor n\delta \rfloor,\ l-\sigma_n \le k-\tau_n \\
&\le C n^{-p} \left| S_{k,\sigma_n} \right|^p && \text{by (4.19).}
\end{aligned}$$

In the previous case we have seen that $\left(\left| \tilde{S}_{m,\sigma_n} \right|^p \right)_{\sigma_n - \tau_n - \lfloor n\delta \rfloor \le m \le \sigma_n - \tau_n - \lfloor x \rfloor}$ is a non-negative submartingale. An index transformation leads to

$$\begin{aligned}
\mathbb{P}\left[E_{1,8}\right] &\le \mathbb{P}\left[\max_{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor} \left| S_{k,\sigma_n} \right|^p \ge C n^p \right] \\
&= \mathbb{P}\left[\max_{\sigma_n - \tau_n - \lfloor n\delta \rfloor \le m \le \sigma_n - \tau_n - \lfloor x \rfloor} \left| \tilde{S}_{m,\sigma_n} \right|^p \ge C n^p \right] \\
&\le C n^{-p} \mathbb{E}\left[\left| \tilde{S}_{\sigma_n - \tau_n - \lfloor x \rfloor, \sigma_n} \right|^p \right] && \text{by Doob In.} \\
&\le C M_p n^{-p} (\sigma_n - \tau_n - \lfloor x \rfloor)^{p/2} && \text{by Cor. 2.16} \\
&\le C n^{-p/2} && \text{by } \sigma_n - \tau_n - \lfloor x \rfloor \le n,\ M_p < \infty.
\end{aligned}$$

We next consider $E_{1,9}$. It holds

$$\begin{aligned}
E_{1,9} &= \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le l-\sigma_n \le k-\tau_n}} \frac{n}{k-\tau_n} A_{1,9}(k,l) \ge \frac{1}{26} L \right\} \\
&\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le l-\sigma_n \le k-\tau_n}} \left| \frac{n}{k-\tau_n} A_{1,9}(k,l) \right| \ge \frac{1}{26} L \right\}.
\end{aligned}$$

Since $l > \sigma_n$ and $k \le \tau_n + \lfloor n\delta \rfloor$, we have $l - k > \sigma_n - \tau_n - \lfloor n\delta \rfloor$. We get

$$\begin{aligned}
\left| \frac{n}{k-\tau_n} A_{1,9}(k,l) \right| &= \left| \frac{2n}{(k-\tau_n)n(l-k)} S_{k,\sigma_n} S_{\sigma_n,l} \right| && \text{by def. of } A_{1,9} \\
&= 2(l-k)^{-1}(k-\tau_n)^{-1} \left| S_{k,\sigma_n} S_{\sigma_n,l} \right| \\
&< 2(\sigma_n - \tau_n - \lfloor n\delta \rfloor)^{-1} (l-\sigma_n)^{-1} \left| S_{k,\sigma_n} S_{\sigma_n,l} \right| \\
&\le C n^{-1} (l-\sigma_n)^{-1} \left| S_{k,\sigma_n} S_{\sigma_n,l} \right| && \text{by (4.19).}
\end{aligned}$$

The penultimate inequality follows from $l-k > \sigma_n - \tau_n - \lfloor n\delta \rfloor$ and $k - \tau_n \geq l - \sigma_n$. It follows that

$$\begin{aligned}
E_{1,9} &\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq l-\sigma_n \leq k-\tau_n}} (l-\sigma_n)^{-1} \left| S_{k,\sigma_n} S_{\sigma_n,l} \right| \geq Cn \right\} \\
&\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq l-\sigma_n \leq \lfloor n\delta \rfloor}} (l-\sigma_n)^{-1} \left| S_{k,\sigma_n} S_{\sigma_n,l} \right| \geq Cn \right\} && \text{by } k-\tau_n \leq \lfloor n\delta \rfloor \\
&= \left\{ \left(\max_{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor} \left| S_{k,\sigma_n} \right| \right) \left(\max_{1 \leq l-\sigma_n \leq \lfloor n\delta \rfloor} (l-\sigma_n)^{-1} \left| S_{\sigma_n,l} \right| \right) \geq Cn \right\} \\
&=: \{UV \geq Cn\}.
\end{aligned}$$

The independence of the observations $X_1, ..., X_n$ and (4.17) lead to the independence of the vectors $\left(S_{\tau_n+\lfloor x \rfloor,\sigma_n}, ..., S_{\tau_n+\lfloor n\delta \rfloor,\sigma_n}\right)$ and $\left(S_{\sigma_n,\sigma_n+1}, ..., S_{\sigma_n,\sigma_n+\lfloor n\delta \rfloor}\right)$, which establishes the independence of U and V. We infer that

$$\begin{aligned}
\mathbb{P}\left[E_{1,9}\right] &\leq \mathbb{P}\left[UV \geq Cn\right] \\
&= \int_{(0,\infty)} \mathbb{P}\left[V \geq Cnu^{-1}\right] \, \mathbb{P}_U(\mathrm{d}u) && \text{by Lem. A.9} \\
&\leq \int_{(0,\infty)} \mathbb{P}\left[V^p \geq Cn^p u^{-p}\right] \, \mathbb{P}_U(\mathrm{d}u). && (4.25)
\end{aligned}$$

To treat the integrand, we write $\tilde{S}_{\sigma_n,m} := \sum_{i=1}^{m}(X_{\sigma_n+i} - \mathbb{E}\left[X_{\sigma_n+i}\right])$ for $1 \leq m \leq \lfloor n\delta \rfloor$. Observe that $\left(\left|\tilde{S}_{\sigma_n,m}\right|^p\right)_{1 \leq m \leq \lfloor n\delta \rfloor}$ is a non-negative submartingale by Lemma 2.14 and $(m^{-p})_{1 \leq m \leq \lfloor n\delta \rfloor}$ is a non-increasing sequence. By an index transformation and similar arguments used to get (4.23) and (4.24), we can deduce that

$$\begin{aligned}
&\mathbb{P}\left[V^p \geq Cn^p u^{-p}\right] \\
&\quad = \mathbb{P}\left[\left(\max_{1\leq l-\sigma_n\leq \lfloor n\delta\rfloor}(l-\sigma_n)^{-1}\left|S_{\sigma_n,l}\right|\right)^p \geq Cn^p u^{-p}\right] \\
&\quad = \mathbb{P}\left[\max_{1\leq m\leq \lfloor n\delta\rfloor} m^{-p}\left|\tilde{S}_{\sigma_n,m}\right|^p \geq Cn^p u^{-p}\right] \\
&\quad \leq CM_p n^{-p}u^p\left[\lfloor n\delta\rfloor^{-p/2} + p\sum_{m=1}^{\lfloor n\delta\rfloor-1} m^{-(p/2+1)}\right] \\
&\quad \leq Cn^{-p}u^p\left[\lfloor n\delta\rfloor^{-p/2} + p\sum_{m=1}^{\infty} m^{-(p/2+1)}\right] && \text{by } M_p<\infty. \qquad (4.26)
\end{aligned}$$

Since the series converges for $p > 2$, there exists a constant $\tilde{C}_1 > 0$ such that $\lfloor n\delta\rfloor^{-p/2} + p\sum_{m=1}^{\infty} m^{-(p/2+1)} \leq \tilde{C}_1$. Consequently,

$$\mathbb{P}\left[V^p \geq Cn^p u^{-p}\right] \leq Cn^{-p}u^p.$$

We conclude that

$$\begin{aligned}
\mathbb{P}\left[E_{1,9}\right] &\leq Cn^{-p}\int_{(0,\infty)} u^p\ \mathbb{P}_U(\mathrm{d}u) && \text{by (4.25)} \\
&\leq Cn^{-p}\int_{\mathbb{R}} u^p\ \mathbb{P}_U(\mathrm{d}u) \\
&= Cn^{-p}\mathbb{E}[U^p]. && (4.27)
\end{aligned}$$

Moreover, in the case $E_{1,7}$ we have seen that $\left(\left|\tilde{S}_{m,\sigma_n}\right|^p\right)_{\sigma_n-\tau_n-\lfloor n\delta\rfloor\leq m\leq\sigma_n-\tau_n-\lfloor x\rfloor}$ is a non-negative submartingale. An index transformation gives

$$\begin{aligned}
\mathbb{E}\left[U^p\right] &= \mathbb{E}\left[\left(\max_{\lfloor x\rfloor\leq k-\tau_n\leq\lfloor n\delta\rfloor}\left|S_{k,\sigma_n}\right|\right)^p\right] \\
&= \mathbb{E}\left[\max_{\sigma_n-\tau_n-\lfloor n\delta\rfloor\leq m\leq\sigma_n-\tau_n-\lfloor x\rfloor}\left|\tilde{S}_{m,\sigma_n}\right|^p\right] \\
&\leq C\mathbb{E}\left[\left|\tilde{S}_{\sigma_n-\tau_n-\lfloor x\rfloor,\sigma_n}\right|^p\right] && \text{by Doob In.} \\
&\leq CM_p(\sigma_n-\tau_n-\lfloor x\rfloor)^{p/2} && \text{by Cor. 2.16} \\
&\leq Cn^{p/2} && \text{by } \sigma_n-\tau_n-\lfloor x\rfloor\leq n,\ M_p<\infty.
\end{aligned}$$

By (4.27), the result is

$$\mathbb{P}\left[E_{1,9}\right] \leq Cn^{-p/2}.$$

We next consider $E_{1,10}$. Since $l > \sigma_n$ and $k \leq \tau_n + \lfloor n\delta \rfloor$, we have $l - k > \sigma_n - \tau_n - \lfloor n\delta \rfloor$. By $(k-\tau_n)^{-1} > 0$ and $S^2_{\sigma_n,l} \geq 0$, we can assert that

$$\begin{aligned}
&\frac{n}{k-\tau_n} A_{1,10}(k,l) \\
&= \frac{n}{(k-\tau_n)n}\left(\frac{1}{l-k} - \frac{1}{n-\sigma_n}\right) S^2_{\sigma_n,l} && \text{by def. of } A_{1,10} \\
&\leq (l-k)^{-1}(k-\tau_n)^{-1} S^2_{\sigma_n,l} && \text{by } (l-k)^{-1} - (n-\sigma_n)^{-1} \leq (l-k)^{-1} \\
&< (\sigma_n - \tau_n - \lfloor n\delta \rfloor)^{-1}(l-\sigma_n)^{-1} S^2_{\sigma_n,l} && \text{by } l-k > \sigma_n - \tau_n - \lfloor n\delta \rfloor,\ k-\tau_n \geq l - \sigma_n \\
&\leq Cn^{-1}(l-\sigma_n)^{-1} S^2_{\sigma_n,l} && \text{by (4.19).}
\end{aligned}$$

We obtain

$$\begin{aligned}
E_{1,10} &= \left\{ \max_{\substack{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq l-\sigma_n \leq k-\tau_n}} \frac{n}{k-\tau_n} A_{1,10}(k,l) \geq \frac{1}{26} L \right\} \\
&\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq l-\sigma_n \leq k-\tau_n}} (l-\sigma_n)^{-1} S^2_{\sigma_n,l} \geq Cn \right\} \\
&\subseteq \left\{ \max_{1 \leq l-\sigma_n \leq \lfloor n\delta \rfloor} (l-\sigma_n)^{-1} S^2_{\sigma_n,l} \geq Cn \right\} && \text{by } k-\tau_n \leq \lfloor n\delta \rfloor \\
&\subseteq \left\{ \max_{1 \leq l-\sigma_n \leq \lfloor n\delta \rfloor} (l-\sigma_n)^{-p/2} |S_{\sigma_n,l}|^p \geq Cn^{p/2} \right\}.
\end{aligned}$$

In the previous case we have seen that $\left(\left|\tilde{S}_{\sigma_n,m}\right|^p\right)_{1 \leq m \leq \lfloor n\delta \rfloor}$ is a non-negative submartingale. Furthermore, $\left(m^{-p/2}\right)_{1 \leq m \leq \lfloor n\delta \rfloor}$ is a non-increasing sequence. An index transformation and similar arguments applied in the case $E_{1,4}$ lead to

$$\begin{aligned}
\mathbb{P}\left[E_{1,10}\right] &\leq \mathbb{P}\left[\max_{1 \leq l-\sigma_n \leq \lfloor n\delta \rfloor} (l-\sigma_n)^{-p/2} |S_{\sigma_n,l}|^p \geq Cn^{p/2}\right] \\
&= \mathbb{P}\left[\max_{1 \leq m \leq \lfloor n\delta \rfloor} m^{-p/2} \left|\tilde{S}_{\sigma_n,m}\right|^p \geq Cn^{p/2}\right] \\
&\leq Cn^{-p/2} \ln(n).
\end{aligned}$$

We next consider $E_{1,11}$. Fix an arbitrary $\kappa \in \left(0, \frac{1}{2}\right)$. It holds

$$\begin{aligned} E_{1,11} &= \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le l-\sigma_n \le k-\tau_n}} \frac{n}{k-\tau_n} A_{1,11}(k,l) \ge \frac{1}{26} L \right\} \\ &\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le l-\sigma_n \le k-\tau_n}} \left| \frac{n}{k-\tau_n} A_{1,11}(k,l) \right|^p \ge \left(\frac{1}{26} L \right)^p \right\}. \end{aligned}$$

Note that the properties of the floor function give $k - \tau_n \ge \lfloor x \rfloor > x - 1 \ge \frac{1}{2}x$ for $x \ge 2$. Hence

$$\begin{aligned} &\left| \frac{n}{k-\tau_n} A_{1,11}(k,l) \right|^p \\ &\quad = \left| \frac{2n(\sigma_n - k)(\beta-\gamma)}{(k-\tau_n)n(l-k)} S_{\sigma_n,l} \right|^p && \text{by def. of } A_{1,11} \\ &\quad = 2^p |\beta-\gamma|^p (k-\tau_n)^{-(1/2-\kappa)p} (k-\tau_n)^{-(1/2+\kappa)p} (\sigma_n-k)^p (l-k)^{-p} \left|S_{\sigma_n,l}\right|^p \\ &\quad < 2^{(3/2-\kappa)p} |\beta-\gamma|^p x^{-(1/2-\kappa)p} (l-\sigma_n)^{-(1/2+\kappa)p} \left|S_{\sigma_n,l}\right|^p \\ &\quad \le C x^{-(1/2-\kappa)p} (l-\sigma_n)^{-(1/2+\kappa)p} \left|S_{\sigma_n,l}\right|^p. \end{aligned}$$

The penultimate inequality follows from $k - \tau_n > \frac{1}{2}x$, $k - \tau_n \ge l - \sigma_n$ and $l > \sigma_n$. Hence

$$\begin{aligned} E_{1,11} &\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le l-\sigma_n \le k-\tau_n}} (l-\sigma_n)^{-(1/2+\kappa)p} \left|S_{\sigma_n,l}\right|^p \ge C x^{(1/2-\kappa)p} \right\} \\ &\subseteq \left\{ \max_{1 \le l-\sigma_n \le \lfloor n\delta \rfloor} (l-\sigma_n)^{-(1/2+\kappa)p} \left|S_{\sigma_n,l}\right|^p \ge C x^{(1/2-\kappa)p} \right\} && \text{by } k-\tau_n \le \lfloor n\delta \rfloor. \end{aligned}$$

We can now proceed analogously to (4.26). We conclude that

$$\begin{aligned} \mathbb{P}\left[E_{1,11}\right] &\le \mathbb{P}\left[\max_{1 \le l-\sigma_n \le \lfloor n\delta \rfloor} (l-\sigma_n)^{-(1/2+\kappa)p} \left|S_{\sigma_n,l}\right|^p \ge C x^{(1/2-\kappa)p} \right] \\ &= \mathbb{P}\left[\max_{1 \le m \le \lfloor n\delta \rfloor} m^{-(1/2+\kappa)p} \left|\tilde{S}_{\sigma_n,m}\right|^p \ge C x^{(1/2-\kappa)p} \right] \\ &\le C x^{-(1/2-\kappa)p} \left[\lfloor n\delta \rfloor^{-\kappa p} + \left(\frac{1}{2} + \kappa \right) p \sum_{m=1}^{\infty} m^{-(\kappa p+1)} \right]. \end{aligned}$$

The series converges, because we find that $\kappa p + 1 > 1$ for $\kappa > 0$ and $p > 2$. Therefore, there exists $\tilde{C}_2 > 0$ such that $\lfloor n\delta \rfloor^{-\kappa p} + (1/2 + \kappa)p \sum_{m=1}^{\infty} m^{-(\kappa p+1)} \le \tilde{C}_2$. This gives

$$\mathbb{P}\left[E_{1,11}\right] \le C x^{-(1/2-\kappa)p}.$$

We next consider $E_{1,12}$. It holds

$$\begin{aligned} E_{1,12} &= \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le l-\sigma_n \le k-\tau_n}} \frac{n}{k-\tau_n} A_{1,12}(k,l) \ge \frac{1}{26} L \right\} \\ &\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le l-\sigma_n \le k-\tau_n}} \left| \frac{n}{k-\tau_n} A_{1,12}(k,l) \right|^p \ge \left(\frac{1}{26} L \right)^p \right\}. \end{aligned}$$

We have

$$\begin{aligned} \left| \frac{n}{k-\tau_n} A_{1,12}(k,l) \right|^p &= \left| \frac{-2n}{(k-\tau_n)n(n-\sigma_n)} S_{\sigma_n,l} S_{l,n} \right|^p && \text{by def. of } A_{1,12} \\ &= 2^p (n-\sigma_n)^{-p} (k-\tau_n)^{-p} \left| S_{\sigma_n,l} S_{l,n} \right|^p \\ &\le 2^p (n-\sigma_n)^{-p} (l-\sigma_n)^{-p} \left| S_{\sigma_n,l} S_{l,n} \right|^p && \text{by } k-\tau_n \ge l-\sigma_n \\ &\le C n^{-p} (l-\sigma_n)^{-p} \left| S_{\sigma_n,l} S_{l,n} \right|^p && \text{by (4.19).} \end{aligned}$$

The independence of the observations $X_1, ..., X_n$ and (4.17) lead to the independence of $S_{\sigma_n,l}$ and $S_{l,n}$ for each $l \in \{\sigma_n + 1, ..., \sigma_n + \lfloor n\delta \rfloor\}$. We see that

$$\begin{aligned} E_{1,12} &\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le l-\sigma_n \le k-\tau_n}} (l-\sigma_n)^{-p} \left| S_{\sigma_n,l} S_{l,n} \right|^p \ge C n^p \right\} \\ &\subseteq \left\{ \max_{1 \le l-\sigma_n \le \lfloor n\delta \rfloor} (l-\sigma_n)^{-p} \left| S_{\sigma_n,l} S_{l,n} \right|^p \ge C n^p \right\} && \text{by } k-\tau_n \le \lfloor n\delta \rfloor. \end{aligned}$$

Applying similar arguments used in the case $E_{1,6}$ yields

$$\begin{aligned} \mathbb{P}\left[E_{1,12} \right] &\le \mathbb{P}\left[\max_{1 \le l-\sigma_n \le \lfloor n\delta \rfloor} (l-\sigma_n)^{-p} \left| S_{\sigma_n,l} S_{l,n} \right|^p \ge C n^p \right] \\ &\le C n^{-p/2}. \end{aligned}$$

We next consider $E_{1,13}$. It holds

$$\begin{aligned} E_{1,13} &= \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le l-\sigma_n \le k-\tau_n}} \frac{n}{k-\tau_n} A_{1,13}(k,l) \ge \frac{1}{26} L \right\} \\ &\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le l-\sigma_n \le k-\tau_n}} \left| \frac{n}{k-\tau_n} A_{1,13}(k,l) \right|^{p/2} \ge \left(\frac{1}{26} L \right)^{p/2} \right\}. \end{aligned}$$

We find that

$$\begin{aligned}
\left|\frac{n}{k-\tau_n}A_{1,13}(k,l)\right|^{p/2} &= \left|\frac{n}{(k-\tau_n)n}\left(\frac{1}{n-l}-\frac{1}{n-\sigma_n}\right)S_{l,n}^2\right|^{p/2} && \text{by def. of } A_{1,13}\\
&= (n-\sigma_n)^{-p/2}(n-l)^{-p/2}(k-\tau_n)^{-p/2}(l-\sigma_n)^{p/2}\left|S_{l,n}\right|^p\\
&\leq (n-\sigma_n)^{-p/2}(n-l)^{-p/2}\left|S_{l,n}\right|^p && \text{by } k-\tau_n \geq l-\sigma_n\\
&\leq Cn^{-p/2}(n-l)^{-p/2}\left|S_{l,n}\right|^p && \text{by (4.19).}
\end{aligned}$$

Hence

$$\begin{aligned}
E_{1,13} &\subseteq \left\{\max_{\substack{\lfloor x\rfloor \leq k-\tau_n \leq \lfloor n\delta\rfloor\\ 1\leq l-\sigma_n\leq k-\tau_n}}(n-l)^{-p/2}\left|S_{l,n}\right|^p \geq Cn^{p/2}\right\}\\
&\subseteq \left\{\max_{1\leq l-\sigma_n\leq\lfloor n\delta\rfloor}(n-l)^{-p/2}\left|S_{l,n}\right|^p \geq Cn^{p/2}\right\} && \text{by } k-\tau_n\leq\lfloor n\delta\rfloor.
\end{aligned}$$

Write $\tilde{S}_{m,n} := \sum_{i=1}^m (X_{n-i+1} - \mathbb{E}\left[X_{n-i+1}\right])$ for $n-\sigma_n-\lfloor n\delta\rfloor \leq m \leq n-\sigma_n-1$. Notice that $\left(\left|\tilde{S}_{m,n}\right|^p\right)_{n-\sigma_n-\lfloor n\delta\rfloor\leq m\leq n-\sigma_n-1}$ is a non-negative submartingale by Lemma 2.14 and $\left(m^{-p/2}\right)_{n-\sigma_n-\lfloor n\delta\rfloor\leq m\leq n-\sigma_n-1}$ is a non-increasing sequence. By similar arguments applied in the case $E_{1,4}$, we obtain

$$\begin{aligned}
\mathbb{P}\left[E_{1,13}\right] &\leq \mathbb{P}\left[\max_{1\leq l-\sigma_n\leq\lfloor n\delta\rfloor}(n-l)^{-p/2}\left|S_{l,n}\right|^p \geq Cn^{p/2}\right]\\
&= \mathbb{P}\left[\max_{n-\sigma_n-\lfloor n\delta\rfloor\leq m\leq n-\sigma_n-1} m^{-p/2}\left|\tilde{S}_{m,n}\right|^p \geq Cn^{p/2}\right]\\
&\leq Cn^{-p/2}\ln(n).
\end{aligned}$$

Altogether, by (4.18), the previous estimates provide

$$\mathbb{P}[E_1] \leq \sum_{i=1}^{13}\mathbb{P}[E_{1,i}] \leq C\left(n^{-p/2} + n^{-p/2}\ln(n) + x^{-p/2} + x^{-(1/2-\kappa)p}\right), \tag{4.28}$$

where $x \geq 4$ and $\kappa \in \left(0, \frac{1}{2}\right)$. We next consider E_2, which is given in (4.16). We will see that we get a substantially deteriorate estimate. Observe that

$$\tau_n + x \leq k \leq \tau_n + n\delta < \sigma_n - n\delta \leq l \leq \sigma_n < n. \tag{4.29}$$

by Lemma 3.19. Hence $\left(\frac{k}{n},\frac{l}{n}\right) \in \Theta^4 \cap \Theta_n$. By Lemma B.1 and the Binomial Formula, a trivial verification shows that

$$\begin{aligned}
&\hat{\delta}_n\left(\frac{k}{n},\frac{l}{n}\right) - \hat{\delta}_n\left(\frac{\tau_n}{n},\frac{\sigma_n}{n}\right)\\
&= \frac{1}{nk}(S_{0,\tau_n}+S_{\tau_n,k})^2 + \frac{2(\tau_n\alpha+(k-\tau_n)\beta)}{nk}(S_{0,\tau_n}+S_{\tau_n,k}) + \frac{1}{n(l-k)}S_{k,l}^2\\
&\quad + \frac{2\beta}{n}S_{k,l} + \frac{1}{n(n-l)}(S_{l,\sigma_n}+S_{\sigma_n,n})^2 + \frac{2((\sigma_n-l)\beta+(n-\sigma_n)\gamma)}{n(n-l)}(S_{l,\sigma_n}+S_{\sigma_n,n})\\
&\quad - \Bigg[\frac{1}{n\tau_n}S_{0,\tau_n}^2 + \frac{2\alpha}{n}S_{0,\tau_n} + \frac{1}{n(\sigma_n-\tau_n)}(S_{\tau_n,k}+S_{k,l}+S_{l,\sigma_n})^2\\
&\qquad + \frac{2\beta}{n}(S_{\tau_n,k}+S_{k,l}+S_{l,\sigma_n}) + \frac{1}{n(n-\sigma_n)}S_{\sigma_n,n}^2 + \frac{2\gamma}{n}S_{\sigma_n,n}\Bigg]\\
&= \frac{1}{n}\left(\frac{1}{k}-\frac{1}{\tau_n}\right)S_{0,\tau_n}^2 + \frac{2(k-\tau_n)(\alpha-\beta)}{nk}S_{0,\tau_n} + \frac{2}{nk}S_{0,\tau_n}S_{\tau_n,k}\\
&\quad + \frac{1}{n}\left(\frac{1}{k}-\frac{1}{\sigma_n-\tau_n}\right)S_{\tau_n,k}^2 + \frac{2\tau_n(\alpha-\beta)}{nk}S_{\tau_n,k} - \frac{2}{n(\sigma_n-\tau_n)}S_{\tau_n,k}S_{k,l}\\
&\quad + \frac{1}{n}\left(\frac{1}{l-k}-\frac{1}{\sigma_n-\tau_n}\right)S_{k,l}^2 - \frac{2}{n(\sigma_n-\tau_n)}S_{k,l}S_{l,\sigma_n} - \frac{2}{n(\sigma_n-\tau_n)}S_{\tau_n,k}S_{l,\sigma_n}\\
&\quad + \frac{1}{n}\left(\frac{1}{n-l}-\frac{1}{\sigma_n-\tau_n}\right)S_{l,\sigma_n}^2 + \frac{2(n-\sigma_n)(\gamma-\beta)}{n(n-l)}S_{l,\sigma_n} + \frac{1}{n(n-l)}S_{l,\sigma_n}S_{\sigma_n,n}\\
&\quad + \frac{1}{n}\left(\frac{1}{n-l}-\frac{1}{n-\sigma_n}\right)S_{\sigma_n,n}^2 + \frac{2(\sigma_n-l)(\beta-\gamma)}{n(n-l)}S_{\sigma_n,n}\\
&=: \sum_{i=1}^{14} A_{2,i}(k,l).
\end{aligned}$$

Observe that for $l=\sigma_n$ we have $A_{2,i}(k,l)=0$, $i\in\{8,...,12\}$. We now proceed analogously to (4.18) and obtain

$$\begin{aligned}
E_2 &\subseteq \bigcup_{i=1}^{14}\left\{\max_{\substack{\lfloor x\rfloor\le k-\tau_n\le\lfloor n\delta\rfloor\\ 1\le\sigma_n-l\le k-\tau_n}} \frac{n}{k-\tau_n}A_{2,i}(k,l) \ge \frac{1}{28}L\right\}\\
&=: \bigcup_{i=1}^{14} E_{2,i}.
\end{aligned} \tag{4.30}$$

We next present further techniques to estimate the probabilities of $E_{2,6}$ and $E_{2,8}$. We first consider $E_{2,6}$. It holds

$$\begin{aligned}E_{2,6} &= \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le \sigma_n - l \le k-\tau_n}} \frac{n}{k-\tau_n} A_{2,6}(k,l) \ge \frac{1}{28} L \right\} \\ &\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le \sigma_n - l \le k-\tau_n}} \left| \frac{n}{k-\tau_n} A_{2,6}(k,l) \right|^p \ge \left(\frac{1}{28} L \right)^p \right\}.\end{aligned}$$

We infer that

$$\begin{aligned}\left| \frac{n}{k-\tau_n} A_{2,6}(k,l) \right|^p &= \left| \frac{-2n}{(k-\tau_n) n (\sigma_n - \tau_n)} S_{\tau_n,k} S_{k,l} \right|^p && \text{by def. of } A_{2,6} \\ &= 2^p (\sigma_n - \tau_n)^{-p} (k-\tau_n)^{-p} \left| S_{\tau_n,k} S_{k,l} \right|^p \\ &\le C n^{-p} (k-\tau_n)^{-p} \left| S_{\tau_n,k} S_{k,l} \right|^p && \text{by (4.19).}\end{aligned}$$

Therefore

$$\begin{aligned}E_{2,6} &\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le \sigma_n - l \le k-\tau_n}} (k-\tau_n)^{-p} \left| S_{\tau_n,k} S_{k,l} \right|^p \ge C n^p \right\} \\ &\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor \\ 1 \le \sigma_n - l \le \lfloor n\delta \rfloor}} (k-\tau_n)^{-p} \left| S_{\tau_n,k} S_{k,l} \right|^p \ge C n^p \right\} && \text{by } k-\tau_n \le \lfloor n\delta \rfloor \\ &\subseteq \left\{ \max_{\lfloor x \rfloor \le k-\tau_n \le \lfloor n\delta \rfloor} \left(\max_{1 \le \sigma_n - l \le \lfloor n\delta \rfloor} (k-\tau_n)^{-p} \left| S_{\tau_n,k} S_{k,l} \right|^p \right) \ge C n^p \right\} \\ &= \bigcup_{k=\tau_n+\lfloor x \rfloor}^{\tau_n+\lfloor n\delta \rfloor} \left\{ \max_{1 \le \sigma_n - l \le \lfloor n\delta \rfloor} \left| S_{\tau_n,k} S_{k,l} \right|^p \ge C n^p (k-\tau_n)^p \right\}. && (4.31)\end{aligned}$$

We next prove that $(|S_{\tau_n,k} S_{k,l}|^p)_{\sigma_n - \lfloor n\delta \rfloor \le l \le \sigma_n - 1}$ is a non-negative submartingale for each $k \in \{\tau_n + \lfloor x \rfloor, ..., \tau_n + \lfloor n\delta \rfloor\}$. For this purpose, fix $k \in \{\tau_n + \lfloor x \rfloor, ..., \tau_n + \lfloor n\delta \rfloor\}$ and let $\mathcal{F}_l := \sigma(X_{\tau_n+1}, ..., X_l)$ be the σ-algebra generated by $X_{\tau_n+1}, ..., X_l$. Then $(\mathcal{F}_l)_{\sigma_n - \lfloor n\delta \rfloor \le l \le \sigma_n - 1}$ is a filtration in $\mathcal{A}$ (σ-algebra in the probability space in our model). The independence of the observations $X_1, ..., X_n$ and (4.29) lead to the independence of $S_{\tau_n,k}$ and $S_{k,l}$ for every $l \in \{\sigma_n - \lfloor n\delta \rfloor, ..., \sigma_n - 1\}$. By Corollary 2.16, there exists a constant $\tilde{C}_3 > 0$ such that

$$\mathbb{E}[|S_{\tau_n,k} S_{k,l}|^p] = \mathbb{E}[|S_{\tau_n,k}|^p] \mathbb{E}[|S_{k,l}|^p] \le \tilde{C}_3 M_p^2 (k-\tau_n)^{p/2} (l-k)^{p/2}$$

for all $l \in \{\sigma_n - \lfloor n\delta \rfloor, ..., \sigma_n - 1\}$, and consequently $\mathbb{E}[|S_{\tau_n,k}S_{k,l}|^p] < \infty$ by $M_p < \infty$. Furthermore, since $k \leq \tau_n + \lfloor n\delta \rfloor < \sigma_n - \lfloor n\delta \rfloor$ by (4.29), it follows that $|S_{\tau_n,k}|^p$ is $\mathcal{F}_l$- measurable. By Lemma 2.14, we see that $(|S_{k,l}|^p)_{\sigma_n - \lfloor n\delta \rfloor \leq l \leq \sigma_n - 1}$ is a non-negative submartingale. We obtain

$$\begin{aligned}
\mathbb{E}\left[|S_{\tau_n,k}S_{k,l+1}|^p \, |\mathcal{F}_l\right] &= |S_{\tau_n,k}|^p \, \mathbb{E}[|S_{k,l+1}|^p \, |\mathcal{F}_l] \\
&\geq |S_{\tau_n,k}|^p \, |S_{k,l}|^p \\
&= |S_{\tau_n,k}S_{k,l}|^p \, .
\end{aligned}$$

We thus conclude that the process $(|S_{\tau_n,k}S_{k,l}|^p)_{\sigma_n - \lfloor n\delta \rfloor \leq l \leq \sigma_n - 1}$ is a non-negative submartingale for each $k \in \{\tau_n + \lfloor x \rfloor, ..., \tau_n + \lfloor n\delta \rfloor\}$ with respect to the filtration $(\mathcal{F}_l)_{\sigma_n - \lfloor n\delta \rfloor \leq l \leq \sigma_n - 1}$. By (4.31) and Doob's Inequality (Lemma 2.10 (i)), it follows that

$$\begin{aligned}
\mathbb{P}\left[E_{2,6}\right] &\leq \sum_{k=\tau_n+\lfloor x \rfloor}^{\tau_n+\lfloor n\delta \rfloor} \mathbb{P}\left[\max_{\sigma_n - \lfloor n\delta \rfloor \leq l \leq \sigma_n - 1} |S_{\tau_n,k}S_{k,l}|^p \geq Cn^p(k-\tau_n)^p\right] \\
&\leq Cn^{-p} \sum_{k=\tau_n+\lfloor x \rfloor}^{\tau_n+\lfloor n\delta \rfloor} (k-\tau_n)^{-p} \mathbb{E}\left[|S_{\tau_n,k}S_{k,\sigma_n-1}|^p\right].
\end{aligned}$$

The independence of the observations $X_1, ..., X_n$ and (4.29) lead to the independence of $S_{\tau_n,k}$ and S_{k,σ_n-1} for each $k \in \{\tau_n + \lfloor x \rfloor, ..., \tau_n + \lfloor n\delta \rfloor\}$. Since $k > \tau_n$, we conclude that $\sigma_n - 1 - k < \sigma_n - \tau_n - 1 \leq n$. It holds

$$\begin{aligned}
\mathbb{P}\left[E_{2,6}\right] &\leq Cn^{-p} \sum_{k=\tau_n+\lfloor x \rfloor}^{\tau_n+\lfloor n\delta \rfloor} (k-\tau_n)^{-p} \mathbb{E}\left[|S_{\tau_n,k}|^p\right] \mathbb{E}\left[|S_{k,\sigma_n-1}|^p\right] \\
&\leq CM_p^2 n^{-p} \sum_{k=\tau_n+\lfloor x \rfloor}^{\tau_n+\lfloor n\delta \rfloor} (k-\tau_n)^{-p/2} (\sigma_n - 1 - k)^{p/2} && \text{by Cor. 2.16} \\
&\leq Cn^{-p/2} \sum_{k=\tau_n+\lfloor x \rfloor}^{\tau_n+\lfloor n\delta \rfloor} (k-\tau_n)^{-p/2} && \text{by } \sigma_n - 1 - k \leq n, \ M_p < \infty.
\end{aligned}$$

Since $\sum_{m=1}^{\infty} m^{-p/2} < \infty$ for $p > 2$, we get

$$\mathbb{P}\left[E_{2,6}\right] \leq Cn^{-p/2}.$$

We next consider $E_{2,8}$. It holds

$$\begin{aligned} E_{2,8} &= \left\{ \max_{\substack{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq \sigma_n - l \leq k - \tau_n}} \frac{n}{k-\tau_n} A_{2,8}(k,l) \geq \frac{1}{28} L \right\} \\ &\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq \sigma_n - l \leq k - \tau_n}} \left| \frac{n}{k-\tau_n} A_{2,8}(k,l) \right|^p \geq \left(\frac{1}{28} L \right)^p \right\}. \end{aligned}$$

We see that

$$\begin{aligned} \left| \frac{n}{k-\tau_n} A_{2,8}(k,l) \right|^p &= \left| \frac{-2n}{(k-\tau_n) n (\sigma_n - \tau_n)} S_{k,l} S_{l,\sigma_n} \right|^p && \text{by def. of } A_{2,8} \\ &= 2^p (\sigma_n - \tau_n)^{-p} (k-\tau_n)^{-p} \left| S_{k,l} S_{l,\sigma_n} \right|^p \\ &\leq 2^p (\sigma_n - \tau_n)^{-p} (\sigma_n - l)^{-p} \left| S_{k,l} S_{l,\sigma_n} \right|^p && \text{by } k - \tau_n \geq \sigma_n - l \\ &\leq C n^{-p} (\sigma_n - l)^{-p} \left| S_{k,l} S_{l,\sigma_n} \right|^p && \text{by (4.19).} \end{aligned}$$

We get

$$\begin{aligned} E_{2,8} &\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq \sigma_n - l \leq k - \tau_n}} (\sigma_n - l)^{-p} \left| S_{k,l} S_{l,\sigma_n} \right|^p \geq C n^p \right\} \\ &\subseteq \left\{ \max_{\substack{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor \\ 1 \leq \sigma_n - l \leq \lfloor n\delta \rfloor}} (\sigma_n - l)^{-p} \left| S_{k,l} S_{l,\sigma_n} \right|^p \geq C n^p \right\} && \text{by } k - \tau_n \leq \lfloor n\delta \rfloor \\ &\subseteq \left\{ \max_{\lfloor x \rfloor \leq k-\tau_n \leq \lfloor n\delta \rfloor} \left(\max_{1 \leq \sigma_n - l \leq \lfloor n\delta \rfloor} (\sigma_n - l)^{-p} \left| S_{k,l} S_{l,\sigma_n} \right|^p \right) \geq C n^p \right\} \\ &= \bigcup_{k=\tau_n+\lfloor x \rfloor}^{\tau_n + \lfloor n\delta \rfloor} \bigcup_{l=\sigma_n - \lfloor n\delta \rfloor}^{\sigma_n - 1} \left\{ \left| S_{k,l} S_{l,\sigma_n} \right|^p \geq C n^p (\sigma_n - l)^p \right\}. \end{aligned}$$

The subadditivity of $\mathbb{P}$ and the Markov Inequality (see Lemma 2.4) imply

$$\begin{aligned} \mathbb{P}\left[E_{2,8} \right] &\leq \sum_{k=\tau_n+\lfloor x \rfloor}^{\tau_n + \lfloor n\delta \rfloor} \sum_{l=\sigma_n - \lfloor n\delta \rfloor}^{\sigma_n - 1} \mathbb{P}\left[\left| S_{k,l} S_{l,\sigma_n} \right|^p \geq C n^p (\sigma_n - l)^p \right] \\ &\leq C n^{-p} \sum_{k=\tau_n+\lfloor x \rfloor}^{\tau_n + \lfloor n\delta \rfloor} \sum_{l=\sigma_n - \lfloor n\delta \rfloor}^{\sigma_n - 1} (\sigma_n - l)^{-p} \mathbb{E}\left[\left| S_{k,l} S_{l,\sigma_n} \right|^p \right]. \end{aligned}$$

The independence of the observations $X_1, ..., X_n$ and (4.29) lead to the independence of $S_{k,l}$ and S_{l,σ_n} for each $k \in \{\tau_n + \lfloor x \rfloor, ..., \tau_n + \lfloor n\delta \rfloor\}$ and $l \in \{\sigma_n - \lfloor n\delta \rfloor, ..., \sigma_n - 1\}$. Since $l < \sigma_n$ and $k > \tau_n$, we have $l - k < \sigma_n - \tau_n \leq n$. We deduce that

$$\begin{aligned}
\mathbb{P}\left[E_{2,8}\right] &\leq Cn^{-p} \sum_{k=\tau_n+\lfloor x \rfloor}^{\tau_n+\lfloor n\delta \rfloor} \sum_{l=\sigma_n-\lfloor n\delta \rfloor}^{\sigma_n-1} (\sigma_n - l)^{-p} \mathbb{E}\left[|S_{k,l}|^p\right] \mathbb{E}\left[|S_{l,\sigma_n}|^p\right] \\
&\leq CM_p^2 n^{-p} \sum_{k=\tau_n+\lfloor x \rfloor}^{\tau_n+\lfloor n\delta \rfloor} \sum_{l=\sigma_n-\lfloor n\delta \rfloor}^{\sigma_n-1} (\sigma_n - l)^{-p/2} (l-k)^{p/2} && \text{by Cor. 2.16} \\
&\leq Cn^{-p/2} (\lfloor n\delta \rfloor - \lfloor x \rfloor + 1) \sum_{l=\sigma_n-\lfloor n\delta \rfloor}^{\sigma_n-1} (\sigma_n - l)^{-p/2} && \text{by } l-k \leq n,\ M_p < \infty \\
&\leq Cn^{-(p/2-1)} \sum_{l=\sigma_n-\lfloor n\delta \rfloor}^{\sigma_n-1} (\sigma_n - l)^{-p/2} && \text{by } \lfloor n\delta \rfloor - \lfloor x \rfloor + 1 \leq n.
\end{aligned}$$

Since $\sum_{m=1}^{\infty} m^{-p/2} < \infty$ for $p > 2$, we can assert that

$$\mathbb{P}\left[E_{2,8}\right] \leq Cn^{-(p/2-1)}.$$

The remaining sets $E_{2,i}$, $i \in \{1, ..., 14\} \backslash \{6, 8\}$, can be handled as before. By (4.30), we get

$$\mathbb{P}[E_2] \leq \sum_{i=1}^{14} \mathbb{P}[E_{2,i}] \leq C\left(n^{-p/2} + n^{-(p/2-1)} + n^{-p/2}\ln(n) + x^{-(1/2-\kappa)p}\right),$$

where $x \geq 4$ and $\kappa \in \left(0, \frac{1}{2}\right)$. Similarly to (4.28) and above, we obtain such upper bounds for the probabilities of E_3 and E_4. By (4.16), we see that

$$\begin{aligned}
\mathbb{P}[E] \leq \sum_{i=1}^{4} \mathbb{P}[E_i] &\leq C\left(n^{-p/2} + n^{-(p/2-1)} + n^{-p/2}\ln(n) + x^{-p/2} + x^{-(1/2-\kappa)p}\right) \\
&\leq C\left(n^{-(p/2-1)} + x^{-(1/2-\kappa)p}\right).
\end{aligned}$$

The last inequality follows from $n^{-p/2} \leq n^{-p/2}\ln(n) \leq n^{-(p/2-1)}$ and $x^{-p/2} \leq x^{-(1/2-\kappa)p}$. The same upper bound can be found for the probability of F on a similar way. By (4.15), the result is

$$\mathbb{P}\left[x \leq \|\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta\right] \leq \mathbb{P}\left[E\right] + \mathbb{P}\left[F\right] \leq C\left(n^{-(p/2-1)} + x^{-(1/2-\kappa)p}\right). \qquad \square$$

We now obtain stochastic boundedness of $\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$.

Proposition 4.17. *Suppose there is some* $p \in (2, \infty)$ *such that* $M_p < \infty$. *Then*

$$\lim_{x\to\infty} \limsup_{n\to\infty} \mathbb{P}[\|\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \geq x] = 0.$$

Proof. The same proceeding as in the proof of Proposition 3.22 leads to

$$\mathbb{P}[\|\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \geq x] \leq \mathbb{P}[x \leq \|\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \leq n\delta] + \mathbb{P}\left[\left\|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right\| > \frac{1}{2}\delta\right] + \mathbb{P}\left[1 > \frac{1}{2}n\delta\right]$$

with $\delta > 0$, $x > 0$ and $n \in \mathbb{N}$. Applying the error estimate in Lemma 4.16 and the weak consistency of $\hat{\boldsymbol{\theta}}_n$ (see Theorem 4.10) we conclude that

$$\limsup_{n\to\infty} \mathbb{P}[\|\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| \geq x] \leq Cx^{-(1/2-\kappa)p},$$

where $C > 0$, $x \geq 4$ and $\kappa \in \left(0, \frac{1}{2}\right)$. Letting $x \to \infty$ yields the claim. $\square$

4.1.4 Consistency of the estimator of expectations

This section contains the proof of weak consistency of $\hat{\boldsymbol{\alpha}}_n$, which is based on the stochastic boundedness of $\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$.

Theorem 4.18. *If* $M_4 < \infty$, *then*

$$\hat{\boldsymbol{\alpha}}_n \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\alpha}.$$

Proof. Let us first recall that

$$\hat{\boldsymbol{\alpha}}_n = \left(\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n\right) \xrightarrow[n\to\infty]{\mathbb{P}} (\alpha, \beta, \gamma) = \boldsymbol{\alpha}$$

if and only if

$$\hat{\alpha}_n \xrightarrow[n\to\infty]{\mathbb{P}} \alpha, \qquad \hat{\beta}_n \xrightarrow[n\to\infty]{\mathbb{P}} \beta \qquad \text{and} \qquad \hat{\gamma}_n \xrightarrow[n\to\infty]{\mathbb{P}} \gamma.$$

We give the proof only for the convergence in probability of $\hat{\beta}_n$ to β. In the same manner we can see the convergence in probability of $\hat{\alpha}_n$ to α and $\hat{\gamma}_n$ to γ. Fix $n \in \mathbb{N}$, $\varepsilon > 0$ and $x \geq 1$. To simplify notation, the fact that some mathematical objects, which are defined in this proof depend on n, ε or x is omitted. Write

$$A := \left\{ \left| \hat{\beta}_n - \beta \right| > \varepsilon \right\}, \quad A_1 := \{ |\hat{\tau}_n - \tau_n| \leq x \} \quad \text{and} \quad A_2 := \{ |\hat{\sigma}_n - \sigma_n| \leq x \} .$$

By the rules of De Morgan, we obtain

$$\begin{aligned} A &= (A \cap (A_1 \cap A_2)) \cup \left(A \cap (A_1 \cap A_2)^{\complement} \right) \\ &\subseteq (A \cap A_1 \cap A_2) \cup \left(A_1^{\complement} \cup A_2^{\complement} \right) . \end{aligned}$$

By definition of $\hat{\beta}_n$, we thus get

$$\begin{aligned} \mathbb{P}\left[\left| \hat{\beta}_n - \beta \right| > \varepsilon \right] \leq \mathbb{P}&\left[\left| \frac{1}{\hat{\sigma}_n - \hat{\tau}_n} \sum_{i=\hat{\tau}_n+1}^{\hat{\sigma}_n} (X_i - \beta) \right| > \varepsilon,\ |\hat{\tau}_n - \tau_n| \leq x,\ |\hat{\sigma}_n - \sigma_n| \leq x \right] \\ &+ \mathbb{P}\left[|\hat{\tau}_n - \tau_n| > x \right] + \mathbb{P}\left[|\hat{\sigma}_n - \sigma_n| > x \right] \\ &=: \mathbb{P}_1 + \mathbb{P}_2 + \mathbb{P}_3. \end{aligned} \tag{4.32}$$

At the end of the proof we apply the stochastic boundedness of $\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$, which is given in Propostion 4.17. Therefore, we only need to estimate the first probability. It follows that

$$\begin{aligned} \mathbb{P}_1 &= \mathbb{P}\left[\bigcup_{k=1}^{n} \bigcup_{\substack{l=1 \\ l \neq k}}^{n} \left\{ \left| \frac{1}{l-k} \sum_{i=k+1}^{l} (X_i - \beta) \right| > \varepsilon, |k - \tau_n| \leq x, |l - \sigma_n| \leq x, \hat{\tau}_n = k, \hat{\sigma}_n = l \right\} \right] \\ &\leq \sum_{k=1}^{n} \sum_{\substack{l=1 \\ l \neq k}}^{n} \mathbb{P}\left[\left| \frac{1}{l-k} \sum_{i=k+1}^{l} (X_i - \beta) \right| > \varepsilon, |k - \tau_n| \leq x, |l - \sigma_n| \leq x, \hat{\tau}_n = k, \hat{\sigma}_n = l \right] \\ &= \sum_{\substack{k=1 \\ |k-\tau_n| \leq x}}^{n} \quad \sum_{\substack{l=1 \\ |l-\sigma_n| \leq x, l \neq k}}^{n} \mathbb{P}\left[\left| \frac{1}{l-k} \sum_{i=k+1}^{l} (X_i - \beta) \right| > \varepsilon,\ \hat{\tau}_n = k,\ \hat{\sigma}_n = l \right] . \end{aligned}$$

Without loss of generality we assume that $\mathbb{P}\left[\hat{\tau}_n = k, \hat{\sigma}_n = l \right] > 0$, since otherwise we get

$\mathbb{P}_1 = 0$. The definition of the conditional probability and Lemma 2.5 lead to

$$\mathbb{P}_1 \le \sum_{\substack{k=1 \\ |k-\tau_n|\le x}}^{n} \sum_{\substack{l=1 \\ |l-\sigma_n|\le x \\ l\neq k}}^{n} \mathbb{P}\left[\left|\frac{1}{l-k}\sum_{i=k+1}^{l}(X_i-\beta)\right| > \varepsilon \;\middle|\hat{\tau}_n = k, \hat{\sigma}_n = l\right] \mathbb{P}\left[\hat{\tau}_n = k, \hat{\sigma}_n = l\right]$$

$$\le \varepsilon^{-2} \sum_{\substack{k=1 \\ |k-\tau_n|\le x}}^{n} \sum_{\substack{l=1 \\ |l-\sigma_n|\le x, l\neq k}}^{n} (l-k)^{-2}\mathbb{E}\left[\mathbb{1}_{\{\hat{\tau}_n = k, \hat{\sigma}_n = l\}}\left(\sum_{i=k+1}^{l}(X_i-\beta)\right)^2\right].$$

Throughout the proof, we use the abbreviations

$$P(k,l) := \mathbb{P}\left[\hat{\tau}_n = k, \hat{\sigma}_n = l\right] \quad \text{and} \quad E(k,l) := \mathbb{E}\left[\left(\sum_{i=k+1}^{l}(X_i-\beta)\right)^4\right]. \tag{4.33}$$

By the Cauchy–Schwarz Inequality (see Lemma 2.11 (ii)), we infer that

$$\mathbb{P}_1 \le \varepsilon^{-2} \sum_{\substack{k=1 \\ |k-\tau_n|\le x}}^{n} \sum_{\substack{l=1 \\ |l-\sigma_n|\le x, l\neq k}}^{n} (l-k)^{-2} P(k,l)^{1/2} E(k,l)^{1/2}. \tag{4.34}$$

Lemma A.2 ensures that $\sigma_n - \tau_n \xrightarrow[n\to\infty]{} \infty$. Accordingly, there exists $n_0 = n_0(x) \in \mathbb{N}$ such that $\tau_n + x < \sigma_n - x$ for all $n \ge n_0$. From now on, let $n \ge n_0$. In (4.34) the summation indices $(k,l) \in \{1,...,n\}^2$ fulfill

$$\tau_n - x \le k \le \tau_n + x < \sigma_n - x \le l \le \sigma_n + x. \tag{4.35}$$

Set

$$I_1^- := \{k \in \mathbb{N} |\ 1 \le k \le n,\ \tau_n - x \le k \le \tau_n - 1\},$$
$$I_1^+ := \{k \in \mathbb{N} |\ 1 \le k \le n,\ \tau_n \le k \le \tau_n + x\},$$
$$I_2^- := \{l \in \mathbb{N} |\ 1 \le l \le n,\ \sigma_n - x \le l \le \sigma_n\} \qquad \text{and}$$
$$I_2^+ := \{l \in \mathbb{N} |\ 1 \le l \le n,\ \sigma_n + 1 \le l \le \sigma_n + x\}.$$

Observe that the cardinality of the sets amount

$$\left|I_1^-\right| = \left|I_2^+\right| = \lfloor x \rfloor \quad \text{and} \quad \left|I_1^+\right| = \left|I_2^-\right| = \lfloor x \rfloor + 1. \tag{4.36}$$

We split the sums in (4.34) and obtain

$$
\begin{aligned}
\mathbb{P}_1 &\le \varepsilon^{-2} \sum_{k\in I_1^-}\sum_{l\in I_2^-}(l-k)^{-2}P(k,l)^{1/2}E(k,l)^{1/2} \\
&\quad + \varepsilon^{-2} \sum_{k\in I_1^-}\sum_{l\in I_2^+}(l-k)^{-2}P(k,l)^{1/2}E(k,l)^{1/2} \\
&\quad + \varepsilon^{-2} \sum_{k\in I_1^+}\sum_{l\in I_2^-}(l-k)^{-2}P(k,l)^{1/2}E(k,l)^{1/2} \\
&\quad + \varepsilon^{-2} \sum_{k\in I_1^+}\sum_{l\in I_2^+}(l-k)^{-2}P(k,l)^{1/2}E(k,l)^{1/2} \\
&=: \sum_{i=1}^{4} D_i. \qquad (4.37)
\end{aligned}
$$

Let $C > 0$ be a generic constant. We first estimate D_3. Note that $\mathbb{E}[X_i] = \beta$ for all $i \in \{k+1, ..., l\}$ with $k \in I_1^+$ and $l \in I_2^-$. By Equation (4.33), Corollary 2.16 and $M_4 < \infty$, we have

$$
E(k,l) \le CM_4(l-k)^2 \le C(l-k)^2
$$

for all $k \in I_1^+$ and $l \in I_2^-$. Hence

$$
D_3 \le C\varepsilon^{-2} \sum_{k\in I_1^+}\sum_{l\in I_2^-}(l-k)^{-1}P(k,l)^{1/2}.
$$

The Cauchy–Schwarz Inequality (Lemma 2.11 (i)) implies

$$
D_3 \le C\varepsilon^{-2} \left(\sum_{k\in I_1^+}\sum_{l\in I_2^-}(l-k)^{-2}\right)^{1/2} \left(\sum_{k\in I_1^+}\sum_{l\in I_2^-}P(k,l)\right)^{1/2}. \qquad (4.38)
$$

Since the events $\{\hat{\tau}_n = k, \hat{\sigma}_n = l\}$ are disjoint for all $k \in I_1^+$ and $l \in I_2^-$, we can estimate by Equation (4.33)

$$
\left(\sum_{k\in I_1^+}\sum_{l\in I_2^-}P(k,l)\right)^{1/2} = \mathbb{P}\left[\bigcup_{k\in I_1^+}\bigcup_{l\in I_2^-}\{\hat{\tau}_n = k, \hat{\sigma}_n = l\}\right]^{1/2} \le 1. \qquad (4.39)
$$

For $k \in I_1^+$ we have $k \leq \tau_n + \lfloor x \rfloor$. It follows that

$$\begin{aligned} D_3 &\leq C\varepsilon^{-2} \left((\lfloor x \rfloor + 1) \sum_{l \in I_2^-} (l - (\tau_n + \lfloor x \rfloor))^{-2} \right)^{1/2} && \text{by (4.38), (4.36), (4.39)} \\ &= C\varepsilon^{-2} (\lfloor x \rfloor + 1)^{1/2} \left(\sum_{m=\sigma_n - \tau_n - 2\lfloor x \rfloor}^{\sigma_n - \tau_n - \lfloor x \rfloor} m^{-2} \right)^{1/2} && \text{by index transformation} \\ &\leq C\varepsilon^{-2} (\lfloor x \rfloor + 1)^{1/2} (\sigma_n - \tau_n - 2\lfloor x \rfloor - 1)^{-1/2} && \text{by Lem. A.5.} \end{aligned} \tag{4.40}$$

Here and subsequently, let $S_{u,v} := \sum_{i=u+1}^{v} (X_i - \beta)$ for $u, v \in \mathbb{N}$ with $u < v$. We next consider D_2. We first estimate $E(k,l) = \mathbb{E}\left[S_{k,l}^4\right]$ for fixed $k \in I_1^-$ and $l \in I_2^+$. By proper splitting of $S_{k,l}$ and the Binomial Formula, we get

$$\begin{aligned} S_{k,l}^4 &= (S_{k,\sigma_n} + S_{\sigma_n,l})^4 \\ &= S_{k,\sigma_n}^4 + 4S_{k,\sigma_n}^3 S_{\sigma_n,l} + 6S_{k,\sigma_n}^2 S_{\sigma_n,l}^2 + 4S_{k,\sigma_n} S_{\sigma_n,l}^3 + S_{\sigma_n,l}^4 \\ &\leq S_{k,\sigma_n}^4 + 4\left|S_{k,\sigma_n}\right|^3 \left|S_{\sigma_n,l}\right| + 6S_{k,\sigma_n}^2 S_{\sigma_n,l}^2 + 4\left|S_{k,\sigma_n}\right| \cdot \left|S_{\sigma_n,l}\right|^3 + S_{\sigma_n,l}^4. \end{aligned}$$

Furthermore, we have

$$\begin{aligned} S_{k,\sigma_n}^4 &= (S_{k,\tau_n} + S_{\tau_n,\sigma_n})^4 \\ &= S_{k,\tau_n}^4 + 4S_{k,\tau_n}^3 S_{\tau_n,\sigma_n} + 6S_{k,\tau_n}^2 S_{\tau_n,\sigma_n}^2 + 4S_{k,\tau_n} S_{\tau_n,\sigma_n}^3 + S_{\tau_n,\sigma_n}^4 \\ &\leq S_{k,\tau_n}^4 + 4S_{k,\tau_n}^3 S_{\tau_n,\sigma_n} + 6S_{k,\tau_n}^2 S_{\tau_n,\sigma_n}^2 + 4\left|S_{k,\tau_n}\right| \cdot \left|S_{\tau_n,\sigma_n}\right|^3 + S_{\tau_n,\sigma_n}^4. \end{aligned}$$

The independence of $X_1, ..., X_n$ ensures the independence of S_{k,σ_n} and $S_{\sigma_n,l}$ as well as S_{k,τ_n} and S_{τ_n,σ_n}. We thus get

$$\begin{aligned} E(k,l) &= \mathbb{E}\left[S_{k,l}^4\right] \\ &\leq \mathbb{E}\left[S_{k,\tau_n}^4\right] + 4\mathbb{E}\left[S_{k,\tau_n}^3\right] \mathbb{E}\left[S_{\tau_n,\sigma_n}\right] + 6\mathbb{E}\left[S_{k,\tau_n}^2\right] \mathbb{E}\left[S_{\tau_n,\sigma_n}^2\right] \\ &\quad + 4\mathbb{E}\left[\left|S_{k,\tau_n}\right|\right] \mathbb{E}\left[\left|S_{\tau_n,\sigma_n}\right|^3\right] + \mathbb{E}\left[S_{\tau_n,\sigma_n}^4\right] + 4\mathbb{E}\left[\left|S_{k,\sigma_n}\right|^3\right] \mathbb{E}\left[\left|S_{\sigma_n,l}\right|\right] \\ &\quad + 6\mathbb{E}\left[S_{k,\sigma_n}^2\right] \mathbb{E}\left[S_{\sigma_n,l}^2\right] + 4\mathbb{E}\left[\left|S_{k,\sigma_n}\right|\right] \mathbb{E}\left[\left|S_{\sigma_n,l}\right|^3\right] + \mathbb{E}\left[S_{\sigma_n,l}^4\right]. \end{aligned}$$

Since $\mathbb{E}[X_i] = \beta$ for all $i \in \{\tau_n + 1, ..., \sigma_n\}$, the absolute moments of S_{τ_n,σ_n} can be estimated by Corollary 2.16 and we have $\mathbb{E}[S_{\tau_n,\sigma_n}] = 0$. We estimate the other absolute moments by Lemma 2.17. The result is

$$\begin{aligned} E(k,l) \leq\ & CM_4(\tau_n - k)^4 + CM_2^2(\tau_n - k)^2(\sigma_n - \tau_n) + CM_1M_3(\tau_n - k)(\sigma_n - \tau_n)^{3/2} \\ & + CM_4(\sigma_n - \tau_n)^2 + CM_1M_3(\sigma_n - k)^3(l - \sigma_n) + CM_2^2(\sigma_n - k)^2(l - \sigma_n)^2 \\ & + CM_1M_3(\sigma_n - k)(l - \sigma_n)^3 + CM_4(l - \sigma_n)^4. \end{aligned} \tag{4.41}$$

From $M_4 < \infty$ we deduce that $M_p < \infty$ for $1 \leq p < 4$. By definition of D_2, (4.41) and Lemma A.7, we can assert that

$$\begin{aligned} D_2 &= \varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^+} (l-k)^{-2} P(k,l)^{1/2} E(k,l)^{1/2} \\ &\leq C\varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^+} (l-k)^{-2} (\tau_n - k)^2 P(k,l)^{1/2} \\ &\quad + C\varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^+} (l-k)^{-2} (\tau_n - k)(\sigma_n - \tau_n)^{1/2} P(k,l)^{1/2} \\ &\quad + C\varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^+} (l-k)^{-2} (\tau_n - k)^{1/2}(\sigma_n - \tau_n)^{3/4} P(k,l)^{1/2} \\ &\quad + C\varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^+} (l-k)^{-2} (\sigma_n - \tau_n) P(k,l)^{1/2} \\ &\quad + C\varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^+} (l-k)^{-2} (\sigma_n - k)^{3/2}(l - \sigma_n)^{1/2} P(k,l)^{1/2} \\ &\quad + C\varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^+} (l-k)^{-2} (\sigma_n - k)(l - \sigma_n) P(k,l)^{1/2} \\ &\quad + C\varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^+} (l-k)^{-2} (\sigma_n - k)^{1/2}(l - \sigma_n)^{3/2} P(k,l)^{1/2} \\ &\quad + C\varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^+} (l-k)^{-2} (l - \sigma_n)^2 P(k,l)^{1/2} \\ &\quad =: \sum_{i=1}^{8} D_{2,i}. \end{aligned} \tag{4.42}$$

The further proceeding is presented for $D_{2,3}$ and $D_{2,5}$. The Cauchy–Schwarz Inequality (Lemma 2.11 (i)) and similar arguments used in (4.39) lead to

$$\begin{aligned} D_{2,3} &= C\varepsilon^{-2} \sum_{k\in I_1^-}\sum_{l\in I_2^+} (l-k)^{-2}(\tau_n-k)^{1/2}(\sigma_n-\tau_n)^{3/4}P(k,l)^{1/2} \\ &\le C\varepsilon^{-2}\left(\sum_{k\in I_1^-}\sum_{l\in I_2^+}(l-k)^{-4}(\tau_n-k)(\sigma_n-\tau_n)^{3/2}\right)^{1/2}\left(\sum_{k\in I_1^-}\sum_{l\in I_2^+}P(k,l)\right)^{1/2} \\ &\le C\varepsilon^{-2}\left(\sum_{k\in I_1^-}\sum_{l\in I_2^+}(l-k)^{-4}(\tau_n-k)(\sigma_n-\tau_n)^{3/2}\right)^{1/2}. \end{aligned} \tag{4.43}$$

Since $\tau_n - k \le \lfloor x\rfloor$ and $l > \sigma_n$ for $k \in I_1^-$ and $l \in I_2^+$, we see by (4.36) and an index transformation that

$$\begin{aligned} D_{2,3} &\le C\varepsilon^{-2}\left(\lfloor x\rfloor(\sigma_n-\tau_n)^{3/2}\sum_{l\in I_2^+}\sum_{k\in I_1^-}(\sigma_n-k)^{-4}\right)^{1/2} \\ &= C\varepsilon^{-2}\left(\lfloor x\rfloor^2(\sigma_n-\tau_n)^{3/2}\sum_{m=\sigma_n-\tau_n+1}^{\sigma_n-\tau_n+\lfloor x\rfloor} m^{-4}\right)^{1/2}. \end{aligned}$$

By Lemma A.5, we obtain

$$D_{2,3} \le C\varepsilon^{-2}\left(\lfloor x\rfloor^2(\sigma_n-\tau_n)^{3/2}(\sigma_n-\tau_n)^{-3}\right)^{1/2} = C\varepsilon^{-2}\lfloor x\rfloor(\sigma_n-\tau_n)^{-3/4}.$$

We next consider $D_{2,5}$. As in (4.43), we get

$$\begin{aligned} D_{2,5} &= C\varepsilon^{-2}\sum_{k\in I_1^-}\sum_{l\in I_2^+}(l-k)^{-2}(\sigma_n-k)^{3/2}(l-\sigma_n)^{1/2}P(k,l)^{1/2} \\ &\le C\varepsilon^{-2}\left(\sum_{k\in I_1^-}\sum_{l\in I_2^+}(l-k)^{-4}(\sigma_n-k)^3(l-\sigma_n)\right)^{1/2}. \end{aligned}$$

Since $l-\sigma_n \le \lfloor x\rfloor$ and $l > \sigma_n$ for $l \in I_2^+$, we deduce that

$$D_{2,5} \le C\varepsilon^{-2}\left(\lfloor x\rfloor\sum_{k\in I_1^-}\sum_{l\in I_2^+}(\sigma_n-k)^{-1}\right)^{1/2}.$$

For $k \in I_1^-$ we have $k < \tau_n$. By (4.36), we infer that

$$\begin{aligned} D_{2,5} &\leq C\varepsilon^{-2} \left(\lfloor x \rfloor \sum_{k \in I_1^-} \sum_{l \in I_2^+} (\sigma_n - \tau_n)^{-1} \right)^{1/2} \\ &= C\varepsilon^{-2} \lfloor x \rfloor^{3/2} (\sigma_n - \tau_n)^{-1/2}. \end{aligned}$$

A similar proceeding leads to the estimate of the remaining terms. We get

$$\begin{aligned} D_2 &\leq \sum_{i=1}^{8} D_{2,i} \\ &\leq C\varepsilon^{-2} \lfloor x \rfloor^{5/2} (\sigma_n - \tau_n)^{-3/2} + C\varepsilon^{-2} \lfloor x \rfloor^{3/2} (\sigma_n - \tau_n)^{-1} \\ &\quad + C\varepsilon^{-2} \lfloor x \rfloor (\sigma_n - \tau_n)^{-3/4} + C\varepsilon^{-2} \lfloor x \rfloor^{1/2} (\sigma_n - \tau_n)^{-1/2} \\ &\quad + C\varepsilon^{-2} \lfloor x \rfloor^{3/2} (\sigma_n - \tau_n)^{-1/2} + C\varepsilon^{-2} \lfloor x \rfloor^{3/2} (\sigma_n - \tau_n)^{-1/2} \\ &\quad + C\varepsilon^{-2} \lfloor x \rfloor^{2} (\sigma_n - \tau_n)^{-1} + C\varepsilon^{-2} \lfloor x \rfloor^{5/2} (\sigma_n - \tau_n)^{-3/2}. \end{aligned}$$

The estimate of D_1 and D_4 runs as before. By (4.37), we obtain

$$\begin{aligned} D_1 &= \varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^-} (l-k)^{-2} P(k,l)^{1/2} E(k,l)^{1/2} \\ &\leq C\varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^-} (l-k)^{-2} (\tau_n - k)^2 P(k,l)^{1/2} \\ &\quad + C\varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^-} (l-k)^{-2} (\tau_n - k)(l - \tau_n)^{1/2} P(k,l)^{1/2} \\ &\quad + C\varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^-} (l-k)^{-2} (\tau_n - k)^{1/2} (l - \tau_n)^{3/4} P(k,l)^{1/2} \\ &\quad + C\varepsilon^{-2} \sum_{k \in I_1^-} \sum_{l \in I_2^-} (l-k)^{-2} (l - \tau_n) P(k,l)^{1/2} \\ &\leq C\varepsilon^{-2} \lfloor x \rfloor^{5/2} (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-3/2} + C\varepsilon^{-2} \lfloor x \rfloor^{3/2} (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-1} \\ &\quad + C\varepsilon^{-2} \lfloor x \rfloor (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-3/4} + C\varepsilon^{-2} \lfloor x \rfloor^{1/2} (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-1/2}. \end{aligned}$$

and

$$\begin{aligned}
D_4 &= \varepsilon^{-2} \sum_{k\in I_1^+} \sum_{l\in I_2^+} (l-k)^{-2} P(k,l)^{1/2} E(k,l)^{1/2} \\
&\le C\varepsilon^{-2} \sum_{k\in I_1^+} \sum_{l\in I_2^+} (l-k)^{-2} (\sigma_n - k) P(k,l)^{1/2} \\
&\quad + C\varepsilon^{-2} \sum_{k\in I_1^+} \sum_{l\in I_2^+} (l-k)^{-2} (\sigma_n - k)^{3/4} (l-\sigma_n)^{1/2} P(k,l)^{1/2} \\
&\quad + C\varepsilon^{-2} \sum_{k\in I_1^+} \sum_{l\in I_2^+} (l-k)^{-2} (\sigma_n - k)^{1/2} (l-\sigma_n) P(k,l)^{1/2} \\
&\quad + C\varepsilon^{-2} \sum_{k\in I_1^+} \sum_{l\in I_2^+} (l-k)^{-2} (l-\sigma_n)^2 P(k,l)^{1/2} \\
&\le C\varepsilon^{-2} \lfloor x \rfloor^{1/2} (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-1/2} + C\varepsilon^{-2} \lfloor x \rfloor (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-3/4} \\
&\quad + C\varepsilon^{-2} \lfloor x \rfloor^{3/2} (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-1} + C\varepsilon^{-2} \lfloor x \rfloor^{5/2} (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-3/2}.
\end{aligned}$$

In summary, by (4.37), (4.40) and the estimates above, there exist $n_0 = n_0(x) \in \mathbb{N}$ and a constant $C > 0$ such that

$$\begin{aligned}
\mathbb{P}_1 &\le \sum_{i=1}^{4} D_4 \\
&\le C\varepsilon^{-2} \Big(\lfloor x \rfloor^{5/2} (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-3/2} + \lfloor x \rfloor^{3/2} (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-1} \\
&\quad + \lfloor x \rfloor (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-3/4} + \lfloor x \rfloor^{1/2} (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-1/2} \\
&\quad + \lfloor x \rfloor^{5/2} (\sigma_n - \tau_n)^{-3/2} + \lfloor x \rfloor^{3/2} (\sigma_n - \tau_n)^{-1} + \lfloor x \rfloor (\sigma_n - \tau_n)^{-3/4} \\
&\quad + \lfloor x \rfloor^{1/2} (\sigma_n - \tau_n)^{-1/2} + \lfloor x \rfloor^{3/2} (\sigma_n - \tau_n)^{-1/2} + \lfloor x \rfloor^{3/2} (\sigma_n - \tau_n)^{-1/2} \\
&\quad + \lfloor x \rfloor^{2} (\sigma_n - \tau_n)^{-1} + \lfloor x \rfloor^{5/2} (\sigma_n - \tau_n)^{-3/2} + (\lfloor x \rfloor + 1)^{1/2} (\sigma_n - \tau_n - 2\lfloor x \rfloor - 1)^{-1/2} \\
&\quad + \lfloor x \rfloor^{1/2} (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-1/2} + \lfloor x \rfloor (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-3/4} \\
&\quad + \lfloor x \rfloor^{3/2} (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-1} + \lfloor x \rfloor^{5/2} (\sigma_n - \tau_n - \lfloor x \rfloor - 1)^{-3/2} \Big) \qquad (4.44)
\end{aligned}$$

for all $n \ge n_0$, $\varepsilon > 0$ and $x \ge 1$. To see convergence in probability of $\hat{\beta}_n$ to β, we now apply the stochastic boundedness of $\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$. By (4.32) and the maximum norm, we infer that

$$\mathbb{P}_2 \leq \mathbb{P}\left[\|\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| > x\right] \qquad \text{and} \qquad \mathbb{P}_3 \leq \mathbb{P}\left[\|\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| > x\right]$$

for all $x > 0$. By (4.32), we conclude that

$$\begin{aligned} \limsup_{n\to\infty} \mathbb{P}\left[\left|\hat{\beta}_n - \beta\right| > \varepsilon\right] &\leq \limsup_{n\to\infty} \left(\mathbb{P}_1 + \mathbb{P}_2 + \mathbb{P}_3\right) \\ &\leq \limsup_{n\to\infty} \mathbb{P}_1 + 2 \limsup_{n\to\infty} \mathbb{P}\left[\|\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n\| > x\right] \qquad (4.45) \end{aligned}$$

for all $\varepsilon > 0$ and $x \geq 1$. Since $\sigma_n - \tau_n \xrightarrow[n\to\infty]{} \infty$ by Lemma A.2, we deduce by (4.44) that

$$\limsup_{n\to\infty} \mathbb{P}_1 = 0$$

for all $\varepsilon > 0$ and $x \geq 1$. By (4.45) and Proposition 4.17, letting $x \to \infty$ gives

$$\lim_{n\to\infty} \mathbb{P}\left[\left|\hat{\beta}_n - \beta\right| > \varepsilon\right] = 0,$$

which means

$$\hat{\beta}_n \xrightarrow[n\to\infty]{\mathbb{P}} \beta.$$

In the same manner we can see that

$$\hat{\alpha}_n \xrightarrow[n\to\infty]{\mathbb{P}} \alpha \quad \text{and} \quad \hat{\gamma}_n \xrightarrow[n\to\infty]{\mathbb{P}} \gamma. \qquad \square$$

Corollary 4.19. *If $M_4 < \infty$, then*

$$\left(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\alpha}}_n\right) \xrightarrow[n\to\infty]{\mathbb{P}} (\boldsymbol{\theta}, \boldsymbol{\alpha}).$$

Proof. The claim follows from Theorems 4.10 and 4.18 and the properties of convergence in probability. $\square$

4.2 Another estimation approach for the multiple change-point

In the previous section we have seen the point estimation of $\boldsymbol{\theta}$ and $\boldsymbol{\tau}_n$, respectively. Another aim is to estimate $\boldsymbol{\tau}_n$ by an asymptotic confidence region (as specified in Section 3.4 in the case of known expectations). Unfortunately, the rescaled process with respect to $\hat{M}_n$ is hard to handle to examine convergence in distribution of $\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$ (compare Section 3.3 in the case of known expectations). Therefore, in this section we construct another estimator of $\boldsymbol{\tau}_n$ based on the consistent estimator $\hat{\boldsymbol{\alpha}}_n$ of expectations, which allows us to proceed similarly to Chapter 3.

4.2.1 Estimation of the multiple change-point

We begin with the estimation of the moments of change $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ again. Now, the main idea is to replace the unknown expectations $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)$ in the criterion function $\bar{S}_n$, defined in (3.1), by their associated estimators $\hat{\boldsymbol{\alpha}}_n = \left(\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n\right)$. Let us denote by S_n^* the random criterion function

$$S_n^*(k,l) := \sum_{i=1}^{k} (X_i - \hat{\alpha}_n)^2 + \sum_{i=k+1}^{l} \left(X_i - \hat{\beta}_n\right)^2 + \sum_{i=l+1}^{n} (X_i - \hat{\gamma}_n)^2 \,, \; (k,l) \in \Delta_n. \quad (4.46)$$

Note that S_n^* features the same structure as $\bar{S}_n$. Consequently, our further approach is very similar to Chapter 3. It is evident that S_n^* has at least one minimizer. Similarly to Equation (3.3), we use a choice function $\phi^* : \operatorname{Argmin}(S_n^*) \longrightarrow \Delta_n$ if more than one minimizing point of S_n^* exists. Here and subsequently,

$$\boldsymbol{\tau}_n^* := (\tau_n^*, \sigma_n^*) := \operatorname*{argmin}_{(k,l)\in\Delta_n} S_n^*(k,l) \quad (4.47)$$

stands for $\boldsymbol{\tau}_n^* = \phi^* \left(\operatorname{Argmin}(S_n^*)\right)$.

We first observe that $(\boldsymbol{\tau}_n^*, \hat{\boldsymbol{\alpha}}_n)$ is also a least squares estimator of $(\boldsymbol{\tau}_n, \boldsymbol{\alpha})$.

Lemma 4.20. *Let $n \in \mathbb{N}$. Then*

$$(\boldsymbol{\tau}_n^*, \hat{\boldsymbol{\alpha}}_n) \in \operatorname{Argmin}(S_n).$$

Proof. Fix $n \in \mathbb{N}$. By Equations (4.4) and (4.47), we have $\hat{\boldsymbol{\tau}}_n \in \Delta_n$ and $\boldsymbol{\tau}_n^* \in \Delta_n$. Equations (4.1), (4.46) and (4.47) lead to

$$S_n\left(\boldsymbol{\tau}_n^*, \hat{\boldsymbol{\alpha}}_n\right) = S_n^*\left(\boldsymbol{\tau}_n^*\right) = \min_{(k,l)\in\Delta_n} S_n^*(k,l) \leq S_n^*\left(\hat{\boldsymbol{\tau}}_n\right).$$

In addition, Equations (4.46) and (4.1) and (4.6) yield

$$S_n^*\left(\hat{\boldsymbol{\tau}}_n\right) = S_n\left(\hat{\boldsymbol{\tau}}_n, \hat{\boldsymbol{\alpha}}_n\right) = \min_{(k,l,a_1,a_2,a_3)\in\Delta_n\times\mathbb{R}^3} S_n(k,l,a_1,a_2,a_3) \leq S_n\left(\boldsymbol{\tau}_n^*, \hat{\boldsymbol{\alpha}}_n\right).$$

Especially, we obtain

$$S_n\left(\boldsymbol{\tau}_n^*, \hat{\boldsymbol{\alpha}}_n\right) = \min_{(k,l,a_1,a_2,a_3)\in\Delta_n\times\mathbb{R}^3} S_n(k,l,a_1,a_2,a_3),$$

which is our claim. □

Adapted from Chapter 3, for simplicity of notation, we apply the abbreviations

$$a_{n,1}^*(X_i) := 2\left(\hat{\beta}_n - \hat{\alpha}_n\right) X_i + \hat{\alpha}_n^2 - \hat{\beta}_n^2 \qquad \text{and} \tag{4.48}$$
$$a_{n,2}^*(X_i) := 2\left(\hat{\gamma}_n - \hat{\beta}_n\right) X_i + \hat{\beta}_n^2 - \hat{\gamma}_n^2$$

for $i \in \{1, ..., n\}$ and write

$$M_n^*(k,l) := \sum_{i=1}^{k} a_{n,1}^*(X_i) + \sum_{i=1}^{l} a_{n,2}^*(X_i), \qquad (k,l) \in \Delta_n. \tag{4.49}$$

Lemma 4.21. *Let $n \in \mathbb{N}$. Then*

$$\operatorname{Argmin}\left(S_n^*\right) = \operatorname{Argmin}\left(M_n^*\right).$$

Proof. The proof of Lemma 3.1 works for S_n^*, M_n^*, $a_{n,1}^*$, $a_{n,2}^*$, $\hat{\alpha}_n$, $\hat{\beta}_n$ and $\hat{\gamma}_n$ instead of $\bar{S}_n$, $\bar{M}_n$, a_1, a_2, α, β and γ. □

According to the previous lemma, we use the representation

$$\boldsymbol{\tau}_n^* = (\tau_n^*, \sigma_n^*) = \underset{(k,l)\in\Delta_n}{\operatorname{argmin}} M_n^*(k,l). \tag{4.50}$$

to estimate $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$.
A simulation study (see Chapter 5 for more details) provides the following conjecture. If all distributions are continuous, then the estimators $\boldsymbol{\tau}_n^*$ and $\hat{\boldsymbol{\tau}}_n$ are almost surely identical.

Conjecture 4.22. Let $n \in \mathbb{N}$ and let Q_1, Q_2, Q_3 be continuous distributions. Then

$$\operatorname{Argmin}(M_n^*) = \{\hat{\boldsymbol{\tau}}_n\} \quad \text{almost surely,}$$

which means $\boldsymbol{\tau}_n^* = \hat{\boldsymbol{\tau}}_n$ almost surely. Furthermore, it holds $\mathbb{P}[\boldsymbol{\tau}_n^* \neq \hat{\boldsymbol{\tau}}_n] > 0$.

The further approach to estimate the multiple change-point $\boldsymbol{\theta} = (\theta_1, \theta_2)$ is analogous to the last part of Section 3.1. The estimator of the multiple change-point is given by

$$\boldsymbol{\theta}_n^* := \frac{1}{n}\boldsymbol{\tau}_n^*.$$

Moreover, we define

$$\rho_n^*(s,t) := \frac{1}{n} M_n^*(\lfloor ns \rfloor, \lfloor nt \rfloor), \quad (s,t) \in \Theta_n,$$

where Θ_n is given by (3.6).

Lemma 4.23. *Let $n \in \mathbb{N}$. Then*

$$\boldsymbol{\theta}_n^* = \underset{(s,t)\in\Theta_n}{\operatorname{argmin}} \rho_n^*(s,t).$$

Proof. The proof of Lemma 3.4 works by replacing (3.3), $\bar{\boldsymbol{\tau}}_n = (\bar{\tau}_n, \bar{\sigma}_n)$, $\bar{\boldsymbol{\theta}}_n$, $\bar{\rho}_n$ and $\bar{M}_n$ by (4.50), $\boldsymbol{\tau}_n^* = (\tau_n^*, \sigma_n^*)$, $\boldsymbol{\theta}_n^*$, ρ_n^* and M_n^*. □

Remark 4.24. The factor n^{-1} in the definition of ρ_n^* does not influence the minimizing points of M_n^*, but the proof of consistency of $\boldsymbol{\theta}_n^*$ requires this factor.

Lemma 4.25. *ρ_n^*, $n \in \mathbb{N}$, is a stochastic process with trajectories in the multivariate Skorokhod space $D(\Theta_n)$.*

Proof. The proof of Lemma 3.6 remains valid for ρ_n^* and M_n^* instead of $\bar{\rho}_n$ and $\bar{M}_n$. □

4.2.2 Consistency of the multiple change-point estimator

This section deals with the weak conistency of $\boldsymbol{\theta}_n^*$. For this purpose, we apply Theorem 2.1 again. To get uniform convergence in probability of ρ_n^* (assumption (i) of Theorem 2.1), we give a decomposition of ρ_n^* first.

Lemma 4.26. *Let $n \in \mathbb{N}$ and $(s,t) \in \Theta_n$. Then*

$$\rho_n^*(s,t) = \delta_n^*(s,t) + \varrho_n^*(s,t),$$

where δ_n^* and ϱ_n^* are specified in Lemma B.2.

Proof. Fix $n \in \mathbb{N}$. We first recall that

$$\mathbb{E}[X_i] = \begin{cases} \alpha, & 1 \le i \le \tau_n, \\ \beta, & \tau_n + 1 \le i \le \sigma_n, \\ \gamma, & \sigma_n + 1 \le i \le n. \end{cases} \tag{4.51}$$

Definitions of ρ_n^* and M_n^* yield

$$\rho_n^*(s,t) = \frac{1}{n} M_n^*(\lfloor ns \rfloor, \lfloor nt \rfloor) = \frac{1}{n}\left(\sum_{i=1}^{\lfloor ns \rfloor} a_{n,1}^*(X_i) + \sum_{i=1}^{\lfloor nt \rfloor} a_{n,2}^*(X_i) \right)$$

for all $(s,t) \in \Theta_n$. We only discuss the case $(s,t) \in \Theta^2 \cap \Theta_n$. Lemma A.1 (ii) gives

$$1 \leq \lfloor ns \rfloor \leq \tau_n < \lfloor nt \rfloor \leq \sigma_n < n.$$

We split the sums into segments according to above and obtain

$$\rho_n^*(s,t) = \frac{1}{n} \left(\sum_{i=1}^{\lfloor ns \rfloor} a_{n,1}^*(X_i) + \sum_{i=1}^{\tau_n} a_{n,2}^*(X_i) + \sum_{i=\tau_n+1}^{\lfloor nt \rfloor} a_{n,2}^*(X_i) \right).$$

We now use the expectations to center X_i, i.e., $X_i = (X_i - \mathbb{E}[X_i]) + \mathbb{E}[X_i]$, $i \in \{1, ..., \lfloor nt \rfloor\}$. By definitions of $a_{n,1}^*$ and $a_{n,2}^*$, an easy computation shows that $\rho_n^*(s,t) = \delta_n^*(s,t) + \varrho_n^*(s,t)$, where

$$\begin{aligned} \delta_n^*(s,t) := & \frac{2}{n}\left(\hat{\beta}_n - \hat{\alpha}_n\right) \sum_{i=1}^{\lfloor ns \rfloor}(X_i - \alpha) + \frac{2}{n}\left(\hat{\gamma}_n - \hat{\beta}_n\right) \sum_{i=1}^{\tau_n}(X_i - \alpha) \\ & + \frac{2}{n}\left(\hat{\gamma}_n - \hat{\beta}_n\right) \sum_{i=\tau_n+1}^{\lfloor nt \rfloor}(X_i - \beta) \end{aligned} \tag{4.52}$$

and

$$\begin{aligned} \varrho_n^*(s,t) := & \left(2\alpha\left(\hat{\beta}_n - \hat{\alpha}_n\right) + \hat{\alpha}_n^2 - \hat{\beta}_n^2\right) \frac{\lfloor ns \rfloor}{n} + 2(\alpha - \beta)\left(\hat{\gamma}_n - \hat{\beta}_n\right) \frac{\tau_n}{n} \\ & + \left(2\beta\left(\hat{\gamma}_n - \hat{\beta}_n\right) + \hat{\beta}_n^2 - \hat{\gamma}_n^2\right) \frac{\lfloor nt \rfloor}{n}. \end{aligned} \tag{4.53}$$

The details and the other cases are left to the reader. □

We prove in the following both lemmas that δ_n^* uniformly converges in probability to zero and ϱ_n^* to the limit process ρ, given in (3.13).

Lemma 4.27. *If $M_4 < \infty$, then*

$$\sup_{(s,t) \in \Theta_n} |\delta_n^*(s,t)| \xrightarrow[n \to \infty]{\mathbb{P}} 0.$$

Proof. Fix $n \in \mathbb{N}$. Our proof starts with the observation that the partition of Θ gives

$$\sup_{(s,t) \in \Theta_n} |\delta_n^*(s,t)| = \max_{i \in \{1,...,6\}} \sup_{(s,t) \in \Theta^i \cap \Theta_n} |\delta_n^*(s,t)|. \tag{4.54}$$

We look at case $(s,t) \in \Theta^2 \cap \Theta_n$, which leads to $1 \le \lfloor ns \rfloor \le \tau_n < \lfloor nt \rfloor \le \sigma_n < n$. Set

$$\delta^*_{n,1}(s) := \frac{1}{n}\sum_{i=1}^{\lfloor ns \rfloor}(X_i - \alpha), \ \delta^*_{n,2} := \frac{1}{n}\sum_{i=1}^{\tau_n}(X_i - \alpha) \text{ and } \delta^*_{n,3}(t) := \frac{1}{n}\sum_{i=\tau_n+1}^{\lfloor nt \rfloor}(X_i - \beta).$$

By (4.52) and the Triangle Inequality, we obtain

$$\sup_{(s,t)\in\Theta^2\cap\Theta_n} |\delta^*_n(s,t)| \le 2\left|\hat{\beta}_n - \hat{\alpha}_n\right| \sup_{(s,t)\in\Theta^2\cap\Theta_n} |\delta^*_{n,1}(s)| + 2\left|\hat{\gamma}_n - \hat{\beta}_n\right| \cdot \left|\delta^*_{n,2}\right| + 2\left|\hat{\gamma}_n - \hat{\beta}_n\right| \sup_{(s,t)\in\Theta^2\cap\Theta_n} |\delta^*_{n,3}(t)|. \tag{4.55}$$

By the weak consistency of $\hat{\boldsymbol{\alpha}}_n = \left(\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n\right)$ (see Theorem 4.18) and the properties of convergence in probability, it is sufficient to show that

$$\sup_{(s,t)\in\Theta^2\cap\Theta_n} |\delta^*_{n,1}(s)| \xrightarrow[n\to\infty]{\mathbb{P}} 0, \quad |\delta^*_{n,2}| \xrightarrow[n\to\infty]{\mathbb{P}} 0 \quad \text{and} \quad \sup_{(s,t)\in\Theta^2\cap\Theta_n} |\delta^*_{n,3}(t)| \xrightarrow[n\to\infty]{\mathbb{P}} 0. \tag{4.56}$$

To apply the first Kolmogorov Inequality (Lemma 2.7) and the Chebyshev Inequality (Lemma 2.6), we observe that we have sums of independent and centered random variables. We conclude for all $\varepsilon > 0$ that

$$\begin{aligned}
&\mathbb{P}\left[\sup_{(s,t)\in\Theta^2\cap\Theta_n} |\delta^*_{n,1}(s)| > \varepsilon\right] \\
&\quad \le \mathbb{P}\left[\max_{1\le k\le\tau_n}\left|\sum_{i=1}^{k}(X_i-\alpha)\right| > n\varepsilon\right] \\
&\quad \le \varepsilon^{-2}n^{-2}\sum_{i=1}^{\tau_n}\mathbb{V}\left[X_i-\alpha\right] && \text{by first Kolmogorov In.} \\
&\quad \le M_2\varepsilon^{-2}n^{-2}\tau_n && \text{by } \mathbb{V}[X_i] \le M_2 \\
&\quad \le M_2\varepsilon^{-2}n^{-1} && \text{by } \tau_n \le n.
\end{aligned} \tag{4.57}$$

Furthermore, we get for all $\varepsilon > 0$

$$\begin{aligned}
\mathbb{P}\left[|\delta^*_{n,2}| > \varepsilon\right] &= \mathbb{P}\left[\left|\sum_{i=1}^{\tau_n}(X_i - \alpha)\right| > n\varepsilon\right] \\
&\leq \varepsilon^{-2} n^{-2} \sum_{i=1}^{\tau_n} \mathbb{V}\left[X_i - \alpha\right] && \text{by Chebyshev In.} \\
&\leq M_2 \varepsilon^{-2} n^{-2} \tau_n && \text{by } \mathbb{V}[X_i] \leq M_2 \\
&\leq M_2 \varepsilon^{-2} n^{-1} && \text{by } \tau_n \leq n. \qquad (4.58)
\end{aligned}$$

Moreover, by an index transformation, we see for all $\varepsilon > 0$ that

$$\begin{aligned}
&\mathbb{P}\left[\sup_{(s,t)\in\Theta^2\cap\Theta_n} |\delta^*_{n,3}(t)| > \varepsilon\right] \qquad (4.59) \\
&\leq \mathbb{P}\left[\max_{\tau_n+1\leq l\leq\sigma_n}\left|\sum_{i=\tau_n+1}^{l}(X_i - \beta)\right| > n\varepsilon\right] \\
&= \mathbb{P}\left[\max_{1\leq l-\tau_n\leq\sigma_n-\tau_n}\left|\sum_{i=1}^{l-\tau_n}(X_{\tau_n+i} - \beta)\right| > n\varepsilon\right] \\
&\leq \varepsilon^{-2} n^{-2} \sum_{i=1}^{\sigma_n-\tau_n} \mathbb{V}\left[X_i - \beta\right] && \text{by first Kolmogorov In.} \\
&\leq M_2 \varepsilon^{-2} n^{-2}(\sigma_n - \tau_n) && \text{by } \mathbb{V}[X_i] \leq M_2 \\
&\leq M_2 \varepsilon^{-2} n^{-1} && \text{by } \sigma_n - \tau_n \leq n. \qquad (4.60)
\end{aligned}$$

From $M_4 < \infty$ we see that $M_2 < \infty$. To deduce (4.56) from (4.57), (4.58) and (4.59), let $n \to \infty$. The rest of the proof runs as before. We find that

$$\sup_{(s,t)\in\Theta^i\cap\Theta_n} |\delta^*_n(s,t)| \xrightarrow[n\to\infty]{\mathbb{P}} 0$$

for all $i \in \{1, 3, 4, 5, 6\}$. The assertion follows by (4.54). □

Lemma 4.28. *If $M_4 < \infty$, then*

$$\sup_{(s,t)\in\Theta_n} |\varrho^*_n(s,t) - \rho(s,t)| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Proof. By the partition of Θ (see (3.11)), we have

$$\sup_{(s,t)\in\Theta_n} |\varrho_n^*(s,t)-\rho(s,t)| = \max_{i\in\{1,\dots,6\}} \sup_{(s,t)\in\Theta^i\cap\Theta_n} |\varrho_n^*(s,t)-\rho(s,t)|. \tag{4.61}$$

We consider the case $(s,t)\in\Theta^2\cap\Theta_n$ again. As defined in (3.13), we see that

$$\rho(s,t) = -s(\alpha-\beta)^2 - t(\beta-\gamma)^2 + \theta_1\left((\alpha-\beta)^2+(\beta-\gamma)^2-(\alpha-\gamma)^2\right).$$

Equation (4.53) and the Triangle Inequality lead to

$$\begin{aligned}
&\sup_{(s,t)\in\Theta^2\cap\Theta_n} |\varrho_n^*(s,t)-\rho(s,t)| \\
&\leq \sup_{(s,t)\in\Theta^2\cap\Theta_n} \left|\frac{\lfloor ns\rfloor}{n}\left(2\alpha\left(\hat{\beta}_n-\hat{\alpha}_n\right)+\hat{\alpha}_n^2-\hat{\beta}_n^2\right) - s\left(-(\alpha-\beta)^2\right)\right| \\
&\quad + \sup_{(s,t)\in\Theta^2\cap\Theta_n} \left|\frac{\lfloor nt\rfloor}{n}\left(2\beta\left(\hat{\gamma}_n-\hat{\beta}_n\right)+\hat{\beta}_n^2-\hat{\gamma}_n^2\right) - t\left(-(\beta-\gamma)^2\right)\right| \\
&\quad + \left|2\frac{\tau_n}{n}(\alpha-\beta)\left(\hat{\gamma}_n-\hat{\beta}_n\right) - \theta_1\left((\alpha-\beta)^2+(\beta-\gamma)^2-(\alpha-\gamma)^2\right)\right|.
\end{aligned}$$

Note that it holds $AB-ab = B(A-a)+(B-b)a$ for all $A,B,a,b\in\mathbb{R}$. Therefore, the Triangle Inequality gives

$$\begin{aligned}
&\sup_{(s,t)\in\Theta^2\cap\Theta_n} |\varrho_n^*(s,t)-\rho(s,t)| \\
&\leq \left|2\alpha\left(\hat{\beta}_n-\hat{\alpha}_n\right)+\hat{\alpha}_n^2-\hat{\beta}_n^2\right| \sup_{(s,t)\in\Theta^2\cap\Theta_n}\left|\frac{\lfloor ns\rfloor}{n}-s\right| \\
&\qquad + \left|2\alpha\left(\hat{\beta}_n-\hat{\alpha}_n\right)+\hat{\alpha}_n^2-\hat{\beta}_n^2+(\alpha-\beta)^2\right| \sup_{(s,t)\in\Theta^2\cap\Theta_n}|s| \\
&\quad + \left|2\beta\left(\hat{\gamma}_n-\hat{\beta}_n\right)+\hat{\beta}_n^2-\hat{\gamma}_n^2\right| \sup_{(s,t)\in\Theta^2\cap\Theta_n}\left|\frac{\lfloor nt\rfloor}{n}-t\right| \\
&\qquad + \left|2\beta\left(\hat{\gamma}_n-\hat{\beta}_n\right)+\hat{\beta}_n^2-\hat{\gamma}_n^2+(\beta-\gamma)^2\right| \sup_{(s,t)\in\Theta^2\cap\Theta_n}|t| \\
&\quad + \left|2(\alpha-\beta)\left(\hat{\gamma}_n-\hat{\beta}_n\right)\right|\cdot\left|\frac{\tau_n}{n}-\theta_1\right| \\
&\qquad + \left|2(\alpha-\beta)\left(\hat{\gamma}_n-\hat{\beta}_n\right) - \left((\alpha-\beta)^2+(\beta-\gamma)^2-(\alpha-\gamma)^2\right)\right|\theta_1.
\end{aligned}$$

We further estimate

$$\sup_{(s,t)\in\Theta^2\cap\Theta_n}|s|\leq 1 \qquad \text{and} \qquad \sup_{(s,t)\in\Theta^2\cap\Theta_n}|t|\leq 1.$$

Lemma A.1 (iii) and (iv) yield

$$\begin{aligned}
&\sup_{(s,t)\in\Theta^2\cap\Theta_n} |\varrho_n^*(s,t)-\rho(s,t)| \\
&\leq \left|2\alpha\left(\hat{\beta}_n-\hat{\alpha}_n\right)+\hat{\alpha}_n^2-\hat{\beta}_n^2\right|\frac{1}{n}+\left|2\alpha\left(\hat{\beta}_n-\hat{\alpha}_n\right)+\hat{\alpha}_n^2-\hat{\beta}_n^2+(\alpha-\beta)^2\right| \\
&\quad+\left|2\beta\left(\hat{\gamma}_n-\hat{\beta}_n\right)+\hat{\beta}_n^2-\hat{\gamma}_n^2\right|\frac{1}{n}+\left|2\beta\left(\hat{\gamma}_n-\hat{\beta}_n\right)+\hat{\beta}_n^2-\hat{\gamma}_n^2+(\beta-\gamma)^2\right| \\
&\quad+\left|2(\alpha-\beta)\left(\hat{\gamma}_n-\hat{\beta}_n\right)\right|\frac{1}{n} \\
&\quad+\left|2(\alpha-\beta)\left(\hat{\gamma}_n-\hat{\beta}_n\right)-\left((\alpha-\beta)^2+(\beta-\gamma)^2-(\alpha-\gamma)^2\right)\right|\theta_1.
\end{aligned}$$

By the consistency of $\hat{\boldsymbol{\alpha}}_n=\left(\hat{\alpha}_n,\hat{\beta}_n,\hat{\gamma}_n\right)$ (see Theorem 4.18) and the properties of convergence in probability, a trivial verification shows that

$$\sup_{(s,t)\in\Theta^2\cap\Theta_n} |\varrho_n^*(s,t)-\rho(s,t)| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

In the same manner we can see that

$$\sup_{(s,t)\in\Theta^i\cap\Theta_n} |\varrho_n^*(s,t)-\rho(s,t)| \xrightarrow[n\to\infty]{\mathbb{P}} 0$$

for all $i\in\{1,3,4,5,6\}$. Combining this with (4.61) finishes the proof. □

We can now state and prove weak consistency of $\boldsymbol{\theta}_n^*$, which is one of our main results.

Theorem 4.29. *If $M_4<\infty$, then*

$$\boldsymbol{\theta}_n^* \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\theta}.$$

Proof. We apply Theorem 2.1. ρ_n^*, $n\in\mathbb{N}$, is a stochastic process with trajectories in the multivariate Skorokhod space $D(\Theta_n)$ by Lemma 4.25. ρ has trajectories in the multivariate Skorokhod space $D(\Theta)$, since ρ is continuous, as is easy to check. Moreover, $(\Theta_n)_{n\in\mathbb{N}}\subseteq\Theta$ is a sequence of sets such that $\Theta_n\subseteq\Theta_{n+1}$ for every $n\in\mathbb{N}$ with $\bigcup_{n\in\mathbb{N}}\Theta_n=\Theta$. By Lemma

4.23, $\boldsymbol{\theta}_n^*$ is a minimizer of ρ_n^* for any $n \in \mathbb{N}$. By the decomposition of ρ_n^* (see Lemma 4.26) and the Triangle Inequality, for each $n \in \mathbb{N}$ we conclude that

$$\begin{aligned}\sup_{(s,t)\in\Theta_n} |\rho_n^*(s,t) - \rho(s,t)| &= \sup_{(s,t)\in\Theta_n} |\delta_n^*(s,t) + \varrho_n^*(s,t) - \rho(s,t)| \\ &\leq \sup_{(s,t)\in\Theta_n} (|\delta_n^*(s,t)| + |\varrho_n^*(s,t) - \rho(s,t)|) \\ &\leq \sup_{(s,t)\in\Theta_n} |\delta_n^*(s,t)| + \sup_{(s,t)\in\Theta_n} |\varrho_n^*(s,t) - \rho(s,t)| \,.\end{aligned}$$

Letting $n \to \infty$, Lemmas 4.27 and 4.28 lead to

$$\sup_{(s,t)\in\Theta_n} |\rho_n^*(s,t) - \rho(s,t)| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

In addition, $\boldsymbol{\theta} \in \Theta$ is the well-separated minimizer of ρ by Propsition 3.11. An application of Theorem 2.1 gives the claim. □

Corollary 4.30. *If $M_4 < \infty$, then*

$$(\boldsymbol{\theta}_n^*, \hat{\boldsymbol{\alpha}}_n) \xrightarrow[n\to\infty]{\mathbb{P}} (\boldsymbol{\theta}, \boldsymbol{\alpha}).$$

Proof. The assertion follows from Theorems 4.29 and 4.18 and the properties of convergence in probability. □

4.2.3 Convergence in distribution

This section is devoted to the study of convergence in distribution of $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$.

The approach to get another main result of this work is very similar to Section 3.3, but some proofs are technical harder. To apply Theorem 2.3, we have to consider the rescaled process, which is minimized by $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$. For this purpose, recall the notation

$$H_n = \left\{(k,l) \in \mathbb{Z}^2 \middle| k \geq 1 - \tau_n,\ l - k \geq 1 - (\sigma_n - \tau_n),\ n - l \geq \sigma_n + 1\right\}.$$

The rescaled process Γ_n^* is defined by

$$\Gamma_n^*(k,l) := M_n^*(\tau_n + k, \sigma_n + l) - M_n^*(\tau_n, \sigma_n), \qquad (k,l) \in H_n.$$

Lemma 4.31. *Let $n \in \mathbb{N}$. Then*

$$\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n \in \operatorname{Argmin}\left(\Gamma_n^*\right).$$

Proof. The proof of Lemma 3.14 remains valid for $\boldsymbol{\tau}_n^* = (\tau_n^*, \sigma_n^*)$, Γ_n^* and M_n^* instead of $\bar{\boldsymbol{\tau}}_n = (\bar{\tau}_n, \bar{\sigma}_n)$, $\bar{\Gamma}_n$ and $\bar{M}_n$. □

Γ_n^* has the following form.

Lemma 4.32. *Let $n \in \mathbb{N}$ and $(k,l) \in H_n$. Then*

$$\Gamma_n^*(k,l) = \Gamma_{n,1}^*(k) + \Gamma_{n,2}^*(l)$$

with

$$\Gamma_{n,1}^*(k) := \begin{cases} \sum_{i=1}^{k} a_{n,1}^*(X_{\tau_n+i}), & k \geq 0, \\ -\sum_{i=1}^{-k} a_{n,1}^*(X_{\tau_n-i+1}), & k < 0 \end{cases} \quad \text{and} \quad \Gamma_{n,2}^*(l) := \begin{cases} \sum_{i=1}^{l} a_{n,2}^*(X_{\sigma_n+i}), & l \geq 0, \\ -\sum_{i=1}^{-l} a_{n,2}^*(X_{\sigma_n-i+1}), & l < 0, \end{cases}$$

where $a_{n,1}^*$ and $a_{n,2}^*$ are given by (4.48).

Proof. The proof of Lemma 3.15 works by replacing $\bar{\Gamma}_n$, $\bar{\Gamma}_{n,1}$, $\bar{\Gamma}_{n,2}$, $\bar{M}_n$, a_1 and a_2 by Γ_n^*, $\Gamma_{n,1}^*$, $\Gamma_{n,2}^*$, M_n^*, $a_{n,1}^*$ and $a_{n,2}^*$. □

We next prove convergence in distribution of all finite-dimensional distributions of Γ_n^* (assumption (i) of Theorem 2.3). For this purpose, we previously show that the rescaled processes in the case of known and unknown expectations are stochastically equivalent.

Lemma 4.33. *Let $m \in \mathbb{N}$ and $(k_1, l_1), ..., (k_m, l_m) \in \mathbb{Z}^2$. If $M_4 < \infty$, then*

$$\max_{1 \leq r \leq m} \left|\Gamma_n^*(k_r, l_r) - \bar{\Gamma}_n(k_r, l_r)\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Proof. Fix $m \in \mathbb{N}$ and $(k_1, l_1), ..., (k_m, l_m) \in \mathbb{Z}^2$. The proof of Lemma 3.17 provides that

$$(k_1, l_1), ..., (k_m, l_m) \in H_n \tag{4.62}$$

for a sufficiently large $n \in \mathbb{N}$. It suffices to show that

$$\left|\Gamma_n^*(k_r, l_r) - \bar{\Gamma}_n(k_r, l_r)\right| \xrightarrow[n\to\infty]{\mathbb{P}} 0 \tag{4.63}$$

for all $r \in \{1, ..., m\}$. Fix an arbitrary $r \in \{1, ..., m\}$ and fix $n \in \mathbb{N}$ sufficiently large for a moment. As an example, we show (4.63) for $k_r \geq 0$ and $l_r < 0$. The remaining three cases $k_r \geq 0$, $l_r \geq 0$ and $k_r < 0$, $l_r \geq 0$ and $k_r < 0$, $l_r < 0$ follows analogously. By definitions of Γ_n^* and $\bar{\Gamma}_n$, we have

$$\begin{aligned}&\left|\Gamma_n^*(k_r, l_r) - \bar{\Gamma}_n(k_r, l_r)\right| \\ &= \left|\sum_{i=1}^{k_r} a_{n,1}^*(X_{\tau_n+i}) - \sum_{i=1}^{-l_r} a_{n,2}^*(X_{\sigma_n-i+1}) - \left(\sum_{i=1}^{k_r} a_1(X_{\tau_n+i}) - \sum_{i=1}^{-l_r} a_2(X_{\sigma_n-i+1})\right)\right|.\end{aligned}$$

Moreover, the proof of Lemma 3.17 shows that there exists $n_0 = n_0(k_1, l_1, ..., k_m, l_m) \in \mathbb{N}$ such that for all $n \geq n_0$ we get $X_{\tau_n+i} \sim Q_2$ for each $i \in \{1, ..., k_r\}$ and $X_{\sigma_n-i+1} \sim Q_2$ for each $i \in \{1, ..., -l_r\}$ and condition (4.62) is fulfilled. From now on, let $n \geq n_0$. Centering the observations and applying the Triangle Inequality gives by an easy computation

$$\begin{aligned}&\left|\Gamma_n^*(k_r, l_r) - \bar{\Gamma}_n(k_r, l_r)\right| \\ &\leq 2\left(|\hat{\alpha}_n - \alpha| + \left|\hat{\beta}_n - \beta\right|\right)\left|\sum_{i=1}^{k_r}(X_{\tau_n+i} - \beta)\right| + 2\beta k_r\left(|\hat{\alpha}_n - \alpha| + \left|\hat{\beta}_n - \beta\right|\right) \\ &\quad + k_r\left(\left|\hat{\alpha}_n^2 - \alpha^2\right| + \left|\hat{\beta}_n^2 - \beta^2\right|\right) + 2\left(\left|\hat{\beta}_n - \beta\right| + |\hat{\gamma}_n - \gamma|\right)\left|\sum_{i=1}^{-l}(X_{\sigma_n-i+1} - \beta)\right| \\ &\quad - 2\beta l_r\left(\left|\hat{\beta}_n - \beta\right| + |\hat{\gamma}_n - \gamma|\right) - l_r\left(\left|\hat{\beta}_n^2 - \beta^2\right| + \left|\hat{\gamma}_n^2 - \gamma^2\right|\right).\end{aligned} \tag{4.64}$$

To continue, we have to recall the following property of convergence in probability. Let $(Z_n)_{n\in\mathbb{N}}$ and $(V_n)_{n\in\mathbb{N}}$ be sequences of arbitrary random variables with $Z_n \xrightarrow[n\to\infty]{\mathbb{P}} 0$ and $\lim_{x\to\infty}\limsup_{n\to\infty}\mathbb{P}[|V_n| > x] = 0$. Then $Z_nV_n \xrightarrow[n\to\infty]{\mathbb{P}} 0$.

Furthermore, by Chebyshev's Inequality (see Lemma 2.6) and $\mathbb{V}[X_i] \leq M_2$, $1 \leq i \leq n$, we have for all $x > 0$

$$\mathbb{P}\left[\left|\sum_{i=1}^{k_r}(X_{\tau_n+i} - \beta)\right| > x\right] \leq x^{-2}\sum_{i=1}^{k_r}\mathbb{V}\left[X_{\tau_n+i} - \beta\right] \leq k_r x^{-2} M_2$$

and

$$\mathbb{P}\left[\left|\sum_{i=1}^{-l_r}(X_{\sigma_n-i+1} - \beta)\right| > x\right] \leq x^{-2}\sum_{i=1}^{-l_r}\mathbb{V}\left[X_{\sigma_n-i+1} - \beta\right] \leq -l_r x^{-2} M_2.$$

From $M_4 < \infty$ we see that $M_2 < \infty$. We can deduce that

$$\lim_{x\to\infty}\limsup_{n\to\infty}\mathbb{P}\left[\left|\sum_{i=1}^{k_r}(X_{\tau_n+i} - \beta)\right| > x\right] = 0$$

and

$$\lim_{x\to\infty}\limsup_{n\to\infty}\mathbb{P}\left[\left|\sum_{i=1}^{-l_r}(X_{\sigma_n-i+1} - \beta)\right| > x\right] = 0.$$

Combining (4.64) with the weak consistency of $\hat{\boldsymbol{\alpha}}_n = \left(\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n\right)$ (see Theorem 4.18) leads to (4.63) by an application of the mentioned property of convergence in probability. Since $r \in \{1, ..., m\}$ and $(k_1, l_1), ..., (k_m, l_m)$ are arbitrary, we get the claim. □

Proposition 4.34. *Let $m \in \mathbb{N}$ and $(k_1, l_1), ..., (k_m, l_m) \in \mathbb{Z}^2$. If $M_4 < \infty$, then*

$$(\Gamma_n^*(k_1, l_1), ..., \Gamma_n^*(k_m, l_m)) \xrightarrow[n\to\infty]{\mathcal{L}} (\Gamma(k_1, l_1), ..., \Gamma(k_m, l_m)),$$

where Γ is given by (3.18).

Proof. We apply one of Cramér's Theorems, which says that all finite-dimensional distributions of two stochastically equivalent processes converge in distribution to the finite-dimensional distributions of the same process (see for instance Gänssler and Stute [19, p. 352, Theorem 8.6.2]). An application of Proposition 3.18 and Lemma 4.33 completes the proof. □

The next aim is to prove stochastic boundedness of $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$, which establish assumption (ii) of Theorem 2.3. First recall the notation

$$H_{n,x,\delta} = \{(k,l) \in H_n | x \leq \|(k,l)\| \leq n\delta\}$$

for $n \in \mathbb{N}$, $x > 0$ and $\delta > 0$.

Lemma 4.35. *Let $x > 0$, $\delta > 0$ and $n \in \mathbb{N}$. Then*

$$\{x \leq \|\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n\| \leq n\delta\} \subseteq \bigcup_{(k,l)\in H_{n,x,\delta}} \{-\Gamma_n^*(k,l) \geq 0\}.$$

Proof. The proof of Lemma 3.20 works for $\boldsymbol{\tau}_n^*$, Γ_n^*, M_n^* and Lemma 4.31 instead of $\bar{\boldsymbol{\tau}}_n$, $\bar{\Gamma}_n$, $\bar{M}_n$ and Lemma 3.14. □

We get the following error estimate.

Lemma 4.36. *Suppose that $M_2 < \infty$. Then there exist $n_0 \in \mathbb{N}$, $\delta > 0$ and $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0]$ there exist a constant $C_\varepsilon > 0$, which depends on ε, and another constant $C > 0$ such that for all $n \geq n_0$ we have*

$$\begin{aligned}
&\mathbb{P}[x \leq \|\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n\| \leq n\delta] \\
&\quad \leq C_\varepsilon x^{-1} + C x^{-1} \\
&\qquad + 8\left(\mathbb{P}\left[|\hat{\alpha}_n - \alpha| > \varepsilon\right] + \mathbb{P}\left[\left|\hat{\beta}_n - \beta\right| > \varepsilon\right] + \mathbb{P}\left[|\hat{\gamma}_n - \gamma| > \varepsilon\right]\right),
\end{aligned}$$

for all $x \geq 2$.

Proof. The proof is similar in spirit to the proof of Lemma 3.21 but technical harder. Let $x \geq 2$. By Lemma 3.19, there exist $\delta > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ the conditions hold in Lemma 3.19. Let us regard $n \geq n_0$ and $\delta > 0$ as fixed. By Lemma 4.35, we first observe that

$$\begin{aligned}
\{x \leq \|\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n\| \leq n\delta\} &\subseteq \bigcup_{(k,l)\in H_{n,x,\delta}} \{-\Gamma_n^*(k,l) \geq 0\} \\
&\subseteq \bigcup_{\substack{x\leq|k|\leq n\delta\\ |l|\leq n\delta}} \{-\Gamma_n^*(k,l) \geq 0\} \cup \bigcup_{\substack{|k|\leq n\delta\\ x\leq|l|\leq n\delta}} \{-\Gamma_n^*(k,l) \geq 0\} \\
&=: E \cup F. \qquad (4.65)
\end{aligned}$$

To simplify notation, the fact that all in this proof defined sets, random variables and probabilities depend on n, x, δ or ε is omitted. We give the proof only for the estimate of the probability of E; the other case follows the same pattern. We find that

$$\begin{aligned}
E \subseteq &\bigcup_{\substack{x\leq k\leq n\delta\\ 0\leq l\leq n\delta}} \{-\Gamma_n^*(k,l) \geq 0\} \cup \bigcup_{\substack{x\leq k\leq n\delta\\ -n\delta\leq l<0}} \{-\Gamma_n^*(k,l) \geq 0\} \\
&\cup \bigcup_{\substack{-n\delta\leq k\leq -x\\ 0\leq l\leq n\delta}} \{-\Gamma_n^*(k,l) \geq 0\} \cup \bigcup_{\substack{-n\delta\leq k\leq -x\\ -n\delta\leq l<0}} \{-\Gamma_n^*(k,l) \geq 0\} \\
=: &E^{(++)} \cup E^{(+-)} \cup E^{(-+)} \cup E^{(--)}. \qquad (4.66)
\end{aligned}$$

We only describe our proceeding for the estimate of the probability of $E^{(++)}$ in detail. Inserting the expectations α, β and γ on proper positions we conclude by Lemma 4.32 and the definitions of $a_{n,1}^*$ and $a_{n,2}^*$ that

$$\begin{aligned}
&E^{(++)} \\
&= \bigcup_{x\leq k\leq n\delta} \bigcup_{0\leq l\leq n\delta} \left\{ \sum_{i=1}^{k} -a_{n,1}^*(X_{\tau_n+i}) + \sum_{i=1}^{l} -a_{n,2}^*(X_{\sigma_n+i}) \geq 0 \right\} \\
&= \bigcup_{x\leq k\leq n\delta} \bigcup_{0\leq l\leq n\delta} \\
&\quad \left\{ \sum_{i=1}^{k} \left[2\left(\hat{\alpha}_n - \alpha + \beta - \hat{\beta}_n\right) X_{\tau_n+i} + \hat{\beta}_n^2 - \beta^2 + \alpha^2 - \hat{\alpha}_n^2 - a_1(X_{\tau_n+i}) \right] \right. \\
&\quad \left. + \sum_{i=1}^{l} \left[2\left(\hat{\beta}_n - \beta + \gamma - \hat{\gamma}_n\right) X_{\sigma_n+i} + \hat{\gamma}_n^2 - \gamma^2 + \beta^2 - \hat{\beta}_n^2 - a_2(X_{\sigma_n+i}) \right] \geq 0 \right\},
\end{aligned}$$

where a_1 and a_2 are given by (3.2). The problem occurs that both sums contain the estimators $\hat{\alpha}_n$, $\hat{\beta}_n$ and $\hat{\gamma}_n$, which depend on the observatations $X_1, ..., X_n$. Therefore, the independence of both sums cannot be ensured (as in proof of Lemma 3.21). However, we are able to create independence as follows. The Triangle Inequality leads to

$$
\begin{aligned}
&E^{(++)} \\
&\subseteq \bigcup_{x \le k \le n\delta} \bigcup_{0 \le l \le n\delta} \Bigg\{ \left| \sum_{i=1}^{k} \left[2 \left(\hat{\alpha}_n - \alpha + \beta - \hat{\beta}_n \right) X_{\tau_n + i} + \hat{\beta}_n^2 - \beta^2 + \alpha^2 - \hat{\alpha}_n^2 \right] \right| \\
&\qquad + \left| \sum_{i=1}^{l} \left[2 \left(\hat{\beta}_n - \beta + \gamma - \hat{\gamma}_n \right) X_{\sigma_n + i} + \hat{\gamma}_n^2 - \gamma^2 + \beta^2 - \hat{\beta}_n^2 \right] \right| \\
&\qquad + \sum_{i=1}^{k} -a_1(X_{\tau_n+i}) + \sum_{i=1}^{l} -a_2(X_{\sigma_n+i}) \ge 0 \Bigg\} \\
&\subseteq \bigcup_{x \le k \le n\delta} \bigcup_{0 \le l \le n\delta} \Bigg\{ \sum_{i=1}^{k} \left[2 \left(|\hat{\alpha}_n - \alpha| + \left| \hat{\beta}_n - \beta \right| \right) |X_{\tau_n+i}| + \left| \hat{\alpha}_n^2 - \alpha^2 \right| + \left| \hat{\beta}_n^2 - \beta^2 \right| \right] \\
&\qquad + \sum_{i=1}^{l} \left[2 \left(\left| \hat{\beta}_n - \beta \right| + |\hat{\gamma}_n - \gamma| \right) |X_{\sigma_n+i}| + \left| \hat{\beta}_n^2 - \beta^2 \right| + \left| \hat{\gamma}_n^2 - \gamma^2 \right| \right] \\
&\qquad + \sum_{i=1}^{k} -a_1(X_{\tau_n+i}) + \sum_{i=1}^{l} -a_2(X_{\sigma_n+i}) \ge 0 \Bigg\} \\
&=: \tilde{E}^{(++)}.
\end{aligned}
\tag{4.67}
$$

Write for fixed $\varepsilon > 0$

$$
A_1 := \{ |\hat{\alpha}_n - \alpha| \le \varepsilon \}, \quad A_2 := \left\{ \left| \hat{\beta}_n - \beta \right| \le \varepsilon \right\} \quad \text{and} \quad A_3 := \{ |\hat{\gamma}_n - \gamma| \le \varepsilon \}.
$$

By the rules of De Morgan, we obtain

$$
\begin{aligned}
\tilde{E}^{(++)} &= \left(\tilde{E}^{(++)} \cap (A_1 \cap A_2 \cap A_3) \right) \cup \left(\tilde{E}^{(++)} \cap (A_1 \cap A_2 \cap A_3)^{\complement} \right) \\
&\subseteq \left(\tilde{E}^{(++)} \cap A_1 \cap A_2 \cap A_3 \right) \cup \left(A_1^{\complement} \cup A_2^{\complement} \cup A_3^{\complement} \right) \\
&=: \tilde{E}_1^{(++)} \cup A.
\end{aligned}
\tag{4.68}
$$

We next treat the set $\tilde{E}_1^{(++)}$. Note that the Binomial Formula and the Triangle Inequality gives for all $a, b \in \mathbb{R}$

$$|a^2 - b^2| = |a - b| \cdot |a - b + 2b| \leq |a - b| \cdot (|a - b| + 2|b|) = |a - b|^2 + 2|a - b| \cdot |b|.$$

On the events A_1, A_2 and A_3 we have

$$\begin{aligned}
\left|\hat{\alpha}_n^2 - \alpha^2\right| &\leq \left|\hat{\alpha}_n - \alpha\right|^2 + 2\left|\hat{\alpha}_n - \alpha\right| \cdot |\alpha| \leq \varepsilon^2 + 2\varepsilon|\alpha|, \\
\left|\hat{\beta}_n^2 - \beta^2\right| &\leq \left|\hat{\beta}_n - \beta\right|^2 + 2\left|\hat{\beta}_n - \beta\right| \cdot |\beta| \leq \varepsilon^2 + 2\varepsilon|\beta| \qquad \text{and} \\
\left|\hat{\gamma}_n^2 - \gamma^2\right| &\leq \left|\hat{\gamma}_n - \gamma\right|^2 + 2\left|\hat{\gamma}_n - \gamma\right| \cdot |\gamma| \leq \varepsilon^2 + 2\varepsilon|\gamma|.
\end{aligned}$$

Hence

$$\begin{aligned}
\tilde{E}_1^{(++)} &= \tilde{E}^{(++)} \cap A_1 \cap A_2 \cap A_3 \\
&\subseteq \bigcup_{x \leq k \leq n\delta} \bigcup_{0 \leq l \leq n\delta} \left\{ \sum_{i=1}^{k} \left[4\varepsilon \left|X_{\tau_n+i}\right| + 2\varepsilon^2 + 2\varepsilon(|\alpha| + |\beta|) - a_1(X_{\tau_n+i})\right] \right. \\
&\qquad\qquad \left. + \sum_{i=1}^{l} \left[4\varepsilon \left|X_{\sigma_n+i}\right| + 2\varepsilon^2 + 2\varepsilon(|\beta| + |\gamma|) - a_2(X_{\sigma_n+i})\right] \geq 0 \right\}.
\end{aligned}$$

Set

$$Y_1^{(+)} := \max_{x \leq k \leq n\delta} \sum_{i=1}^{k} \eta_{1,i}^{(+)} \quad \text{and} \quad Y_2^{(+)} := \max_{0 \leq l \leq n\delta} \sum_{i=1}^{l} \eta_{2,i}^{(+)},$$

where

$$\begin{aligned}
\eta_{1,i}^{(+)} &:= 4\varepsilon \left|X_{\tau_n+i}\right| + 2\varepsilon^2 + 2\varepsilon(|\alpha| + |\beta|) - a_1(X_{\tau_n+i}) \quad \text{and} \\
\eta_{2,i}^{(+)} &:= 4\varepsilon \left|X_{\sigma_n+i}\right| + 2\varepsilon^2 + 2\varepsilon(|\beta| + |\gamma|) - a_2(X_{\sigma_n+i}).
\end{aligned}$$

It follows analogously to (3.25) that

$$\tilde{E}_1^{(++)} \subseteq \left\{ Y_1^{(+)} + Y_2^{(+)} \geq 0 \right\}. \tag{4.69}$$

Since $\delta > 0$ is suffiently small (see Lemma 3.19) and the estimators $\hat{\alpha}_n$, $\hat{\beta}_n$ and $\hat{\gamma}_n$ are eliminated, we can conclude that

$$Y_1^{(+)} = Y_1^{(+)}\left(X_{\tau_n+1}, ..., X_{\tau_n+\lfloor n\delta \rfloor}\right) \quad \text{and} \quad Y_2^{(+)} = Y_2^{(+)}\left(X_{\sigma_n+1}, ..., X_{\sigma_n+\lfloor n\delta \rfloor}\right)$$

as two measurable transformations of two independent vectors $\left(X_{\tau_n+1}, ..., X_{\tau_n+\lfloor n\delta \rfloor}\right)$ and $\left(X_{\sigma_n+1}, ..., X_{\sigma_n+\lfloor n\delta \rfloor}\right)$ are also independent. By (4.69) and Lemma A.8, we deduce that

$$\mathbb{P}\left[\tilde{E}_1^{(++)}\right] \leq \mathbb{P}\left[Y_1^{(+)} + Y_2^{(+)} \geq 0\right] = \int_{\mathbb{R}} \mathbb{P}\left[Y_1^{(+)} \geq -y\right] \; \mathbb{P}_{Y_2^{(+)}}(\mathrm{d}y). \tag{4.70}$$

For abbreviation, we write $Z_{1,i} := \eta_{1,i}^{(+)} - \mathbb{E}\left[\eta_{1,i}^{(+)}\right]$, $1 \leq i \leq \lfloor n\delta \rfloor$. We next consider the integrand. We find for all $y \in \mathbb{R}$ that

$$\begin{aligned}\mathbb{P}\left[Y_1^{(+)} \geq -y\right] &= \mathbb{P}\left[\max_{x \leq k \leq n\delta} \sum_{i=1}^{k} \eta_{1,i}^{(+)} \geq -y\right] \\ &= \mathbb{P}\left[\bigcup_{x \leq k \leq n\delta} \left\{\sum_{i=1}^{k} Z_{1,i} \geq \sum_{i=1}^{k} \mathbb{E}\left[-\eta_{1,i}^{(+)}\right] - y\right\}\right].\end{aligned}$$

By Lemma 3.19 (ii), we conclude that $\tau_n + 1 \leq \tau_n + i < \sigma_n$ for $1 \leq i \leq k$ with $x \leq k \leq n\delta$. In the proof of Lemma 3.8 we have seen that

$$\mathbb{E}[a_1(X_{\tau_n+i})] = (\alpha-\beta)^2 \quad \text{for} \quad 1 \leq i \leq k \quad \text{with} \quad x \leq k \leq n\delta.$$

We thus get for $1 \leq i \leq k$ with $x \leq k \leq n\delta$

$$\begin{aligned}\mathbb{E}\left[-\eta_{1,i}^{(+)}\right] &= \mathbb{E}[a_1(X_{\tau_n+i})] - \left(4\varepsilon\mathbb{E}\left[|X_{\tau_n+i}|\right] + 2\varepsilon^2 + 2\varepsilon(|\alpha| + |\beta|)\right) \\ &\geq (\alpha-\beta)^2 - \left(4\varepsilon M_1 + 2\varepsilon^2 + 2\varepsilon(|\alpha| + |\beta|)\right).\end{aligned}$$

From $M_2 < \infty$ we see that $M_1 < \infty$. Since $\varepsilon > 0$ is arbitrary, there exists $\varepsilon_1 = \varepsilon_1(\alpha, \beta) > 0$ such that

$$4\varepsilon M_1 + 2\varepsilon^2 + 2\varepsilon(|\alpha| + |\beta|) \leq \frac{1}{2}(\alpha-\beta)^2$$

for all $\varepsilon \leq \varepsilon_1$. From now on, let $\varepsilon \leq \varepsilon_1$. We obtain for $1 \leq i \leq k$ with $x \leq k \leq n\delta$

$$\mathbb{E}\left[-\eta_{1,i}^{(+)}\right] \geq \frac{1}{2}(\alpha-\beta)^2.$$

By model assumptions $\alpha \neq \beta$ and $\beta \neq \gamma$, it holds $\mu := \min\left\{(\alpha-\beta)^2, (\beta-\gamma)^2\right\} > 0$. It follows for all $y \in \mathbb{R}$ that

$$\begin{aligned}\mathbb{P}\left[Y_1^{(+)} \geq -y\right] &\leq \mathbb{P}\left[\bigcup_{x\leq k\leq n\delta}\left\{\sum_{i=1}^{k} Z_{1,i} \geq \frac{1}{2}k(\alpha-\beta)^2 - y\right\}\right]\\ &\leq \mathbb{P}\left[\bigcup_{x\leq k\leq n\delta}\left\{\sum_{i=1}^{k} Z_{1,i} \geq \frac{1}{2}k\mu - y\right\}\right]\\ &=: \mathbb{P}(y). \end{aligned} \tag{4.71}$$

We distinguish several cases for y to get an estimate for $\mathbb{P}(y)$. Let $C_\varepsilon = C_\varepsilon(\alpha,\beta,\gamma) > 0$ be a generic constant, which depends on ε, and let $C = C(\alpha,\beta,\gamma) > 0$ be another generic constant.

(i) In the case $y \leq 0$ we have $-y \geq 0$. Applying the same arguments used in case (i) in the proof of Lemma 3.21 leads to

$$\mathbb{P}(y) \leq 4\mu^{-2}\left(\lfloor x\rfloor^{-2}\sum_{i=1}^{\lfloor x\rfloor}\mathbb{V}[Z_{1,i}] + \sum_{i=\lfloor x\rfloor+1}^{\lfloor n\delta\rfloor} i^{-2}\mathbb{V}[Z_{1,i}]\right). \tag{4.72}$$

We use the abbreviation $\tilde{Z}_{1,i} := 4\varepsilon\,|X_{\tau_n+i}| - 2(\beta-\alpha)X_{\tau_n+i}$, $1 \leq i \leq \lfloor n\delta\rfloor$, to estimate the variance. Fix for a moment $i \in \{1,...,\lfloor n\delta\rfloor\}$. By definition of $Z_{1,i}$ and the calculation rules of variances, a trivial verification shows that

$$\mathbb{V}[Z_{1,i}] = \mathbb{V}\left[\tilde{Z}_{1,i}\right] = \mathbb{E}\left[\tilde{Z}_{1,i}^2\right] - \left(\mathbb{E}\left[\tilde{Z}_{1,i}\right]\right)^2 \leq \mathbb{E}\left[\tilde{Z}_{1,i}^2\right] = \mathbb{E}\left[\left|\tilde{Z}_{1,i}\right|^2\right].$$

Moreover, by the Triangle Inequality and the Binomial Formula, we can assert that

$$\begin{aligned}\mathbb{E}\left[\left|\tilde{Z}_{1,i}\right|^2\right] &\leq \mathbb{E}\left[\left(4\varepsilon\,|X_{\tau_n+i}| + |2(\beta-\alpha)X_{\tau_n+i}|\right)^2\right]\\ &= 16\varepsilon^2\mathbb{E}\left[|X_{\tau_n+i}|^2\right] + 16\varepsilon|\alpha-\beta|\mathbb{E}\left[|X_{\tau_n+i}|^2\right] + 4(\alpha-\beta)^2\mathbb{E}\left[|X_{\tau_n+i}|^2\right]\\ &\leq \left(16\varepsilon^2 + 16\varepsilon|\alpha-\beta|\right)M_2 + 4(\alpha-\beta)^2 M_2\\ &\leq C_\varepsilon + C,\end{aligned}$$

since $M_2 < \infty$. By (4.72), an analogous proceeding as in case (i) in the proof of Lemma 3.21 gives for $x \geq 2$

$$\mathbb{P}(y) \leq C_\varepsilon x^{-1} + C x^{-1}.$$

(ii) Let $y > 0$. By $k \geq x$, we have

$$\frac{1}{2}k\mu - y = k\left(\frac{1}{2}\mu - \frac{y}{k}\right) \geq k\left(\frac{1}{2}\mu - \frac{y}{x}\right).$$

(a) Let $0 < y < \frac{1}{4}\mu x$. It follows that $\frac{1}{2}k\mu - y \geq \frac{1}{4}k\mu$. As in (i), we obtain

$$\begin{aligned}\mathbb{P}(y) &\leq \mathbb{P}\left[\bigcup_{x\leq k\leq n\delta}\left\{\sum_{i=1}^{k} Z_{1,i} \geq \frac{1}{4}k\mu\right\}\right]\\ &\leq C_\varepsilon x^{-1} + Cx^{-1}.\end{aligned}$$

(b) In the case $y \geq \frac{1}{4}\mu x$ we estimate $\mathbb{P}(y) \leq 1$.

Applying (4.70) and (4.71) gives

$$\begin{aligned}&\mathbb{P}\left[\tilde{E}_1^{(++)}\right]\\ &\leq \int_{(-\infty,0]} \mathbb{P}(y)\ \mathbb{P}_{Y_2^{(+)}}(\mathrm{d}y) + \int_{\left(0,\frac{1}{4}\mu x\right)} \mathbb{P}(y)\ \mathbb{P}_{Y_2^{(+)}}(\mathrm{d}y) + \int_{\left[\frac{1}{4}\mu x,\infty\right)} \mathbb{P}(y)\ \mathbb{P}_{Y_2^{(+)}}(\mathrm{d}y)\\ &\leq (C_\varepsilon x^{-1} + Cx^{-1})\mathbb{P}\left[Y_2^{(+)} \leq 0\right]\\ &\quad + (C_\varepsilon x^{-1} + Cx^{-1})\mathbb{P}\left[0 < Y_2^{(+)} < \frac{1}{4}\mu x\right] + \mathbb{P}\left[Y_2^{(+)} \geq \frac{1}{4}\mu x\right]\\ &\leq C_\varepsilon x^{-1} + Cx^{-1} + \mathbb{P}\left[Y_2^{(+)} \geq \frac{1}{4}\mu x\right]. \end{aligned} \tag{4.73}$$

For abbreviation, we write $Z_{2,i} := \eta_{2,i}^{(+)} - \mathbb{E}\left[\eta_{2,i}^{(+)}\right]$, $1 \leq i \leq \lfloor n\delta\rfloor$. We now handle the probability in the last estimate. By definition, we have

$$\begin{aligned}\left\{Y_2^{(+)} \geq \frac{1}{4}\mu x\right\} &= \left\{\max_{0\leq l\leq n\delta}\sum_{i=1}^{l}\eta_{2,i}^{(+)} \geq \frac{1}{4}\mu x\right\}\\ &= \bigcup_{0\leq l\leq n\delta}\left\{\sum_{i=1}^{l} Z_{2,i} \geq \sum_{i=1}^{l}\mathbb{E}\left[-\eta_{2,i}^{(+)}\right] + \frac{1}{4}\mu x\right\}\end{aligned}$$

Note that $\left\{\sum_{i=1}^{l} Z_{2,i} \geq \sum_{i=1}^{l}\mathbb{E}\left[-\eta_{2,i}^{(+)}\right] + \frac{1}{4}\mu x\right\} = \emptyset$ for $l = 0$, because $\frac{1}{4}\mu x > 0$. From Lemma 3.19 (iii) we deduce that $\sigma_n + 1 \leq \sigma_n + i < n$ for $1 \leq i \leq l$ with $1 \leq l \leq n\delta$. The proof of Lemma 3.8 provides

$$\mathbb{E}[a_2(X_{\sigma_n+i})] = (\beta-\gamma)^2 \quad \text{for} \quad 1 \leq i \leq l \quad \text{with} \quad 1 \leq l \leq n\delta.$$

By similar arguments used for the estimate of $\mathbb{E}\left[-\eta_{1,i}^{(+)}\right]$, there exists $\varepsilon_2 = \varepsilon_2(\beta,\gamma) > 0$ such that

$$\mathbb{E}\left[-\eta_{2,i}^{(+)}\right] \geq \frac{1}{2}(\beta-\gamma)^2 \geq \frac{1}{2}\mu$$

for all $\varepsilon \leq \varepsilon_2$ and $1 \leq i \leq l$ with $1 \leq l \leq n\delta$. From now on, let $\varepsilon \leq \min\{\varepsilon_1, \varepsilon_2\}$. It follows that

$$\begin{aligned}\left\{Y_2^{(+)} \geq \frac{1}{4}\mu x\right\} &\subseteq \bigcup_{1\leq l\leq n\delta}\left\{\sum_{i=1}^{l} Z_{2,i} \geq \frac{1}{2}l\mu + \frac{1}{4}\mu x\right\}\\ &\subseteq \bigcup_{1\leq l\leq n\delta}\left\{(2l+x)^{-1}\left|\sum_{i=1}^{l} Z_{2,i}\right| \geq \frac{1}{4}\mu\right\}.\end{aligned}$$

The further proceeding is analogous to the approach of $Y_2^{(+)}$ in the proof of Lemma 3.21. We get

$$\mathbb{P}\left[Y_2^{(+)} \geq \frac{1}{4}\mu x\right] \leq 16\mu^{-2}\sum_{l=1}^{\lfloor n\delta\rfloor}(2l+x)^{-2}\mathbb{V}[Z_{2,l}].$$

The variance can be handled as in case (i). We obtain

$$\mathbb{V}[Z_{2,l}] \leq \left(16\varepsilon^2 + 16\varepsilon|\beta-\gamma|\right) M_2 + 4(\beta-\gamma)^2 M_2 \leq C_\varepsilon + C$$

for $1 \leq l \leq \lfloor n\delta\rfloor$. We infer in much the same way as in proof of Lemma 3.21

$$\mathbb{P}\left[Y_2^{(+)} \geq \frac{1}{4}\mu x\right] \leq C_\varepsilon x^{-1} + Cx^{-1}.$$

By (4.73), we see that

$$\mathbb{P}\left[\tilde{E}_1^{(++)}\right] \leq C_\varepsilon x^{-1} + Cx^{-1}.$$

Summarizing, by (4.67) and (4.68), we have for all $\varepsilon \leq \min\{\varepsilon_1, \varepsilon_2\}$

$$\begin{aligned}\mathbb{P}\left[E^{(++)}\right] &\leq \mathbb{P}\left[\tilde{E}_1^{(++)}\right] + \mathbb{P}[A]\\ &\leq C_\varepsilon x^{-1} + Cx^{-1}\\ &\quad + \mathbb{P}\left[|\hat{\alpha}_n - \alpha| > \varepsilon\right] + \mathbb{P}\left[\left|\hat{\beta}_n - \beta\right| > \varepsilon\right] + \mathbb{P}\left[|\hat{\gamma}_n - \gamma| > \varepsilon\right].\end{aligned}$$

The rest of the proof runs as before. We outline the proof for $E^{(+-)}$, $E^{(-+)}$ and $E^{(--)}$. Set

$$Y_1^{(-)} := \max_{-n\delta \le k \le -x} \sum_{i=1}^{-k} \eta_{1,i}^{(-)} \quad \text{and} \quad Y_2^{(-)} := \max_{-n\delta \le l < 0} \sum_{i=1}^{-l} \eta_{2,i}^{(-)},$$

where

$$\eta_{1,i}^{(-)} := 4\varepsilon \left| X_{\tau_n - i+1} \right| + 2\varepsilon^2 + 2\varepsilon(|\alpha| + |\beta|) + a_1(X_{\tau_n - i+1}) \qquad \text{and}$$
$$\eta_{2,i}^{(-)} := 4\varepsilon \left| X_{\sigma_n - i+1} \right| + 2\varepsilon^2 + 2\varepsilon(|\beta| + |\gamma|) + a_2(X_{\sigma_n - i+1}).$$

A similar approach as in the first part of the proof leads to

$$E^{(+-)} \subseteq \left\{ Y_1^{(+)} + Y_2^{(-)} \ge 0 \right\} \cup A, \quad E^{(-+)} \subseteq \left\{ Y_1^{(-)} + Y_2^{(+)} \ge 0 \right\} \cup A \quad \text{and}$$
$$E^{(--)} \subseteq \left\{ Y_1^{(-)} + Y_2^{(-)} \ge 0 \right\} \cup A,$$

where A is given by (4.68). The pairwise independence of the measurable transformations

$$Y_1^{(+)} = Y_1^{(+)} \left(X_{\tau_n+1}, ..., X_{\tau_n + \lfloor n\delta \rfloor} \right) \quad \text{and} \quad Y_2^{(-)} = Y_2^{(-)} \left(X_{\sigma_n - \lfloor n\delta \rfloor + 1}, ..., X_{\sigma_n} \right),$$
$$Y_1^{(-)} = Y_1^{(-)} \left(X_{\tau_n - \lfloor n\delta \rfloor + 1}, ..., X_{\tau_n} \right) \quad \text{and} \quad Y_2^{(+)} = Y_2^{(+)} \left(X_{\sigma_n+1}, ..., X_{\sigma_n + \lfloor n\delta \rfloor} \right),$$
$$Y_1^{(-)} = Y_1^{(-)} \left(X_{\tau_n - \lfloor n\delta \rfloor + 1}, ..., X_{\tau_n} \right) \quad \text{and} \quad Y_2^{(-)} = Y_2^{(-)} \left(X_{\sigma_n - \lfloor n\delta \rfloor + 1}, ..., X_{\sigma_n} \right)$$

follow from Lemma 3.19 and the independence of the observations $X_1, ..., X_n$. Lemma 3.19 shows that $1 \le \tau_n - i + 1 \le \tau_n$ for $1 \le i \le -k$ with $x \le -k \le n\delta$ and $\tau_n + 1 \le \sigma_n - i + 1 \le \sigma_n$ for $1 \le i \le -l$ with $1 \le -l \le n\delta$. The proof of Lemma 3.8 establishes

$$\mathbb{E}[a_1(X_{\tau_n - i+1})] = -(\alpha - \beta)^2 \quad \text{for} \quad 1 \le i \le -k \quad \text{with} \quad x \le -k \le n\delta \quad \text{and}$$
$$\mathbb{E}[a_2(X_{\sigma_n - i+1})] = -(\beta - \gamma)^2 \quad \text{for} \quad 1 \le i \le -l \quad \text{with} \quad 1 \le -l \le n\delta.$$

Hence there exist $\varepsilon_3 = \varepsilon_3(\alpha, \beta) > 0$ and $\varepsilon_4 = \varepsilon_4(\beta, \gamma) > 0$ such that

$$\mathbb{E}\left[-\eta_{1,i}^{(-)} \right] \ge \frac{1}{2}(\alpha - \beta)^2 \text{ for all } \varepsilon \le \varepsilon_3 \text{ and } 1 \le i \le -k \text{ with } x \le -k \le n\delta \text{ and}$$
$$\mathbb{E}\left[-\eta_{2,i}^{(-)} \right] \ge \frac{1}{2}(\beta - \gamma)^2 \text{ for all } \varepsilon \le \varepsilon_4 \text{ and } 1 \le i \le -l \text{ with } 1 \le -l \le n\delta.$$

Similar arguments used in the estimate of the probability of $E^{(++)}$ lead to

$$\mathbb{P}\left[E^{(+-)}\right] \leq C_\varepsilon x^{-1} + Cx^{-1} + \mathbb{P}\left[|\hat{\alpha}_n - \alpha| > \varepsilon\right] + \mathbb{P}\left[\left|\hat{\beta}_n - \beta\right| > \varepsilon\right] + \mathbb{P}\left[|\hat{\gamma}_n - \gamma| > \varepsilon\right]$$

for all $\varepsilon \leq \min\{\varepsilon_1, \varepsilon_4\}$,

$$\mathbb{P}\left[E^{(-+)}\right] \leq C_\varepsilon x^{-1} + Cx^{-1} + \mathbb{P}\left[|\hat{\alpha}_n - \alpha| > \varepsilon\right] + \mathbb{P}\left[\left|\hat{\beta}_n - \beta\right| > \varepsilon\right] + \mathbb{P}\left[|\hat{\gamma}_n - \gamma| > \varepsilon\right]$$

for all $\varepsilon \leq \min\{\varepsilon_2, \varepsilon_3\}$ and

$$\mathbb{P}\left[E^{(--)}\right] \leq C_\varepsilon x^{-1} + Cx^{-1} + \mathbb{P}\left[|\hat{\alpha}_n - \alpha| > \varepsilon\right] + \mathbb{P}\left[\left|\hat{\beta}_n - \beta\right| > \varepsilon\right] + \mathbb{P}\left[|\hat{\gamma}_n - \gamma| > \varepsilon\right]$$

for all $\varepsilon \leq \min\{\varepsilon_3, \varepsilon_4\}$. Applying (4.66) yields

$$\begin{aligned}\mathbb{P}[E] &\leq \mathbb{P}\left[E^{(++)}\right] + \mathbb{P}\left[E^{(+-)}\right] + \mathbb{P}\left[E^{(-+)}\right] + \mathbb{P}\left[E^{(--)}\right] \\ &\leq C_\varepsilon x^{-1} + Cx^{-1} + 4\left(\mathbb{P}\left[|\hat{\alpha}_n - \alpha| > \varepsilon\right] + \mathbb{P}\left[\left|\hat{\beta}_n - \beta\right| > \varepsilon\right] + \mathbb{P}\left[|\hat{\gamma}_n - \gamma| > \varepsilon\right]\right)\end{aligned}$$

for all $\varepsilon \leq \varepsilon_0^{(1)} := \min\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}$. In the same manner we can see that there exists $\varepsilon_0^{(2)} = \varepsilon_0^{(2)}(\alpha, \beta, \gamma) > 0$ such that

$$\mathbb{P}[F] \leq C_\varepsilon x^{-1} + Cx^{-1} + 4\left(\mathbb{P}\left[|\hat{\alpha}_n - \alpha| > \varepsilon\right] + \mathbb{P}\left[\left|\hat{\beta}_n - \beta\right| > \varepsilon\right] + \mathbb{P}\left[|\hat{\gamma}_n - \gamma| > \varepsilon\right]\right)$$

for all $\varepsilon \leq \varepsilon_0^{(2)}$. Altogether, by (4.65), we have for all $\varepsilon \leq \varepsilon_0 := \min\left\{\varepsilon_0^{(1)}, \varepsilon_0^{(2)}\right\}$

$$\begin{aligned}&\mathbb{P}[x \leq \|\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n\| \leq n\delta] \\ &\quad \leq \mathbb{P}[E] + \mathbb{P}[F] \\ &\quad \leq C_\varepsilon x^{-1} + Cx^{-1} + 8\left(\mathbb{P}\left[|\hat{\alpha}_n - \alpha| > \varepsilon\right] + \mathbb{P}\left[\left|\hat{\beta}_n - \beta\right| > \varepsilon\right] + \mathbb{P}\left[|\hat{\gamma}_n - \gamma| > \varepsilon\right]\right). \square\end{aligned}$$

We obtain stochastic boundedness of $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$ (assumption (ii) of Theorem 2.3).

Proposition 4.37. *If $M_4 < \infty$, then*

$$\lim_{x\to\infty} \limsup_{n\to\infty} \mathbb{P}[\|\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n\| \geq x] = 0.$$

Proof. The same proceeding as in proof of Proposition 3.22 leads to

$$\mathbb{P}[\|\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n\| \geq x] \leq \mathbb{P}[x \leq \|\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n\| \leq n\delta] + \mathbb{P}\left[\|\boldsymbol{\theta}_n^* - \boldsymbol{\theta}\| > \frac{1}{2}\delta\right] + \mathbb{P}\left[1 > \frac{1}{2}n\delta\right]$$

with $\delta > 0$, $x > 0$ and $n \in \mathbb{N}$. Applying the error estimate in Lemma 4.36 and the weak consistency of $\hat{\boldsymbol{\alpha}}_n = \left(\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n\right)$ and $\boldsymbol{\theta}_n^*$ (see Theorems 4.18 and 4.29) implies

$$\limsup_{n\to\infty} \mathbb{P}[\|\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n\| \geq x] \leq C_\varepsilon x^{-1} + Cx^{-1},$$

where $C_\varepsilon > 0$ ($\varepsilon > 0$ sufficiently small), $C > 0$ and $x \geq 2$. Letting $x \to \infty$ we get the claim. □

We can now formulate and prove another main result of this work. If all distributions are continuous, it turns out that $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$ converges in distribution to the minimizer of Γ, where Γ is a sum of random walks, which we have already investigated in Chapter 3.

Theorem 4.38. *If $M_4 < \infty$, then*

$$\limsup_{n\to\infty} \mathbb{P}[\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n \in F] \leq \mathbb{P}[\mathrm{Argmin}(\Gamma) \cap F \neq \emptyset] \qquad \textit{for all } F \subseteq \mathbb{Z}^2.$$

In addition, if Q_1, Q_2 and Q_3 are continuous, then $\mathrm{Argmin}(\Gamma) = \{\boldsymbol{T}\}$ *almost surely and*

$$\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n \xrightarrow[n\to\infty]{\mathcal{L}} \boldsymbol{T} \quad \textit{in} \quad \mathbb{Z}^2.$$

Proof. We apply Theorem 2.3. By Lemmas 4.31, 3.23 and 3.25, we first observe that $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$ is a minimizer of Γ_n^* and Γ has at least one minimizer. Propositions 4.34 and 4.37 establish assumptions (i) and (ii) of Theorem 2.3, which give the first claim. By Proposition 3.26, we conclude the second claim. □

Corollary 4.39. *Suppose that $M_4 < \infty$. Let Q_1, Q_2 and Q_3 be continuous distributions. Then*

$$\tau_n^* - \tau_n \xrightarrow[n\to\infty]{\mathcal{L}} \operatorname*{argmin}_{k\in\mathbb{Z}} \Gamma_1(k) \quad \textit{in} \quad \mathbb{Z} \qquad \textit{and} \qquad \sigma_n^* - \sigma_n \xrightarrow[n\to\infty]{\mathcal{L}} \operatorname*{argmin}_{l\in\mathbb{Z}} \Gamma_2(l) \quad \textit{in} \quad \mathbb{Z}.$$

Proof. The proof of Corollary 3.28 works by replacing Theorem 3.27 and $\bar{\boldsymbol{\tau}}_n = (\bar{\tau}_n, \bar{\sigma}_n)$ by Theorem 4.38 and $\boldsymbol{\tau}_n^* = (\tau_n^*, \sigma_n^*)$. □

4.2.4 Asymptotic confidence region

This section provides an asymptotic confidence region to estimate the moments of change $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ in the case of unknown expectations.

For this purpose, let $F^{-1}_{\|\boldsymbol{T}\|}(\vartheta)$, $\vartheta \in (0,1)$, stand for the ϑ-quantile of the distribution function $F_{\|\boldsymbol{T}\|}$ of $\|\boldsymbol{T}\|$, where $\boldsymbol{T}$ is the almost surely unique minimizer of Γ (see Theorem 4.38).

To get an asymptotic confidence region, we use a similar approach as in Section 3.4, i.e., the convergence in distribution of $\boldsymbol{\tau}^*_n - \boldsymbol{\tau}_n$ (see Theorem 4.38) and the Continuous Mapping Theorem are applied.

Theorem 4.40. *Suppose that $M_4 < \infty$. Let Q_1, Q_2 and Q_3 be continuous distributions and $\vartheta \in (0,1)$. For each $n \in \mathbb{N}$, the random interval*

$$I_n(\vartheta) := \Big[\tau^*_n - F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta), \tau^*_n + F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta)\Big] \times \Big[\sigma^*_n - F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta), \sigma^*_n + F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta)\Big]$$

is an asymptotic confidence region for $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ at level $1-\vartheta$.

Proof. The proof of Theorem 3.29 remains valid for $\boldsymbol{\tau}^*_n = (\tau^*_n, \sigma^*_n)$ and Theorem 4.38 instead of $\bar{\boldsymbol{\tau}}_n = (\bar{\tau}_n, \bar{\sigma}_n)$ and Theorem 3.27. □

The important point to note here is that the quantile $F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta)$, $\vartheta \in (0,1)$ is unknown, but it can be approximated by a Monte-Carlo method as described in Section 3.4. Indeed, we use the estimators $\hat{\alpha}_n$, $\hat{\beta}_n$ and $\hat{\gamma}_n$ (defined in (4.5)) instead of the unknown expectations α, β and γ to generate the process Γ (compare step (i) (2) in Section 3.4).

For further investigation of the asymptotic confidence region one can find a simulation study in Chapter 5, which contains of approximated coverage probabilities for given confidence levels.

Chapter 5

Simulation study

This chapter contains a brief summary of conclusions based on a simulation study in the software environment R, version 3.3.1.

To get empirical results, we choose a multiple change-point $\boldsymbol{\theta} = (\theta_1, \theta_2)$ with $0 < \theta_1 < \theta_2 < 1$ and the distributions Q_1, Q_2, Q_3, where the first moments of adjacent segments are different (see (1.2)). Furthermore, we generate a data set of independent observations $X_1, ..., X_n$ such that the segments $X_1, ..., X_{\lfloor n\theta_1 \rfloor}$, $X_{\lfloor n\theta_1 \rfloor + 1}, ..., X_{\lfloor n\theta_2 \rfloor}$ and $X_{\lfloor n\theta_2 \rfloor + 1}, ..., X_n$ arise from the given distributions (compare (1.1)). For some examples, the results of this simulation study are given in Appendix C.

At first we look closer at the criterion functions $\bar{M}_n$, $\hat{M}_n$ and M_n^* to confirm Lemma 3.3 and the Conjectures 4.4 and 4.22. As an example, in Table C.1 we set $n = 10$ and $\boldsymbol{\theta} = (0.4, 0.8)$ and consider different binomial, poisson, normal and exponential distributions Q_1, Q_2 and Q_3, where the expectations are fixed by $\boldsymbol{\alpha} = (1, 2, 1)$. The sets of all minimizers and maximizers are computed by 10^6 Monte-Carlo repetitions, respectively. If the distributions are chosen to be continuous, as in the second block in Table C.1, then we observe that $\bar{M}_n$ has exactly one minimizer, $\hat{M}_n$ has one maximizer and M_n^* has only one minimizer, namely the smallest maximizer of $\hat{M}_n$ (which is used to construct M_n^*, compare (4.49), (4.48) and (4.5)). However, it is not sufficient to assume that at least one distribution must be continuous, as one can see in the third block in Table C.1.

In order to evaluate the performance of the estimators $\bar{\boldsymbol{\theta}}_n$, $\hat{\boldsymbol{\theta}}_n$ and $\boldsymbol{\theta}_n^*$ of the multiple change-point and the estimator $\hat{\boldsymbol{\alpha}}_n$ of expectations, the bias and root mean square error (RMSE) were estimated over 10^4 Monte-Carlo repetitions. For this purpose, different normal distributions from which the observations arise were chosen. As a simple consequence of Conjecture 4.22, we obtain the same results by computing $\hat{\boldsymbol{\theta}}_n$ and $\boldsymbol{\theta}_n^*$. Therefore, we only consider $\boldsymbol{\theta}_n^*$ in the case of unknown expectations.
The empirical results, listed in Tables C.2 and C.3, indicate some conclusions. We find that the estimation is more accurate for larger sample sizes, which is clear due to the consistency of our estimators (see Theorems 3.12, 4.29 and 4.18). Though, it is evident that $\bar{\boldsymbol{\theta}}_n$ performs better than $\boldsymbol{\theta}_n^*$, because the expectations are assumed to be known to compute $\bar{\boldsymbol{\theta}}_n$. Furthermore, since $\hat{\boldsymbol{\alpha}}_n$ depends on $\hat{\boldsymbol{\tau}}_n = n\hat{\boldsymbol{\theta}}_n$ (compare Equations (4.5) and (4.7)), it is plausible that $\hat{\boldsymbol{\alpha}}_n$ converges slower than the estimators of the multiple change-point.
In general, the distributions and the location of change-points influence the quality of convergence of our estimators. To be more precise, we consider the estimator $\boldsymbol{\theta}_n^*$ for given $\boldsymbol{\theta} = (0.4, 0.8)$ and $\boldsymbol{\alpha} = (0, 1, -1)$ (variance of Q_1, Q_2 and Q_3 is 1). Then even for the sample size $n = 100$ the true multiple change-point is accurately estimated. When we choose change-points, which are closer to the boundary or closer together, the speed of convergence slightly decreases. In our examples in Table C.2 a sample size of more than 1000 is required to get an acceptable result. The same effect can be seen if the difference within expectations is small or the variances are increased.

Finally, we want to discuss the convergence in distribution of $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$ and $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$ based on the asymptotic confidence regions for $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ in the case of known and unknown expectations (see Theorems 3.29 and 4.40). The simulation of these asymptotic confidence regions was described in Sections 3.4 and 4.2.4. We use $m = \max\{\bar{\tau}_n, \bar{\sigma}_n - \bar{\tau}_n, n - \bar{\sigma}_n\}$ (case of known expectations) and $m = \max\{\tau_n^*, \sigma_n^* - \tau_n^*, n - \sigma_n^*\}$ (case of unknown expectations) to generate $N = 10^4$ processes Γ and their minimizers. To evaluate the convergence in distribution, we compute the approximated coverage probability on the basis of 10^3 intervals.

The same framework as for the performance of estimators is considered. The empirical results can be found in Table C.4. We can observe a similar pattern as before. Let us consider for example the case of unknown expectations. The speed of convergence of $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$ decreases if the change-points are closer to the boundary or closer together, the difference within expectations is small or the variances are increased. So, in our examples a sample size more than 5000 is required such that the coverage probability attains the given confidence level. On the contrary, in the first example $\boldsymbol{\theta} = (0.4, 0.8)$ and $\boldsymbol{\alpha} = (0, 1, -1)$ (variance of Q_1, Q_2 and Q_3 is 1) a sample size of more than 500 suffices. Moreover, in the case of known expectations it is interesting that the coverage probability attains the confidence level almost every time. This observation indicates that the location of change-points and distributions hardly influence the quality of convergence of $\bar{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$.

Chapter 6

Outlook

In this chapter we discuss the generalization of the previous results to an arbitrary, but known, number of change-points and give some ideas for further work on this field.

6.1 Generalization

The question naturally arises whether the results of our work can be generalized to an arbitrary, but known, number of change-points $q \in \mathbb{N}$. The detailed reply of this question may be content of further work. However, in this section we formulate conjectures according to the previous results and hint some problems of proofs. Since in practical applications it is common to have unknown expectations, we only focus on this case. For the convenience of the reader we use almost the same notation as in the previous chapters.

We begin with the formulation of the generalized multiple change-point model. Let $(X_{j,n})_{\substack{n\in\mathbb{N} \\ 1\leq j\leq n}}$ be a triangular array of random variables defined on a probability space

$(\Omega, \mathcal{A}, \mathbb{P})$ with values in the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Each row of the triangular array consists of independent random variables, i.e., $X_{1,n}, ..., X_{n,n}$ are independent for every $n \in \mathbb{N}$. Let us denote by Θ and Δ_n, $n \in \mathbb{N}$, the sets

$$\Theta := \{\mathbf{t} = (t_1, ..., t_q) \in \mathbb{R}^q | 0 < t_1 < ... < t_q < 1\} \qquad \text{and}$$
$$\Delta_n := \{\mathbf{k} = (k_1, ..., k_q) \in \mathbb{N}^q | 1 \leq k_1 < ... < k_q \leq n-1\}.$$

We assume that there exists a vector $\boldsymbol{\theta} := (\theta_1, ..., \theta_q) \in \Theta$ such that for all $n \in \mathbb{N}$

$$X_{i,n} \sim Q_r \quad \text{for} \quad \lfloor n\theta_{r-1} \rfloor + 1 \leq i \leq \lfloor n\theta_r \rfloor, \quad 1 \leq r \leq q+1,$$

where $Q_1, ..., Q_{q+1}$ are arbitrary, but unknown, distributions and $\theta_0 := 0, \theta_{q+1} := 1$. We call $\boldsymbol{\theta}$ *multiple change-point* and $\boldsymbol{\tau}_n := (\tau_{n,1}, ..., \tau_{n,q}) := (\lfloor n\theta_1 \rfloor, ..., \lfloor n\theta_q \rfloor) \in \Delta_n$, $n \in \mathbb{N}$, *moments of change*. Furthermore, we suppose that the expectations $\boldsymbol{\alpha} := (\alpha_1, ..., \alpha_{q+1})$ defined by $\alpha_r := \int_{\mathbb{R}} x \; Q_r(\mathrm{d}x)$, $1 \leq r \leq q+1$, exist, are finite and satisfy

$$\alpha_r \neq \alpha_{r+1}$$

for all $r \in \{1, ..., q\}$. The parameters $\boldsymbol{\theta}$ and $\boldsymbol{\alpha}$ are assumed to be unknown.
For simplicity of notation, we write $X_1, ..., X_n$ instead of $X_{1,n}, ..., X_{n,n}$, $n \in \mathbb{N}$, for the n-th row of the triangular array. Moreover, set

$$M_p := \max_{1 \leq r \leq q+1} \left\{ \int_{\mathbb{R}} |x|^p \; Q_r(\mathrm{d}x) \right\}$$

for the maximum of the p-th absolute moments, $p \in [1, \infty)$. Unless otherwise stated we assume that $M_1 < \infty$.

The same approach as in Chapter 4 should also yield results in our multiple change-point model. To simultaneously estimate the moments of change $\boldsymbol{\tau}_n = (\tau_{n,1}, ..., \tau_{n,q})$ and the expectations $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_{q+1})$ by the least squares method, we minimize the criterion function

$$S_n(\mathbf{k},\mathbf{a}) := \sum_{r=1}^{q+1} \sum_{i=k_{r-1}+1}^{k_r} (X_i - a_r)^2, \quad \mathbf{k} = (k_1, ..., k_q) \in \Delta_n, \quad \mathbf{a} = (a_1, ..., a_{q+1}) \in \mathbb{R}^{q+1},$$

where $k_0 := 0$ and $k_{q+1} := n$. To do this, set

$$\hat{M}_n(\mathbf{k}) := \sum_{r=1}^{q+1} (k_r - k_{r-1}) \bar{X}_{k_{r-1},k_r}, \quad \mathbf{k} = (k_1, ..., k_q) \in \Delta_n$$

and choose an arbitrary maximizing point

$$\hat{\boldsymbol{\tau}}_n := (\hat{\tau}_{n,1}, ..., \hat{\tau}_{n,q}) := \operatorname*{argmax}_{\mathbf{k}\in\Delta_n} \hat{M}_n(\mathbf{k}).$$

A generalization of Theorem 4.1 would bring the result that $(\hat{\boldsymbol{\tau}}_n, \hat{\boldsymbol{\alpha}}_n)$ with

$$\hat{\boldsymbol{\alpha}}_n := (\hat{\alpha}_{n,1}, ..., \hat{\alpha}_{n,q+1}) := \left(\bar{X}_{\hat{\tau}_{n,r},\hat{\tau}_{n,r+1}}\right)_{0\le r\le q}$$

is a minimizer of S_n for each $n \in \mathbb{N}$, where $\hat{\tau}_{n,0} := 0$ and $\hat{\tau}_{n,q+1} := n$. Hence $(\hat{\boldsymbol{\tau}}_n, \hat{\boldsymbol{\alpha}}_n)$ is a least squares estimator of $(\boldsymbol{\tau}_n, \boldsymbol{\alpha})$ and $\boldsymbol{\theta}$ can be estimated by $\hat{\boldsymbol{\theta}}_n := \frac{1}{n}\hat{\boldsymbol{\tau}}_n$.

To show consistency of $\hat{\boldsymbol{\theta}}_n$, we have to identify a limit process $\hat{\rho}$ (see (4.8) in the case $q = 2$) such that $\boldsymbol{\theta}$ is the well-separated maximizer of $\hat{\rho}$. The problem here is the partitioning of Θ into disjoint subsets according to the position of $\mathbf{t} = (t_1, ..., t_q) \in \Theta$ relative to the multiple change-point $\boldsymbol{\theta} = (\theta_1, ..., \theta_q) \in \Theta$. The number of subsets is $(q + 1)!$ (compare (3.11) and (3.12) in the case $q = 2$). Hence we cannot discuss all cases to obtain a limit process. Up to now, there is no self-contained representation of $\hat{\rho}$. If we are able to solve this problem, the proofs of weak and strong consistency of $\hat{\boldsymbol{\theta}}_n$ should be very similar to the work of Albrecht [1] and Section 4.1.2.

Conjecture 6.1. Suppose there is some $p \in (4,\infty)$ such that $M_p < \infty$. Then

$$\hat{\boldsymbol{\theta}}_n \xrightarrow[n\to\infty]{a.s.} \boldsymbol{\theta}.$$

Conjecture 6.2. Suppose there is some $p \in (2,\infty)$ such that $M_p < \infty$. Then

$$\hat{\boldsymbol{\theta}}_n \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\theta}.$$

The problem to get further results lies in the estimate of the error probability to show stochastic boundedness of $\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}_n$. As at the beginning of the proof of Lemma 4.16, we have to consider several cases by computing the maximum norm. For $q \geq 3$ there are lots of such cases which all must be handled. The solution of this problem would lead to the estimate of the error probability, and, in consequence, to the stochastic boundedness. Immediately, the proof of weak consistency of $\hat{\boldsymbol{\alpha}}_n$ (see Theorem 4.18) can be simply generalized by rules of convergence in probability.

Conjecture 6.3. If $M_4 < \infty$, then

$$\hat{\boldsymbol{\alpha}}_n \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\alpha}.$$

We now generalize the estimation approach from Section 4.2. The criterion function here is given by

$$S_n^*(\mathbf{k}) := \sum_{r=1}^{q+1} \sum_{i=k_{r-1}+1}^{k_r} \left(X_i - \hat{\alpha}_{n,r}\right)^2, \qquad \mathbf{k} = (k_1, ..., k_q) \in \Delta_n.$$

For abbreviation, set

$$a_{n,r}^*(X_i) := 2\left(\hat{\alpha}_{n,r+1} - \hat{\alpha}_{n,r}\right) X_i + \hat{\alpha}_{n,r}^2 - \hat{\alpha}_{n,r+1}^2$$

for $1 \leq i \leq n$ and $1 \leq r \leq q$ and

$$M_n^*(\mathbf{k}) := \sum_{r=1}^{q} \sum_{i=1}^{k_r} a_{n,r}^*(X_i), \qquad \mathbf{k} = (k_1, ..., k_q) \in \Delta_n.$$

By arguments applied in the proofs of Lemmas 4.21 and 4.20, we see that an arbitrary minimizer

$$\boldsymbol{\tau}_n^* := (\tau_{n,1}^*, ..., \tau_{n,q}^*) := \underset{\mathbf{k}\in\Delta_n}{\operatorname{argmin}}\, M_n^*(\mathbf{k})$$

of M_n^* minimizes S_n^* and $(\boldsymbol{\tau}_n^*, \hat{\boldsymbol{\alpha}}_n)$ is also a least squares estimator of $(\boldsymbol{\tau}_n, \boldsymbol{\alpha})$. The estimator of the multiple change-point $\boldsymbol{\theta}$ is given by $\boldsymbol{\theta}_n^* := \frac{1}{n}\boldsymbol{\tau}_n^*$.

To prove consistency of $\boldsymbol{\theta}_n^*$, we have to identify a limit variable ρ (see (3.13) in the case $q = 2$) such that $\boldsymbol{\theta}$ is the well-separated minimizer of ρ. The same problem occurs as described above (consistency of $\hat{\boldsymbol{\theta}}_n$). If we find a self-contained representation of ρ, we can proceed as in Section 4.2.2 to get weak consistency of $\boldsymbol{\theta}_n^*$.

Conjecture 6.4. If $M_4 < \infty$, then

$$\boldsymbol{\theta}_n^* \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\theta}.$$

The next step is to investigate convergence in distribution of $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$. To this end, consider the rescaled process

$$\Gamma_n^*(\mathbf{k}) := M_n^*(\boldsymbol{\tau}_n + \mathbf{k}) - M_n^*(\boldsymbol{\tau}_n), \qquad \mathbf{k} = (k_1, ..., k_q) \in H_n,$$

where

$$H_n := \{\mathbf{k} = (k_1, ..., k_q) \in \mathbb{Z}^q \big| k_r - k_{r-1} \geq 1 - (\tau_{n,r} - \tau_{n,r-1}) \text{ for all } r \in \{1, ..., q\}, \\ n - k_q \geq \tau_{n,q} + 1\}$$

and $\tau_{n,0} := 0$ and $\tau_{n,q+1} := n$. As in the proof of Lemma 4.31 follows that $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$ minimizes Γ_n^*, which would allow us to apply Theorem 2.3. We now generalize the process Γ given in (3.18) and Remark 3.16. For this purpose, let $(\xi_{i,r})_{i\in\mathbb{N}}$, $r \in \{1, ..., q+1\}$, be $q+1$ independent sequences, which for each r consist of independent and identically distributed random variables with common distribution Q_r. Let

$$\Gamma(\mathbf{k}) := \sum_{r=1}^{q} \Gamma_r(k_r), \qquad \mathbf{k} = (k_1, ..., k_q) \in \mathbb{Z}^q,$$

where

$$\Gamma_r(k_r) := \begin{cases} 2(\alpha_{r+1} - \alpha_r) \sum_{i=1}^{k_r} (\xi_{i,r+1} - \alpha_{r+1}) + k_r(\alpha_r - \alpha_{r+1})^2, & k_r \geq 0, \\ -2(\alpha_{r+1} - \alpha_r) \sum_{i=1}^{-k_r} (\xi_{i,r} - \alpha_r) - k(\alpha_r - \alpha_{r+1})^2, & k_r < 0. \end{cases}$$

To obtain convergence in distribution of $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$, we would apply Theorem 2.3. A very similar proceeding as in the Sections 4.2.3 and 3.3 establishes assumptions (i) and (iii) of Theorem 2.3. To conclude assumption (ii), the main difficulty also appears by computing the maximum norm in the error estimate to infer stochastic boundedness of $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$ (compare the beginning of the proof of Lemma 4.36), since there are many cases to treat. If we can solve this problem, $\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n$ converges in distribution to the minimizer of Γ, where Γ is a sum of q random walks.

Conjecture 6.5. If $M_4 < \infty$, then

$$\limsup_{n\to\infty} \mathbb{P}[\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n \in F] \leq \mathbb{P}[\mathrm{Argmin}(\Gamma) \cap F \neq \emptyset] \qquad \text{for all } F \subseteq \mathbb{Z}^q.$$

In addition, if $Q_1, ..., Q_{q+1}$ are continuous, then $\mathrm{Argmin}(\Gamma) = \{\boldsymbol{T}\}$ almost surely and

$$\boldsymbol{\tau}_n^* - \boldsymbol{\tau}_n \xrightarrow[n\to\infty]{\mathcal{L}} \boldsymbol{T} \quad \text{in} \quad \mathbb{Z}^q.$$

Let us denote by $F^{-1}_{\|\boldsymbol{T}\|}(\vartheta)$, $\vartheta \in (0,1)$, the ϑ-quantile of the distribution function $F_{\|\boldsymbol{T}\|}$ of $\|\boldsymbol{T}\|$. The same arguments used in the proof of Theorem 4.40 would lead to the following asymptotic confidence region for $\boldsymbol{\tau}_n = (\tau_{n,1}, ..., \tau_{n,q})$.

Conjecture 6.6. Suppose that $M_4 < \infty$. Let $Q_1, ..., Q_{q+1}$ be continuous distributions and $\vartheta \in (0,1)$. For each $n \in \mathbb{N}$, the random interval

$$I_n(\vartheta) := \bigtimes_{r=1}^{q} \left[\tau_{n,r}^* - F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta), \tau_{n,r}^* + F^{-1}_{\|\boldsymbol{T}\|}(1-\vartheta)\right]$$

is an asymptotic confidence region for $\boldsymbol{\tau}_n = (\tau_{n,1}, ..., \tau_{n,q})$ at level $1-\vartheta$.

6.2 Further research

This section gives a brief exposition of ideas for further research on this field.

The first aim of further research should be the generalization of our results to an arbitrary, but known, number of change-points, which was indicated in Section 6.1. Up to now, we have mainly focused on the estimation approach of the multiple change-point. However, one can examine strong consistency and convergence in distribution of $\hat{\boldsymbol{\alpha}}_n$. Thus the strong consistency of $\boldsymbol{\theta}_n^*$ can be proved and it is possible to construct an asymptotic confidence region for the expectations. Furthermore, it is desirable to derive a statistical test for the existence of $q \in \mathbb{N}$ change-points adjusted to our model.
Moreover, one may imagine a slight modification of our model, what is known as the so-called *diminishing disorders*. We assume that all distributions $Q_1, ..., Q_{q+1}$ depend on the sample size $n \in \mathbb{N}$ in the sense that the expectations of adjacent segments approach for growing $n \in \mathbb{N}$. Here, weak and strong consistency and convergence in distribution of all estimators can be also investigated.

Appendix A

Technical lemmas

In this chapter we compile some technical lemmas.

Lemma A.1. *The floor function has the following properties:*

(i) Let $x \in \mathbb{R}$. Then

$$x - 1 < \lfloor x \rfloor \leq x < \lfloor x \rfloor + 1.$$

(ii) Let $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$. If $y - x \geq n^{-1}$, then $\lfloor nx \rfloor < \lfloor ny \rfloor$.

(iii) Let $n \in \mathbb{N}$ and $A \subseteq \mathbb{R}$. Then

$$\sup_{x \in A} \left| \frac{\lfloor nx \rfloor}{n} - x \right| \leq \frac{1}{n}.$$

(iv) Let $y \in \mathbb{R}$ and $n \in \mathbb{N}$. Then

$$\left| \frac{\lfloor ny \rfloor}{n} - y \right| \leq \frac{1}{n}.$$

Proof. (i) The inequalities follow from the definition of the floor function.

(ii) Fix $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$. By (i) and $y - x \geq n^{-1}$, we get

$$\lfloor ny \rfloor - \lfloor nx \rfloor > ny - 1 - nx = n(y - x) - 1 \geq 0.$$

(iii) Fix $n \in \mathbb{N}$ and $A \subseteq \mathbb{R}$. Repeated application of (i) leads to

$$\sup_{x \in A} \left| \frac{\lfloor nx \rfloor}{n} - x \right| = \sup_{x \in A} \frac{nx - \lfloor nx \rfloor}{n} \leq \sup_{x \in A} \frac{\lfloor nx \rfloor + 1 - \lfloor nx \rfloor}{n} = \frac{1}{n}.$$

(iv) Fix $y \in \mathbb{R}$ and $n \in \mathbb{N}$. By (iii), we have

$$\left| \frac{\lfloor ny \rfloor}{n} - y \right| \leq \sup_{x \in \mathbb{R}} \left| \frac{\lfloor nx \rfloor}{n} - x \right| \leq \frac{1}{n}.$$

□

Lemma A.2. *It holds*

$$\tau_n \xrightarrow[n \to \infty]{} \infty, \qquad \sigma_n - \tau_n \xrightarrow[n \to \infty]{} \infty \qquad \text{and} \qquad n - \sigma_n \xrightarrow[n \to \infty]{} \infty.$$

Proof. As an example, we show that $\sigma_n - \tau_n \xrightarrow[n \to \infty]{} \infty$. The other claims follows analogously. By definitions $\tau_n = \lfloor n\theta_1 \rfloor$ and $\sigma_n = \lfloor n\theta_2 \rfloor$ and a simple application of Lemma A.1 (i), it is clear that

$$n(\theta_2 - \theta_1) - 1 \leq \sigma_n - \tau_n \leq n(\theta_2 - \theta_1) + 1$$

for $n \in \mathbb{N}$. Letting $n \to \infty$ completes the proof. □

Lemma A.3. *Let* $n \in \mathbb{N}$. *Then*

$$\Delta_n = \left\{ (\lfloor ns \rfloor, \lfloor nt \rfloor) \in \mathbb{N}^2 \middle| (s, t) \in \Theta_n \right\}.$$

Proof. Fix $n \in \mathbb{N}$. We first recall

$$\Delta_n = \left\{ (k, l) \in \mathbb{N}^2 \middle| 1 \leq k < l \leq n - 1 \right\} \qquad \text{and}$$
$$\Theta_n = \left\{ (s, t) \in \Theta \middle| s \geq \frac{1}{n}, t - s \geq \frac{1}{n}, 1 - t \geq \frac{1}{n} \right\}.$$

1. We first show that $\Delta_n \subseteq \left\{(\lfloor ns \rfloor, \lfloor nt \rfloor) \in \mathbb{N}^2 \middle| (s,t) \in \Theta_n\right\}$.

   Let $(k,l) \in \Delta_n$ and define $(s,t) := \left(\frac{k}{n}, \frac{l}{n}\right)$. Then we have $(\lfloor ns \rfloor, \lfloor nt \rfloor) = (k,l) \in \mathbb{N}^2$ and $(s,t) \in \Theta_n$.

2. We next prove that $\left\{(\lfloor ns \rfloor, \lfloor nt \rfloor) \in \mathbb{N}^2 \middle| (s,t) \in \Theta_n\right\} \subseteq \Delta_n$.

   Let $(\lfloor ns \rfloor, \lfloor nt \rfloor) \in \left\{(\lfloor ns \rfloor, \lfloor nt \rfloor) \in \mathbb{N}^2 \middle| (s,t) \in \Theta_n\right\}$. By $(s,t) \in \Theta_n$, we observe that $ns \geq 1$. On the contrary, suppose that $\lfloor ns \rfloor < 1$. Accordingly, $ns < 1$, which is a contradiction. Hence $\lfloor ns \rfloor \geq 1$. Moreover, since $(s,t) \in \Theta_n$, we see that $nt \leq n-1$. By an application of Lemma A.1 (i), we infer that $\lfloor nt \rfloor \leq n-1$. It still remains to show that $\lfloor ns \rfloor < \lfloor nt \rfloor$. By another application of Lemma A.1 (i), we obtain

   $$\lfloor nt \rfloor - \lfloor ns \rfloor > nt - 1 - ns = n(t-s) - 1 \geq 0,$$

   where the last inequality is deduced from $(s,t) \in \Theta_n$. □

Lemma A.4. *Let $Z_1, ..., Z_n$, $n \in \mathbb{N}$, be random variables and $\epsilon > 0$. Then*

$$\left\{\sum_{i=1}^{n} Z_i > \epsilon\right\} \subseteq \bigcup_{i=1}^{n} \left\{Z_i > \tfrac{\epsilon}{n}\right\}.$$

Proof. Fix $n \in \mathbb{N}$ and $\varepsilon > 0$. To obtain a contradiction, suppose that there exists $\omega \in \{\sum_{i=1}^{n} Z_i > \epsilon\}$, but $\omega \notin \bigcup_{i=1}^{n} \left\{Z_i > \frac{\epsilon}{n}\right\}$. Hence $Z_i(\omega) \leq \frac{\epsilon}{n}$ for all $i \in \{1, ..., n\}$. We thus get

$$\sum_{i=1}^{n} Z_i(\omega) \leq \sum_{i=1}^{n} \tfrac{\epsilon}{n} = \epsilon,$$

which contradicts our assumption. □

Lemma A.5. *Let $r \in [1, \infty)$ and $u, v \in \mathbb{N}_2$ with $u < v$. Then*

$$\sum_{m=u}^{v} m^{-r} \leq \begin{cases} 1 + \ln(v), & r = 1, \\ \frac{1}{r-1}(u-1)^{1-r}, & r \in (1, \infty). \end{cases}$$

Proof. Fix $u, v \in \mathbb{N}_2$ with $u < v$. For $r = 1$ we conclude that

$$\sum_{m=u}^{v} m^{-r} \leq \sum_{m=1}^{v} m^{-r} = 1 + \sum_{m=2}^{v} m^{-r} \leq 1 + \int_{1}^{v} x^{-r} \,\mathrm{d}x = 1 + \ln(v).$$

For $r \in (1, \infty)$ we have

$$\sum_{m=u}^{v} m^{-r} \leq \int_{u-1}^{v} x^{-r} \,\mathrm{d}x \leq \int_{u-1}^{\infty} x^{-r} \,\mathrm{d}x = \frac{1}{r-1}(u-1)^{1-r}. \qquad \square$$

Lemma A.6. *Let $m, r \in (0, \infty)$. Then*

$$m^{-r} - (m+1)^{-r} \leq r m^{-(r+1)}.$$

Proof. Fix $m, r \in (0, \infty)$. Define the continuous and monotonically decreasing mapping $f : \mathbb{R}_{>0} \to \mathbb{R}$ by $f(x) := x^{-r}$. By the Mean Value Theorem (see for instance Heuser [20, p. 279, 49.1]), there exists $\eta \in (m, m+1)$ such that

$$m^{-r} - (m+1)^{-r} = f(m) - f(m+1) = \left| \frac{f(m) - f(m+1)}{m - (m+1)} \right| = |f'(\eta)| = r\eta^{-(r+1)} \leq r m^{-(r+1)}.$$

The last inequality follows from $\eta > m$. $\square$

Lemma A.7. *Let $n \in \mathbb{N}$. For any $b_1, ..., b_n \in (0, \infty)$ it holds*

$$\left(\sum_{i=1}^{n} b_i \right)^{1/2} \leq \sum_{i=1}^{n} b_i^{1/2}.$$

Proof. The proof is by induction on $n \in \mathbb{N}$. Let $n = 2$. Then the assertion is equivalent to $b_1 + b_2 \leq b_1 + b_2 + 2b_1^{1/2} b_2^{1/2}$, which is obviously fulfilled. We now suppose the induction hypothesis $\left(\sum_{i=1}^{n} b_i\right)^{1/2} \leq \sum_{i=1}^{n} b_i^{1/2}$. It follows that

$$\left(\sum_{i=1}^{n+1} b_i \right)^{1/2} = \left(\sum_{i=1}^{n} b_i + b_{n+1} \right)^{1/2} \leq \left(\sum_{i=1}^{n} b_i \right)^{1/2} + b_{n+1}^{1/2} \leq \sum_{i=1}^{n} b_i^{1/2} + b_{n+1}^{1/2} = \sum_{i=1}^{n+1} b_i^{1/2}. \ \square$$

Lemma A.8. *Let Y and Z be independent random variables. Then*

$$\mathbb{P}\left[Y + Z \geq 0\right] = \int_{\mathbb{R}} \mathbb{P}\left[Y \geq -z\right] \ \mathbb{P}_Z(\mathrm{d}z).$$

Proof. By the tower property of the conditional expectation, we conclude that

$$\mathbb{P}[Y + Z \geq 0] = \mathbb{E}\left[\mathbb{1}_{\{Y+Z\geq 0\}}\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\{Y+Z\geq 0\}} \middle| Z\right]\right].$$

Define the function $g : \mathbb{R} \to \mathbb{R}$ by $g(z) := \mathbb{E}\left[\mathbb{1}_{\{Y+Z\geq 0\}} | Z = z\right]$. It follows that

$$\mathbb{P}[Y + Z \geq 0] = \mathbb{E}[g(Z)] = \int_{\mathbb{R}} g(z)\ \mathbb{P}_Z(\mathrm{d}z). \tag{A.1}$$

The independence of Y and Z leads to

$$g(z) = \mathbb{E}\left[\mathbb{1}_{\{Y+z\geq 0\}}\right] = \mathbb{P}[Y \geq -z]$$

for all $z \in \mathbb{R}$. Applying (A.1) establishes the claim. □

Lemma A.9. *Let Y and Z be independent random variables with $Y \geq 0$ and $Z \geq 0$ almost surely. Let $\varepsilon > 0$. Then*

$$\mathbb{P}[YZ \geq \varepsilon] = \int_{(0,\infty)} \mathbb{P}\left[Z \geq \varepsilon y^{-1}\right]\ \mathbb{P}_Y(\mathrm{d}y).$$

Proof. Similar arguments applied in the proof of Lemma A.8 gives the assertion. □

Appendix B

Some functions

This chapter contains some criterion functions with their decompositions. To shorten notations, we use for $u, v \in \mathbb{N}_0$ with $u < v$ the abbreviation

$$S_{u,v} := \sum_{i=u+1}^{v} (X_i - \mathbb{E}[X_i]).$$

Lemma B.1. *Let $n \in \mathbb{N}$ and $(s,t) \in \Theta_n$. Then*

$$\hat{\rho}_n(s,t) = \hat{\delta}_n(s,t) + \hat{\varrho}_n(s,t),$$

where

$$\begin{aligned}
\hat{\delta}_n(s,t)|_{\Theta^1\cap\Theta_n} &= \frac{1}{n\lfloor ns\rfloor} S^2_{0,\lfloor ns\rfloor} + \frac{1}{n(\lfloor nt\rfloor - \lfloor ns\rfloor)} S^2_{\lfloor ns\rfloor,\lfloor nt\rfloor} + \frac{2}{n}\alpha S_{0,\lfloor nt\rfloor} \\
&\quad + \frac{1}{n(n-\lfloor nt\rfloor)} S^2_{\lfloor nt\rfloor,n} \\
&\quad + \frac{2((\tau_n - \lfloor nt\rfloor)\alpha + (\sigma_n - \tau_n)\beta + (n-\sigma_n)\gamma)}{n(n-\lfloor nt\rfloor)} S_{\lfloor nt\rfloor,n}, \\
\hat{\delta}_n(s,t)|_{\Theta^2\cap\Theta_n} &= \frac{1}{n\lfloor ns\rfloor} S^2_{0,\lfloor ns\rfloor} + \frac{2}{n}\alpha S_{0,\lfloor ns\rfloor} + \frac{1}{n(\lfloor nt\rfloor - \lfloor ns\rfloor)} S^2_{\lfloor ns\rfloor,\lfloor nt\rfloor} \\
&\quad + \frac{2((\tau_n - \lfloor ns\rfloor)\alpha + (\lfloor nt\rfloor - \tau_n)\beta)}{n(\lfloor nt\rfloor - \lfloor ns\rfloor)} S_{\lfloor ns\rfloor,\lfloor nt\rfloor} + \frac{1}{n(n-\lfloor nt\rfloor)} S^2_{\lfloor nt\rfloor,n} \\
&\quad + \frac{2((\sigma_n - \lfloor nt\rfloor)\beta + (n-\sigma_n)\gamma)}{n(n-\lfloor nt\rfloor)} S_{\lfloor nt\rfloor,n},
\end{aligned}$$

$$
\begin{aligned}
\hat{\delta}_n(s,t)|_{\Theta^3\cap\Theta_n} &= \frac{1}{n\lfloor ns\rfloor}S^2_{0,\lfloor ns\rfloor} + \frac{2}{n}\alpha S_{0,\lfloor ns\rfloor} + \frac{1}{n(\lfloor nt\rfloor - \lfloor ns\rfloor)}S^2_{\lfloor ns\rfloor,\lfloor nt\rfloor} \\
&\quad + \frac{2((\tau_n - \lfloor ns\rfloor)\alpha + (\sigma_n - \tau_n)\beta + (\lfloor nt\rfloor - \sigma_n)\gamma)}{n(\lfloor nt\rfloor - \lfloor ns\rfloor)}S_{\lfloor ns\rfloor,\lfloor nt\rfloor} \\
&\quad + \frac{1}{n(n-\lfloor nt\rfloor)}S^2_{\lfloor nt\rfloor,n} + \frac{2}{n}\gamma S_{\lfloor nt\rfloor,n}, \\
\hat{\delta}_n(s,t)|_{\Theta^4\cap\Theta_n} &= \frac{1}{n\lfloor ns\rfloor}S^2_{0,\lfloor ns\rfloor} + \frac{2(\tau_n\alpha + (\lfloor ns\rfloor - \tau_n)\beta)}{n\lfloor ns\rfloor}S_{0,\lfloor ns\rfloor} \\
&\quad + \frac{1}{n(\lfloor nt\rfloor - \lfloor ns\rfloor)}S^2_{\lfloor ns\rfloor,\lfloor nt\rfloor} + \frac{2}{n}\beta S_{\lfloor ns\rfloor,\lfloor nt\rfloor} + \frac{1}{n(n-\lfloor nt\rfloor)}S^2_{\lfloor nt\rfloor,n} \\
&\quad + \frac{2((\sigma_n - \lfloor nt\rfloor)\beta + (n-\sigma_n)\gamma)}{n(n-\lfloor nt\rfloor)}S_{\lfloor nt\rfloor,n}, \\
\hat{\delta}_n(s,t)|_{\Theta^5\cap\Theta_n} &= \frac{1}{n\lfloor ns\rfloor}S^2_{0,\lfloor ns\rfloor} + \frac{2(\tau_n\alpha + (\lfloor ns\rfloor - \tau_n)\beta)}{n\lfloor ns\rfloor}S_{0,\lfloor ns\rfloor} \\
&\quad + \frac{1}{n(\lfloor nt\rfloor - \lfloor ns\rfloor)}S^2_{\lfloor ns\rfloor,\lfloor nt\rfloor} + \frac{2((\sigma_n - \lfloor ns\rfloor)\beta + (\lfloor nt\rfloor - \sigma_n)\gamma)}{n(\lfloor nt\rfloor - \lfloor ns\rfloor)}S_{\lfloor ns\rfloor,\lfloor nt\rfloor} \\
&\quad + \frac{1}{n(n-\lfloor nt\rfloor)}S^2_{\lfloor nt\rfloor,n} + \frac{2}{n}\gamma S_{\lfloor nt\rfloor,n}, \\
\hat{\delta}_n(s,t)|_{\Theta^6\cap\Theta_n} &= \frac{1}{n\lfloor ns\rfloor}S^2_{0,\lfloor ns\rfloor} + \frac{2(\tau_n\alpha + (\sigma_n - \tau_n)\beta + (\lfloor ns\rfloor - \sigma_n)\gamma)}{n\lfloor ns\rfloor}S_{0,\lfloor ns\rfloor} \\
&\quad + \frac{1}{n(\lfloor nt\rfloor - \lfloor ns\rfloor)}S^2_{\lfloor ns\rfloor,\lfloor nt\rfloor} + \frac{2}{n}\gamma S_{\lfloor ns\rfloor,\lfloor nt\rfloor} \\
&\quad + \frac{1}{n(n-\lfloor nt\rfloor)}S^2_{\lfloor nt\rfloor,n} + \frac{2}{n}\gamma S_{\lfloor nt\rfloor,n}
\end{aligned}
$$

and

$$
\begin{aligned}
\hat{\varrho}_n(s,t)|_{\Theta^1\cap\Theta_n} &= \frac{\lfloor nt\rfloor}{n}\alpha^2 + \frac{n-\lfloor nt\rfloor}{n}\left(\frac{\tau_n - \lfloor nt\rfloor}{n-\lfloor nt\rfloor}\alpha + \frac{\sigma_n - \tau_n}{n-\lfloor nt\rfloor}\beta + \frac{n-\sigma_n}{n-\lfloor nt\rfloor}\gamma\right)^2, \\
\hat{\varrho}_n(s,t)|_{\Theta^2\cap\Theta_n} &= \frac{\lfloor ns\rfloor}{n}\alpha^2 + \frac{\lfloor nt\rfloor - \lfloor ns\rfloor}{n}\left(\frac{\tau_n - \lfloor ns\rfloor}{\lfloor nt\rfloor - \lfloor ns\rfloor}\alpha + \frac{\lfloor nt\rfloor - \tau_n}{\lfloor nt\rfloor - \lfloor ns\rfloor}\beta\right)^2 \\
&\quad + \frac{n-\lfloor nt\rfloor}{n}\left(\frac{\sigma_n - \lfloor nt\rfloor}{n-\lfloor nt\rfloor}\beta + \frac{n-\sigma_n}{n-\lfloor nt\rfloor}\gamma\right)^2,
\end{aligned}
$$

$$
\begin{aligned}
\hat{\varrho}_n(s,t)|_{\Theta^3\cap\Theta_n} &= \frac{\lfloor ns\rfloor}{n}\alpha^2 \\
&\quad + \frac{\lfloor nt\rfloor - \lfloor ns\rfloor}{n}\left(\frac{\tau_n - \lfloor ns\rfloor}{\lfloor nt\rfloor - \lfloor ns\rfloor}\alpha + \frac{\sigma_n - \tau_n}{\lfloor nt\rfloor - \lfloor ns\rfloor}\beta + \frac{\lfloor nt\rfloor - \sigma_n}{\lfloor nt\rfloor - \lfloor ns\rfloor}\gamma\right)^2 \\
&\quad + \frac{n - \lfloor nt\rfloor}{n}\gamma^2, \\
\hat{\varrho}_n(s,t)|_{\Theta^4\cap\Theta_n} &= \frac{\lfloor ns\rfloor}{n}\left(\frac{\tau_n}{\lfloor ns\rfloor}\alpha + \frac{\lfloor ns\rfloor - \tau_n}{\lfloor ns\rfloor}\beta\right)^2 + \frac{\lfloor nt\rfloor - \lfloor ns\rfloor}{n}\beta^2 \\
&\quad + \frac{n - \lfloor nt\rfloor}{n}\left(\frac{\sigma_n - \lfloor nt\rfloor}{n - \lfloor nt\rfloor}\beta + \frac{n - \sigma_n}{n - \lfloor nt\rfloor}\gamma\right)^2, \\
\hat{\varrho}_n(s,t)|_{\Theta^5\cap\Theta_n} &= \frac{\lfloor ns\rfloor}{n}\left(\frac{\tau_n}{\lfloor ns\rfloor}\alpha + \frac{\lfloor ns\rfloor - \tau_n}{\lfloor ns\rfloor}\beta\right)^2 \\
&\quad + \frac{\lfloor nt\rfloor - \lfloor ns\rfloor}{n}\left(\frac{\sigma_n - \lfloor ns\rfloor}{\lfloor nt\rfloor - \lfloor ns\rfloor}\beta + \frac{\lfloor nt\rfloor - \sigma_n}{\lfloor nt\rfloor - \lfloor ns\rfloor}\gamma\right)^2 + \frac{n - \lfloor nt\rfloor}{n}\gamma^2, \\
\hat{\varrho}_n(s,t)|_{\Theta^6\cap\Theta_n} &= \frac{\lfloor ns\rfloor}{n}\left(\frac{\tau_n}{\lfloor ns\rfloor}\alpha + \frac{\sigma_n - \tau_n}{\lfloor ns\rfloor}\beta + \frac{\lfloor ns\rfloor - \sigma_n}{\lfloor ns\rfloor}\gamma\right)^2 + \frac{n - \lfloor ns\rfloor}{n}\gamma^2.
\end{aligned}
$$

Proof. The proof can be found in Albrecht [1, p. 24, Lemma 3.2]. □

Lemma B.2. *Let $n \in \mathbb{N}$ and $(s,t) \in \Theta_n$. Then*

$$
\rho_n^*(s,t) = \delta_n^*(s,t) + \varrho_n^*(s,t),
$$

where

$$
\begin{aligned}
\delta_n^*(s,t)|_{\Theta^1\cap\Theta_n} &= \frac{2}{n}\left(\hat{\beta}_n - \hat{\alpha}_n\right)S_{0,\lfloor ns\rfloor} + \frac{2}{n}\left(\hat{\gamma}_n - \hat{\beta}_n\right)S_{0,\lfloor nt\rfloor}, \\
\delta_n^*(s,t)|_{\Theta^2\cap\Theta_n} &= \frac{2}{n}\left(\hat{\beta}_n - \hat{\alpha}_n\right)S_{0,\lfloor ns\rfloor} + \frac{2}{n}\left(\hat{\gamma}_n - \hat{\beta}_n\right)S_{0,\tau_n} + \frac{2}{n}\left(\hat{\gamma}_n - \hat{\beta}_n\right)S_{\tau_n,\lfloor nt\rfloor}, \\
\delta_n^*(s,t)|_{\Theta^3\cap\Theta_n} &= \frac{2}{n}\left(\hat{\beta}_n - \hat{\alpha}_n\right)S_{0,\lfloor ns\rfloor} + \frac{2}{n}\left(\hat{\gamma}_n - \hat{\beta}_n\right)S_{0,\tau_n} + \frac{2}{n}\left(\hat{\gamma}_n - \hat{\beta}_n\right)S_{\tau_n,\sigma_n} \\
&\quad + \frac{2}{n}\left(\hat{\gamma}_n - \hat{\beta}_n\right)S_{\sigma_n,\lfloor nt\rfloor}, \\
\delta_n^*(s,t)|_{\Theta^4\cap\Theta_n} &= \frac{2}{n}\left(\hat{\gamma}_n - \hat{\alpha}_n\right)S_{0,\tau_n} + \frac{2}{n}\left(\hat{\beta}_n - \hat{\alpha}_n\right)S_{\tau_n,\lfloor ns\rfloor} + \frac{2}{n}\left(\hat{\gamma}_n - \hat{\beta}_n\right)S_{\tau_n,\lfloor nt\rfloor},
\end{aligned}
$$

$$
\begin{aligned}
\delta_n^*(s,t)|_{\Theta^5\cap\Theta_n} &= \frac{2}{n}\left(\hat{\gamma}_n-\hat{\alpha}_n\right)S_{0,\tau_n}+\frac{2}{n}\left(\hat{\beta}_n-\hat{\alpha}_n\right)S_{\tau_n,\lfloor ns\rfloor}+\frac{2}{n}\left(\hat{\gamma}_n-\hat{\beta}_n\right)S_{\tau_n,\sigma_n}\\
&\quad+\frac{2}{n}\left(\hat{\gamma}_n-\hat{\beta}_n\right)S_{\sigma_n,\lfloor nt\rfloor},\\
\delta_n^*(s,t)|_{\Theta^6\cap\Theta_n} &= \frac{2}{n}\left(\hat{\gamma}_n-\hat{\alpha}_n\right)S_{0,\tau_n}+\frac{2}{n}\left(\hat{\gamma}_n-\hat{\alpha}_n\right)S_{\tau_n,\sigma_n}+\frac{2}{n}\left(\hat{\beta}_n-\hat{\alpha}_n\right)S_{\sigma_n,\lfloor ns\rfloor}\\
&\quad+\frac{2}{n}\left(\hat{\gamma}_n-\hat{\beta}_n\right)S_{\sigma_n,\lfloor nt\rfloor}
\end{aligned}
$$

and

$$
\begin{aligned}
\varrho_n^*(s,t)|_{\Theta^1\cap\Theta_n} &= \left(2\alpha\left(\hat{\beta}_n-\hat{\alpha}_n\right)+\hat{\alpha}_n^2-\hat{\beta}_n^2\right)\frac{\lfloor ns\rfloor}{n}+\left(2\alpha\left(\hat{\gamma}_n-\hat{\beta}_n\right)+\hat{\beta}_n^2-\hat{\gamma}_n^2\right)\frac{\lfloor nt\rfloor}{n},\\
\varrho_n^*(s,t)|_{\Theta^2\cap\Theta_n} &= \left(2\alpha\left(\hat{\beta}_n-\hat{\alpha}_n\right)+\hat{\alpha}_n^2-\hat{\beta}_n^2\right)\frac{\lfloor ns\rfloor}{n}+2(\alpha-\beta)\left(\hat{\gamma}_n-\hat{\beta}_n\right)\frac{\tau_n}{n}\\
&\quad+\left(2\beta\left(\hat{\gamma}_n-\hat{\beta}_n\right)+\hat{\beta}_n^2-\hat{\gamma}_n^2\right)\frac{\lfloor nt\rfloor}{n},\\
\varrho_n^*(s,t)|_{\Theta^3\cap\Theta_n} &= \left(2\alpha\left(\hat{\beta}_n-\hat{\alpha}_n\right)+\hat{\alpha}_n^2-\hat{\beta}_n^2\right)\frac{\lfloor ns\rfloor}{n}+2(\alpha-\beta)\left(\hat{\gamma}_n-\hat{\beta}_n\right)\frac{\tau_n}{n}\\
&\quad+2(\beta-\gamma)\left(\hat{\gamma}_n-\hat{\beta}_n\right)\frac{\sigma_n}{n}+\left(2\gamma\left(\hat{\gamma}_n-\hat{\beta}_n\right)+\hat{\beta}_n^2-\hat{\gamma}_n^2\right)\frac{\lfloor nt\rfloor}{n},\\
\varrho_n^*(s,t)|_{\Theta^4\cap\Theta_n} &= 2(\alpha-\beta)\left(\hat{\gamma}_n-\hat{\alpha}_n\right)\frac{\tau_n}{n}+\left(2\beta\left(\hat{\beta}_n-\hat{\alpha}_n\right)+\hat{\alpha}_n^2-\hat{\beta}_n^2\right)\frac{\lfloor ns\rfloor}{n}\\
&\quad+\left(2\beta\left(\hat{\gamma}_n-\hat{\beta}_n\right)+\hat{\beta}_n^2-\hat{\gamma}_n^2\right)\frac{\lfloor nt\rfloor}{n},\\
\varrho_n^*(s,t)|_{\Theta^5\cap\Theta_n} &= 2(\alpha-\beta)\left(\hat{\gamma}_n-\hat{\alpha}_n\right)\frac{\tau_n}{n}+\left(2\beta\left(\hat{\beta}_n-\hat{\alpha}_n\right)+\hat{\alpha}_n^2-\hat{\beta}_n^2\right)\frac{\lfloor ns\rfloor}{n}\\
&\quad+2(\beta-\gamma)\left(\hat{\gamma}_n-\hat{\beta}_n\right)\frac{\sigma_n}{n}+\left(2\gamma\left(\hat{\gamma}_n-\hat{\beta}_n\right)+\hat{\beta}_n^2-\hat{\gamma}_n^2\right)\frac{\lfloor nt\rfloor}{n},\\
\varrho_n^*(s,t)|_{\Theta^6\cap\Theta_n} &= 2(\alpha-\beta)\left(\hat{\gamma}_n-\hat{\alpha}_n\right)\frac{\tau_n}{n}+2(\beta-\gamma)\left(\hat{\gamma}_n-\hat{\alpha}_n\right)\frac{\sigma_n}{n}\\
&\quad+\left(2\gamma\left(\hat{\beta}_n-\hat{\alpha}_n\right)+\hat{\alpha}_n^2-\hat{\beta}_n^2\right)\frac{\lfloor ns\rfloor}{n}+\left(2\gamma\left(\hat{\gamma}_n-\hat{\beta}_n\right)+\hat{\beta}_n^2-\hat{\gamma}_n^2\right)\frac{\lfloor nt\rfloor}{n}.
\end{aligned}
$$

Proof. The proof is given in Lemma 4.26 only for the case $(s,t)\in\Theta^2\cap\Theta_n$, $n\in\mathbb{N}$; the other cases follows the same pattern. □

Appendix C

Results of a simulation study

This chapter provides the results of a simulation study in the software environment R, version 3.3.1. An explanation of each table can be found in Chapter 5.

Distributions			$\bar{M}_n$ has exactly	$\hat{M}_n$ has exactly	M_n^* has only the
Q_1	Q_2	Q_3	one minimizer	one maximizer	minimizer $\hat{\boldsymbol{\tau}}_n$
$B(5, 0.2)$	$B(4, 0.5)$	$B(5, 0.2)$	721418	928041	999737
$Poi(1)$	$Poi(2)$	$Poi(1)$	743371	940360	999801
$Poi(1)$	$B(4, 0.5)$	$B(5, 0.2)$	730525	930016	999789
$N(1, 1)$	$N(2, 1)$	$N(1, 1)$	1000000	1000000	1000000
$Exp(1)$	$Exp(0.5)$	$Exp(1)$	1000000	1000000	1000000
$Exp(1)$	$N(2, 1)$	$Exp(1)$	1000000	1000000	1000000
$Poi(1)$	$Poi(2)$	$N(1, 1)$	784575	985335	1000000
$Exp(1)$	$B(4, 0.5)$	$Exp(1)$	866390	1000000	1000000
$Poi(1)$	$Exp(0.5)$	$B(5, 0.2)$	944946	986814	1000000

Table C.1: Minimizers and maximizers of $\bar{M}_n$, $\hat{M}_n$ and M_n^* based on 10^6 Monte-Carlo repetitions with $n = 10$ and $\boldsymbol{\theta} = (0.4, 0.8)$, respectively.

n	$\bar{\boldsymbol{\theta}}_n$ Bias	$\bar{\boldsymbol{\theta}}_n$ RMSE	$\boldsymbol{\theta}_n^*$ Bias	$\boldsymbol{\theta}_n^*$ RMSE
	$Q_1 = N(0,1)$, $Q_2 = N(1,1)$, $Q_3 = N(-1,1)$, $\boldsymbol{\theta} = (0.4, 0.8)$			
100	(-0.00023,0.00006)	(0.05159,0.01212)	(0.00220,-0.00044)	(0.07237,0.015780)
500	(0.00010,0.00002)	(0.01020,0.00242)	(-0.00007,0.00001)	(0.01096,0.00249)
1000	(0.00002,0.00000)	(0.00504,0.00124)	(0.00003,0.00000)	(0.00518,0.00126)
5000	(0.00000,0.00000)	(0.00104,0.00025)	(0.00000,0.00000)	(0.00105,0.00024)
	$Q_1 = N(0,1)$, $Q_2 = N(1,1)$, $Q_3 = N(-1,1)$, $\boldsymbol{\theta} = (0.02, 0.995)$			
100	(0.01409,-0.00849)	(0.03747,0.01211)	(0.34715,-0.28674)	(0.47619,0.42556)
500	(0.00081,-0.00118)	(0.00870,0.00241)	(0.09855,-0.09769)	(0.26300, 0.26904)
1000	(0.00014,-0.00003)	(0.00476,0.00116)	(0.01293,-0.02102)	(0.09249,0.13102)
5000	(-0.00001,-0.00001)	(0.00101,0.00025)	(0.00003,-0.00002)	(0.00108,0.00028)
	$Q_1 = N(0,1)$, $Q_2 = N(1,1)$, $Q_3 = N(-1,1)$, $\boldsymbol{\theta} = (0.4, 0.41)$			
100	(-0.02767,-0.00507)	(0.07082,0.04941)	(-0.06516,0.15034)	(0.16195,0.25758)
500	(-0.00121,0.00055)	(0.00969,0.00490)	(-0.04467,0.07796)	(0.12205,0.18953)
1000	(-0.00019,0.00013)	(0.00466,0.00164)	(-0.02124,0.03437)	(0.08279,0.12793)
5000	(0.00000,0.00000)	(0.00102,0.00024)	(-0.00010,0.00000)	(0.00140,0.00026)
	$Q_1 = N(0,1)$, $Q_2 = N(0.2,1)$, $Q_3 = N(-0.1,1)$, $\boldsymbol{\theta} = (0.4, 0.8)$			
100	(-0.05802,-0.04673)	(0.21302,0.19812)	(0.04664,-0.21390)	(0.27874,0.34980)
500	(-0.00434,-0.00207)	(0.15000,0.08254)	(0.09646,-0.14945)	(0.27686,0.29455)
1000	(0.00020,0.00121)	(0.10799 0.05214)	(0.08323,-0.07960)	(0.23571,0.21369)
5000	(-0.00008,0.00008)	(0.02485,0.01135)	(0.00093,-0.00046)	(0.03344,0.01300)
	$Q_1 = N(0,1)$, $Q_2 = N(1,9)$, $Q_3 = N(-1,4)$, $\boldsymbol{\theta} = (0.4, 0.8)$			
100	(0.07637,-0.01904)	(0.13208,0.06757)	(0.16493,-0.08421)	(0.21121,0.13958)
500	(0.03235,-0.00532)	(0.06525,0.01824)	(0.06450,-0.01148)	(0.12059,0.03361)
1000	(0.01715,-0.00284)	(0.03557,0.00941)	(0.02321,-0.00373)	(0.05109,0.01151)
5000	(0.00344,-0.00055)	(0.00726,0.00184)	(0.00362,-0.00059)	(0.00783,0.00191)

Table C.2: Bias and RMSE of the estimators for $\boldsymbol{\theta}$ based on 10^4 Monte-Carlo repetitions.

n	$\hat{\boldsymbol{\alpha}}_n = \left(\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n\right)$ Bias	RMSE
	$Q_1 = N(0,1)$, $Q_2 = N(1,1)$, $Q_3 = N(-1,1)$, $\boldsymbol{\theta} = (0.4, 0.8)$	
100	(-0.03245,0.03902,-0.00995)	(0.21268,0.26622,0.23837)
500	(-0.00412,0.00501,-0.00323)	(0.07183,0.07232,0.10094)
1000	(-0.00255,0.00260,-0.00100)	(0.04994,0.05055,0.07091)
5000	(-0.00057,0.00077,-0.00019)	(0.02252,0.02258,0.03171)
	$Q_1 = N(0,1)$, $Q_2 = N(1,1)$, $Q_3 = N(-1,1)$, $\boldsymbol{\theta} = (0.02, 0.995)$	
100	(0.41605,0.28725,0.86562)	(0.97340,1.57808,1.60021)
500	(0.01621,0.04968,0.17739)	(0.59334,0.85231,0.99755)
1000	(-0.03681,0.00483,-0.00685)	(0.32479,0.34236,0.61582)
5000	(-0.00892,0.00067,-0.00879)	(0.10336,0.01424,0.20663)
	$Q_1 = N(0,1)$, $Q_2 = N(1,1)$, $Q_3 = N(-1,1)$, $\boldsymbol{\theta} = (0.4, 0.41)$	
100	(-0.02554,-1.32110,0.02672)	(0.58107,2.07587,0.58404)
500	(-0.01511,-0.55697,-0.00485)	(0.36609,1.51170,0.37467)
1000	(-0.01050,-0.16806,0.00248)	(0.24122,1.04335,0.24829)
5000	(-0.00075,0.02567,-0.00021)	(0.02228,0.15521,0.01818)
	$Q_1 = N(0,1)$, $Q_2 = N(0.2,1)$, $Q_3 = N(-0.1,1)$, $\boldsymbol{\theta} = (0.4, 0.8)$	
100	(0.00465,0.02073,0.04219)	(0.50418,1.88874,0.51612)
500	(0.01036,0.04757,0.03723)	(0.29624,1.74360,0.32200)
1000	(-0.00124,0.04277,0.01302)	(0.21621,1.38223,0.20856)
5000	(-0.00359,0.00633,-0.00451)	(0.02393,0.04574,0.03272)
	$Q_1 = N(0,1)$, $Q_2 = N(1,9)$, $Q_3 = N(-1,4)$, $\boldsymbol{\theta} = (0.4, 0.8)$	
100	(0.09506,1.95064,0.20147)	(0.26613,4.10713,0.72572)
500	(0.03204,0.41988,0.01171)	(0.12747,1.39036,0.23704)
1000	(0.00749,0.08324,-0.00190)	(0.06577,0.29948,0.14719)
5000	(0.00105,0.01198,-0.00028)	(0.02325,0.06900,0.06344)

Table C.3: Bias and RMSE of the estimator of $\boldsymbol{\alpha}$ based on 10^4 Monte-Carlo repetitions.

	Confidence level					
	0.90		0.95		0.99	
	Expectations are assumed to be					
n	known	unknown	known	unknown	known	unknown
	$Q_1 = N(0,1)$, $Q_2 = N(1,1)$, $Q_3 = N(-1,1)$, $\boldsymbol{\theta} = (0.4, 0.8)$					
100	0.929	0.832	0.971	0.894	0.997	0.950
500	0.912	0.881	0.962	0.940	0.992	0.986
1000	0.932	0.925	0.974	0.969	0.996	0.995
5000	0.910	0.902	0.949	0.945	0.998	0.996
	$Q_1 = N(0,1)$, $Q_2 = N(1,1)$, $Q_3 = N(-1,1)$, $\boldsymbol{\theta} = (0.02, 0.995)$					
100	0.987	0.252	0.995	0.292	1.000	0.347
500	0.977	0.631	0.997	0.689	1.000	0.756
1000	0.936	0.802	0.974	0.858	0.997	0.927
5000	0.925	0.900	0.964	0.950	0.990	0.983
	$Q_1 = N(0,1)$, $Q_2 = N(1,1)$, $Q_3 = N(-1,1)$, $\boldsymbol{\theta} = (0.4, 0.41)$					
100	0.935	0.256	0.953	0.361	0.975	0.545
500	0.974	0.475	0.991	0.578	1.000	0.695
1000	0.950	0.659	0.988	0.743	1.000	0.842
5000	0.913	0.871	0.958	0.928	0.990	0.979
	$Q_1 = N(0,1)$, $Q_2 = N(0.2,1)$, $Q_3 = N(-0.1,1)$, $\boldsymbol{\theta} = (0.4, 0.8)$					
100	0.970	0.091	0.985	0.141	0.996	0.232
500	0.958	0.205	0.978	0.240	0.998	0.307
1000	0.944	0.342	0.975	0.403	1.000	0.486
5000	0.890	0.823	0.951	0.903	0.995	0.972

Continued on the next page

	$Q_1 = N(0,1)$, $Q_2 = N(1,9)$, $Q_3 = N(-1,4)$, $\boldsymbol{\theta} = (0.4, 0.8)$					
100	0.914	0.233	0.964	0.282	0.996	0.358
500	0.857	0.706	0.920	0.759	0.987	0.836
1000	0.853	0.838	0.917	0.903	0.970	0.963
5000	0.865	0.854	0.931	0.930	0.984	0.989

Table C.4: Approximated coverage probabilities (based on 10^3 intervals) for given confidence levels with respect to the asymptotic confidence regions for $\boldsymbol{\tau}_n = (\tau_n, \sigma_n)$ in the case of known and unknown expectations.

Bibliography

[1] N. Albrecht. Strong consistency of a nonparametric estimator in multiple change point models. Diploma thesis, Technische Universität Dresden, 2010.

[2] M. Basseville and I. V. Nikiforov. *Detection of Abrupt Changes: Theory and Application*. Prentice-Hall, Inc., 1993.

[3] J. V. Braun and H.-G. Müller. Statistical Methods for DNA Sequence Segmentation. *Statistical Science*, Vol. 13(2):142–162, 1998.

[4] B. Brodsky. *Change-Point Analysis in Nonstationary Stochastic Models*. CRC Press, 2017.

[5] B. E. Brodsky and B. S. Darkhovsky. *Nonparametric Methods in Change-Point Problems*. Kluwer Academic Publishers, 1993.

[6] J. Chen and A. K. Gupta. *Parametric Statistical Change Point Analysis: With Applications to Genetics, Medicine, and Finance*. Birkhäuser Basel, 2012.

[7] Y. S. Chow and H. Teicher. *Probability Theory: Independence, Interchangeability, Martingales*. Springer-Verlag New York, 1997.

[8] M. Csörgö and L. Horváth. *Limit Theorems in Change-Point Analysis*. John Wiley & Sons, 1997.

[9] M. Döring. *Mehrdimensionale Change-Point-Schätzung mit U-Statistiken*. PhD thesis, Technische Universität Dresden, 2007.

[10] M. Döring. Multiple change-point estimation with U-statistics. *Journal of Statistical Planning and Inference*, 140(7):2003 – 2017, 2010.

[11] M. Döring. Convergence in distribution of multiple change point estimators. *Journal of Statistical Planning and Inference*, 141(7):2238 – 2248, 2011.

[12] D. Ferger. Change-point estimators in case of small disorders. *Journal of Statistical Planning and Inference*, 40(1):33–49, 1994.

[13] D. Ferger. Exponential and polynomial tailbounds for change-point estimators. *Journal of Statistical Planning and Inference*, 92(1):73–109, 2001.

[14] D. Ferger. A continuous mapping theorem for the argmax-functional in the non-unique case. *Statistica Neerlandica*, 58(1):83–96, 2004.

[15] D. Ferger. Weighted Least Squares Estimators for a Change-Point. *Economic Quality Control*, 20(2):255–270, 2005.

[16] D. Ferger. Arginf-Sets of Multivariate Cadlag Processes and their Convergence in Hyperspace Topolgies. *Theory of Stochastic Processes*, 20(2):13–41, 2015.

[17] D. Ferger. Unpublished Manuscript, Technische Universität Dresden, 2018.

[18] S. Fremdt. *Asymptotic Methods in Change-Point Analysis*. PhD thesis, Universiät zu Köln, 2012.

[19] P. Gänssler and W. Stute. *Wahrscheinlichkeitstheorie*. Springer-Verlag, 1977.

[20] H. Heuser. *Lehrbuch der Analysis Teil 1*. B. G. Teubner, 2000.

[21] D. V. Hinkley. Inference about the change-point in a sequence of random variables. *Biometrika*, 57(1):1–17, 1970.

[22] E. S. Page. Continuous inspection schemes. *Biometrika*, 41(1/2):100–115, 1954.

[23] E. S. Page. A test for a change in a parameter occuring at an unknown point. *Biometrika*, 42(3/4):523–527, 1955.

[24] K. D. Schmidt. *Maß und Wahrscheinlichkeit*. Springer-Verlag, 2011.

[25] E. Seijo and B. Sen. A continuous mapping theorem for the smallest argmax functional. *Electronic Journal of Statistics*, 5:421–439, 2011.

[26] A. W. van der Vaart. *Asymptotic statistics*. Cambridge University Press, 1998.

Affirmation

Hereby I affirm that I wrote the present thesis without any inadmissible help by a third party and without using any other means than indicated. Thoughts that were taken directly or indirectly from other sources are indicated as such. This thesis has not been presented to any other examination board in this or a similar form, neither in Germany nor in any other country.

I have written this dissertation at Dresden University of Technology under the scientific supervision of Prof. Dr. Dietmar Ferger.

I accept the requirements for obtaining a PhD (Promotionsordnung) of the Faculty of Science of the TU Dresden, issued February 23, 2011.

Versicherung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die vorliegende Dissertation habe ich an der Technischen Universität Dresden unter der wissenschaftlichen Betreuung von Prof. Dr. Dietmar Ferger angefertigt.

Ich erkenne die Promotionsordnung der Fakultät Mathematik und Naturwissenschaften der TU Dresden vom 23. Februar 2011 an.

Dresden, 17. April 2018