Statistische Analyse epidemischer Strukturbrüche in Zufallsfeldern auf der Basis eines schwachen Invarianzprinzips

Masterarbeit
zur Erlangung des akademischen Grades
Master of Science (M.Sc.)
im Studiengang Mathematik

von
Béatrice Bucchia

angefertigt im Mathematischen Institut
der Universität zu Köln
unter Anleitung von

Prof. Dr. Josef Steinebach

Köln, den 20.12.2012

Inhaltsverzeichnis

1 Einleitung

1.1 Einleitung

In dieser Arbeit geht es um die statistische Analyse des epidemischen Strukturbruchproblems für stochastische Prozesse mit mehrdimensionalem Parameterraum. Diese Fragestellung gehört zum Gebiet der Changepoint-Analyse, welches sich mit der Frage beschäftigt, ob gegebene Beobachtungen homogen sind oder sich in Regionen mit unterschiedlichen Verteilungen aufteilen lassen. Ist letzteres der Fall, so ist man daran interessiert, diese Regionen durch Schätzer genauer zu bestimmen. Oft wird der Unterschied in der Struktur der Daten, abhängig vom gewählten Modell, an Verteilungsparametern wie dem Erwartungswert oder der Varianz gemessen. Wir werden in dieser Arbeit Veränderungen im Erwartungswert der Daten betrachten.

In der klassischen Literatur (vgl. z. B. Csörgő und Horváth (1997), Brodsky und Darkhovsky (1993)) werden die Beobachtungen meist als Realisationen eines Prozesses mit eindimensionalem Parameterraum zu bestimmten Zeitpunkten aufgefasst. In diesem Fall ist eine „epidemische“ Veränderung ein Strukturbruch, bei dem die Daten zunächst in ihrem „Normalzustand“ sind, dann, ähnlich wie bei einer Epidemie, in einen veränderten Zustand wechseln und schließlich wieder in ihren „Normalzustand“ zurückfallen. Möchte man nicht Zeitpunkte, sondern allgemeiner Regionen mit Veränderungen feststellen, so fasst man die Beobachtungen als Realistionen eines Zufallsfeldes mit mehrdimensionalem Parameterraum auf. Dann entspricht der Parameter des Feldes z. B. den Koordinaten eines Punktes in der Ebene. Im eindimensionalen Fall werden die unterschiedlichen Regionen eindeutig durch die Zeitpunkte der Veränderung (Changepoints), also durch ihren „Rand“, festgelegt. Ist der Parameterraum mehrdimensional, so lassen sich allgemeine Regionen auch durch ihren Rand abgrenzen, aber die Beschreibung dieses Randes ist i. Allg. komplizierter. Um das Problem zu vereinfachen, nimmt man daher oft an, dass die Regionen eine bestimmte Form haben, die sich durch wenige Parameter beschreiben lässt. In dieser Arbeit werden wir stets annehmen, dass entweder kein Strukturbruch vorliegt, oder der Parameterraum in zwei Mengen zerfällt: Die Region, in der der „Normalzustand“ herrscht, und ein „Rechteck“ $(a_1, b_1] \times \cdots \times (a_d, b_d]$ ($d \in \mathbb{N}$), in dem der Erwartungswert der Daten verändert ist. Dabei ist das „Rechteck“ eine mehrdimensionale Verallgemeinerung des im eindimensionalen Fall betrachteten Intervalls, in dem der epidemische Strukturbruch stattfindet. Die Punkte $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ werden wir in Anlehnung an den eindimensionalen Fall als Changepoints bezeichnen, da sie die Region mit verändertem Erwartungswert charakterisieren.

Wir werden dabei sogenannte „off-line“ Changepoint-Analyse betreiben, bei der im Gegensatz zur „on-line“ Changepoint-Analyse alle Beobachtungen bereits vorliegen und a posteriori entschieden werden muss, ob sie eine Veränderung aufweisen.

Das Problem der Feststellung eines epidemischen Strukturbruchs wurde von Levin und Kline (1985) eingeführt und seitdem von einigen Autoren aufgegriffen (siehe Csörgő und Horváth (1997), Brodsky und Darkhovsky (1993), Hušková (1995) und Antoch und Hušková (1996) für mehr Referenzen). Das Testproblem für den epidemischen Strukturbruch auf Zufallsfeldern wurde von Jarušková und Piterbarg (2011) unter der Annahme betrachtet, dass die Beobachtungen unabhängig sind. Diese Voraussetzung wollen wir nun abschwächen, indem wir sie durch die Annahme eines schwachen Invarianzprinzips (siehe Definition 9) ersetzen, welches Aussagen über die asymptotische Verteilung gewichteter Partialsummen der Felder macht. Dies ermöglicht es uns, unabhängig von vorgegebenen Abhängigkeitsstrukturen oder Verteilungen asymptotische Aussagen über die betrachteten Statistiken und Schätzer zu treffen. Da das Invarianzprinzip insbesondere von Zufallsfeldern mit unabhängigen Zufallsvariablen erfüllt wird (vgl. z. B. Wichura (1969)), ergeben sich entsprechende Aussagen über unabhängige Zufallsvariablen als Spezialfall.

Die Arbeit ist wie folgt aufgebaut: Im ersten Kapitel führen wir zunächst die hier benutzten Notationen und Konventionen, sowie das in dieser Arbeit betrachtete Changepoint-Modell ein. Das nächste Kapitel enthält einen Exkurs über den Raum $D[0,1]^d$, der eine Erweiterung des klassischen Skorohod-Raums der càdlàg Prozesse für Prozesse mit mehrdimensionalem Parameterraum ist. Die Betrachtung von $D[0,1]^d$ in dieser Arbeit ergibt sich daraus, dass die hier benutzten Statistiken auf den Partialsummen, und damit auf einer Diskretisierung der Partialsummenprozesse (siehe (1.1)) beruhen. Betrachtet man den zu einem stochastischen Prozess $\{X_{\mathbf{j}}\}_{\mathbf{j}\in\mathbb{Z}^d}$ gehörigen Partialsummenprozess, so hat dieser Pfade in $D[0,1]^d$. Nachdem wir $D[0,1]^d$ als polnischen Raum eingeführt haben, definieren wir im nächsten Abschnitt, was wir unter Verteilungskonvergenz in $D[0,1]^d$ verstehen werden, und zitieren ein von Bickel und Wichura (1971) hergeleitetes Hilfsmittel zum Beweis der Verteilungskonvergenz eines Prozesses in $D[0,1]^d$. Ab dem dritten Kapitel wenden wir uns der Betrachtung des Changepoint-Problems als Test- und Schätzproblem zu. Dabei setzen wir stets voraus, dass die betrachteten Prozesse ein schwaches Invarianzprinzip (siehe Definition 9) erfüllen. Zunächst betrachten wir das Testproblem. Als Teststatistik nehmen wir eine leichte Modifikation der in Jarušková und Piterbarg (2011) betrachteten Statistik. Wir ersetzen die dortige Voraussetzung an die Beobachtungen, unabhängig zu sein, durch die Forderung, die Beobachtungen mögen einem Invarianzprinzip genügen, und zeigen, dass unter diesen Bedingungen das asymptotische Verhalten unverändert bleibt. Dabei folgen wir dem Vorgehen in Jarušková und Piterbarg (2011) und betrachten nicht die exakte Verteilung der Statistik, sondern konzentrieren uns auf ihre asymptotische Verteilung, für deren Tail-Verhalten wir eine Approximation herleiten. Dann zeigen wir die Konsistenz des Tests unter der Alternative. Im vierten Kapitel schätzen wir dann die Changepoints unter einer lokalen Alternative. Schließlich betrachten wir noch Beispiele für Prozesse, die das von uns vorausgesetzte schwache Invarianzprinzip erfüllen. Als Beispiel für Prozesse mit eindimensionalem Parameterraum, die das schwache Invarianzprinzip erfüllen, betrachten wir die von Berkes et al. (2011) eingeführten schwach $\mathcal{M}$-abhängigen Prozesse. Berkes et al. (2011) zeigten für diese Prozesse ein starkes Invarianzprinzip. Wir leiten aus diesem das schwache Invarianzprinzip her und stellen einige Beispiele von schwach $\mathcal{M}$-abhängigen Prozessen vor. Als Beispiel für Prozesse mit mehrdimensionalem Parameterraum, die das schwache Invarianzprinzip erfüllen, betrachten wir Prozesse, die einer Assoziationsbedingung genügen. Der Anhang enthält einige technische Beweise.

Grundlegende Resultate und Begriffe der Wahrscheinlichkeitstheorie, der Statistik und der Theorie stochastischer Prozesse werden vorausgesetzt.

1.2 Definitionen, Konventionen und Hintergrund

Im folgenden Abschnitt führen wir einige Notationen und Definitionen ein, die wir später voraussetzen wollen. Da wir in dieser Arbeit Zufallsfelder betrachten, werden immer wieder mehrdimensionale Ausdrücke auftauchen, für die wir einige Kurzschreibweisen festlegen. Um Vektoren optisch von reellen Zahlen abzugrenzen, werden diese fett gedruckt. Zunächst einige allgemeine Definitionen. Für $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$, $n \in \mathbb{N}$ und $x \in \mathbb{R}$ definieren wir:

$$\begin{aligned}
\mathbf{a} \leq \mathbf{b} \;&:\Leftrightarrow\; \forall i = 1, \dots, d: \quad a_i \leq b_i \\
\mathbf{a} < \mathbf{b} \;&:\Leftrightarrow\; \forall i = 1, \dots, d: \quad a_i < b_i \\
\mathbf{a} \pm \mathbf{b} &:= (a_1 \pm b_1, \dots, a_d \pm b_d) \\
\mathbf{ab} &:= (a_1 b_1, \dots, a_d b_d) \\
\mathbf{a} \wedge \mathbf{b} &:= \prod_{i=1}^{d} \min(a_i, b_i) \\
n\mathbf{a} &:= (na_1, \dots, na_d) \\
\lfloor \mathbf{a} \rfloor &:= (\lfloor a_1 \rfloor, \dots, \lfloor a_d \rfloor) \\
[\mathbf{a}] &:= \prod_{i=1}^{d} a_i \\
\|\mathbf{a}\| &:= \max_{i=1,\dots,d} |a_i| \\
\underline{\mathbf{x}} &:= (x, \dots, x) \in \mathbb{R}^d
\end{aligned}$$

Für Teilmengen I, J von $\mathbb{R}^d$ definieren wir den Abstand

$$d(I, J) := \inf\{\|\mathbf{j} - \mathbf{i}\| : \mathbf{i} \in I, \mathbf{j} \in J\}.$$

Für eine endliche Teilmenge $I \subset \mathbb{R}^d$ bezeichne $|I|$ die Kardinalität von I. I heißt *Rechteck*, wenn $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ mit $\mathbf{a} \leq \mathbf{b}$ existieren, so dass

$$I = (a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_d, b_d].$$

Ein *Rechteck I in $\mathbb{Z}^d$* sei eine Menge der obigen Form, wobei $\mathbf{a}, \mathbf{b} \in \mathbb{Z}^d$ und ein „Intervall“ $(z_1, z_2] \subset \mathbb{Z}$ eine Menge der Form

$$(z_1, z_2] = \{z_1 + 1, \ldots, z_2\}$$

sei. Wir schreiben $I = (\mathbf{a}, \mathbf{b}]$. Analog seien auch $I = [\mathbf{a}, \mathbf{b}]$ usw. zu verstehen. Für $\mathbf{a}, \mathbf{b} \in \mathbb{Z}^d$ und $x_\mathbf{j} \in \mathbb{R}$ sei

$$\sum_{\mathbf{a} \leq \mathbf{j} \leq \mathbf{b}} x_\mathbf{j} := \begin{cases} \sum_{j_1=a_1}^{b_1} \cdots \sum_{j_d=a_d}^{b_d} x_\mathbf{j}, & \text{falls } \mathbf{a} \leq \mathbf{b} \\ 0, & \text{sonst.} \end{cases}$$

Für eine Menge E sei I_E die Indikatorfunktion auf E. Für Folgen $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ bezeichnet $a_n \ll b_n$ die Eigenschaft, dass $\limsup_{n \to \infty} |a_n / b_n| < \infty$ und $a_n \sim b_n$ die Eigenschaft, dass $\lim_{n \to \infty} a_n / b_n = 1$. Wenn nichts anderes angegeben ist, sei im Folgenden bei $o(1)$, $\mathcal{O}(1)$, $o_P(1)$, $\mathcal{O}_P(1)$ usw. immer die Konvergenz für $n \to \infty$ gemeint.

Für ein Feld $\{a_\mathbf{j}\}_{\mathbf{j} \in \mathbb{Z}^d}$ von reellen Zahlen schreiben wir $a = \sum_{\mathbf{j} \in \mathbb{Z}^d} a_\mathbf{j}$, wenn für alle $\varepsilon > 0$ ein $\mathbf{n}_0 \in \mathbb{N}^d$ existiert, so dass für alle $\mathbf{m}, \mathbf{n} \geq \mathbf{n}_0$ gilt

$$\left| a - \sum_{-\mathbf{m} \leq \mathbf{j} \leq \mathbf{n}} a_\mathbf{j} \right| < \varepsilon.$$

Wir sagen, die Reihe $\sum_{\mathbf{j} \in \mathbb{Z}^d} a_\mathbf{j}$ *konvergiert absolut*, wenn $\sum_{\mathbf{j} \in \mathbb{Z}^d} |a_\mathbf{j}| < \infty$.

Nun zu wahrscheinlichkeitstheoretischen Definitionen:
Definition 1. *Für einen metrischen Raum $(\mathcal{X}, \mathcal{B})$ und einen Parameterraum T bezeichnen wir eine Familie $\{X_t\}_{t \in T}$ von Zufallsvariablen $X_t : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ als stochastischen Prozess.*
Definition 2. *Eine Familie $\{X_\mathbf{j}\}_{\mathbf{j} \in \mathbb{Z}^d}$ $(d \in \mathbb{N})$ von reellen Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ bezeichnen wir als (reellwertiges) Zufallsfeld.*

Wie aus obigen Definitionen erkennbar, fassen wir stochastische Prozesse als Familien von Zufallsvariablen auf, deren Parameterraum mehrdimensional sein kann, aber nicht sein muss. Offensichtlich sind Zufallsfelder insbesondere auch stochastische Prozesse. Durch den Begriff Zufallsfeld wird betont, dass der Parameterraum mehrdimensional sein kann. Da wir im Folgenden nicht immer zwischen ein- und mehrdimensionalem Parameterraum unterscheiden wollen, werden wir auch für $d = 1$ von Zufallsfeldern reden. Beziehen wir uns explizit auf eine Familie $\{X_j\}_{j \in \mathbb{Z}}$

mit eindimensionalem Parameterraum, so werden wir diese als stochastischen Prozess bezeichnen. Wenn im Folgenden von einer Zufallsvariablen, einem Zufallsfeld oder einem stochastischen Prozess die Rede ist, werden wir den zugrunde gelegten Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ nicht immer explizit erwähnen. Wir werden stets voraussetzen, dass $(\Omega, \mathcal{A}, P)$ vollständig ist. Wenn nichts anderes gesagt wird, sei im Folgenden mit einer Zufallsvariable stets eine reellwertige Zufallsvariable gemeint.

Definition 3. *Für $p > 0$ und eine Zufallsvariable Y bezeichne $\|Y\|_p = (E[|Y|^p])^{1/p}$ die p-Norm von Y und L^p sei der Raum der Zufallsvariablen mit endlicher p-Norm.*

Definition 4. *(vgl. Bulinski und Shashkin (2007), S. 175) Ein Feld $X = \{X_{\mathbf{j}}\}_{\mathbf{j}\in\mathbb{Z}^d}$, das aus quadratisch integrierbaren Zufallsvariablen besteht, heißt schwach stationär, wenn es eine Konstante c und eine Funktion r gibt, so dass für alle $\mathbf{i}, \mathbf{j} \in \mathbb{Z}^d$ gilt:*

$$E[X_{\mathbf{j}}] = c \text{ und } \operatorname{Cov}(X_{\mathbf{i}}, X_{\mathbf{j}}) = r(\mathbf{i} - \mathbf{j})$$

Ein Feld $X = \{X_{\mathbf{j}}\}_{\mathbf{j}\in\mathbb{Z}^d}$ heißt (stark) stationär, wenn für alle $n \in \mathbb{N}$, $\mathbf{i}_1, \dots, \mathbf{i}_n \in \mathbb{Z}^d$ und $\mathbf{h} \in \mathbb{Z}^d$ die Vektoren $(X_{\mathbf{i}_1}, \dots, X_{\mathbf{i}_n})$ und $(X_{\mathbf{i}_1+\mathbf{h}}, \dots, X_{\mathbf{i}_n+\mathbf{h}})$ identisch verteilt sind.

Definition 5. *Ein Zufallsfeld $\{X_{\mathbf{j}}\}_{\mathbf{j}\in\mathbb{Z}^d}$ wird als Gauß'sches Feld bezeichnet, wenn seine endlich-dimensionalen Verteilungen Gauß'sch, d. h. multivariat normalverteilt, sind. Dabei bezeichnen wir einen Vektor $(X_1, \dots, X_k)$, $k \in \mathbb{N}$, als multivariat normalverteilt, wenn alle seine Linearkombinationen $\sum_{i=1}^k a_i X_i$, $a_1, \dots, a_k \in \mathbb{R}$, normalverteilte Zufallsvariablen sind.*

Als mehrdimensionale Verallgemeinerung des bekannten Wiener-Prozesses werden Wiener Felder für unsere Konvergenzbetrachtungen eine zentrale Rolle spielen. Diese sind wie folgt definiert (siehe Adler (1990), S. 6 f.):

Definition 6. *Für $d \in \mathbb{N}$ seien $\mathcal{B}^d$ die Menge der Borel-Mengen auf $\mathbb{R}^d$ und λ^d das Lebesgue-Maß auf $\mathcal{B}^d$. Ein Gauß'sches Weißes Rauschen ist ein Prozess $\{W(A) : A \in \mathcal{B}^d\}$ auf einem vollständigen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$, der für disjunkte Mengen $A, B \in \mathcal{B}^d$ mit endlichem Lebesgue-Maß folgende Eigenschaften besitzt:*

(i) $W(A)$ ist normalverteilt mit Erwartungswert 0 und Varianz $\lambda^d(A)$.

(ii) $W(A \cup B) = W(A) + W(B)$ P-fast sicher

(iii) $W(A)$ und $W(B)$ sind unabhängig.

Das zentrierte Gauß'sche Feld $\{W(\mathbf{t})\}_{\mathbf{t}\in[0,\infty)^d}$ mit $W(\mathbf{t}) := W((\underline{\mathbf{0}}, \mathbf{t}])$ bezeichnen wir als Wiener Feld.

Bemerkung 7. *Aus den Eigenschaften (i)-(iii) folgt für Mengen $A, B \in \mathcal{B}^d$ mit endlichem Lebesgue-Maß:*

$$\begin{aligned} E[W(A)W(B)] &= E[(W(A \cap B) + W(A \setminus B))(W(A \cap B) + W(B \setminus A))] \\ &= E[W(A \cap B)^2] \\ &= \lambda^d(A \cap B) \end{aligned}$$

Für ein Wiener Feld und $\mathbf{s}, \mathbf{t} \in [0,1]^d$ *gilt somit:*

$$E[W(\mathbf{s})W(\mathbf{t})] = \lambda^d((\underline{\mathbf{0}}, \mathbf{s}] \cap (\underline{\mathbf{0}}, \mathbf{t}]) = \mathbf{s} \wedge \mathbf{t}$$

Für das Weiße Rauschen und $\underline{\mathbf{0}} \leq \mathbf{s} \leq \mathbf{t} \leq \underline{\mathbf{1}}$ *besteht aufgrund von Eigenschaften (i) und (ii) folgender Zusammenhang zwischen dem Prozess* $\{W(A) : A \in \mathcal{B}^d\}$ *am Index* $(\mathbf{s}, \mathbf{t}] \subset \mathbb{R}^d$ *und dem Zuwachs (siehe* (2.2)*) von* $\{W(\mathbf{t})\}_{\mathbf{t} \in [0,\infty)^d}$ *auf dem Rechteck* $(\mathbf{s}, \mathbf{t}]$*:*

$$W((\mathbf{s}, \mathbf{t}]) = \sum_{\boldsymbol{\varepsilon} \in \{0,1\}^d} (-1)^{d - \sum_{i=1}^d \varepsilon_i} W(\mathbf{s} + \boldsymbol{\varepsilon}(\mathbf{t} - \mathbf{s})) \quad P - f.s.$$

Bemerkung 8. *Wie bei Wiener-Prozessen mit eindimensionalem Parameterraum sind auch die Pfade eines Wiener Feldes fast sicher stetig (vgl. Adler (1990), Theorem 1.2, S. 8) und es existiert eine stetige Modifikation (vgl. Khoshnevisan (2002), Theorem 3.2.1, S.174).*

In dieser Arbeit werden wir stets Invarianzprinzipien voraussetzen. Diese ermöglichen es uns, Teststatistiken und Schätzer durch den Übergang auf asymptotische Verteilungen zu untersuchen, ohne für die betrachteten Zufallsfelder bestimmte Verteilungen oder Abhängigkeitsstrukturen fordern zu müssen. Nun konkretisieren wir, was unter einem schwachen Invarianzprinzip gemeint sein wird. Für ein Zufallsfeld $X = \{X_{\mathbf{j}}\}_{\mathbf{j} \in \mathbb{Z}^d}$ sei für $n \in \mathbb{N}$

$$\tilde{S}_n(\mathbf{t}) := \frac{1}{n^{d/2}} \sum_{\underline{\mathbf{1}} \leq \mathbf{j} \leq \lfloor n\mathbf{t} \rfloor} X_{\mathbf{j}} \tag{1.1}$$

der zugehörige Partialsummenprozess. Die folgende Definition wurde aus Bulinski und Shashkin (2007) (Definition 5.1.4, S. 255) entnommen. Eine genaue Definition davon, was wir unter $\stackrel{D[0,1]^d}{\longrightarrow}$ verstehen, findet sich im Abschnitt 2.2.

Definition 9. *Wir sagen, dass ein Zufallsfeld* $X = \{X_{\mathbf{j}}\}_{\mathbf{j} \in \mathbb{Z}^d}$ *das* <u>*schwache Invarianzprinzip*</u> *(auch Funktionaler Zentraler Grenzwertsatz genannt) erfüllt, wenn es ein* $\sigma^2 \geq 0$ *und ein Wiener Feld* $\{W(\mathbf{t})\}_{\mathbf{t} \in [0,1]^d}$ *gibt, so dass für den zu X gehörigen Partialsummenprozess* $\{\tilde{S}_n(\mathbf{t})\}_{\mathbf{t} \in [0,1]^d}$ *gilt:*

$$\{\tilde{S}_n(\mathbf{t})\}_{\mathbf{t} \in [0,1]^d} \stackrel{D[0,1]^d}{\longrightarrow} \{\sigma W(\mathbf{t})\}_{\mathbf{t} \in [0,1]^d}, \ n \to \infty \tag{1.2}$$

Definition 10. *(siehe Bulinski und Shashkin (2007), S.90, Newman (1980), S. 122) Für ein zentriertes, schwach stationäres Zufallsfeld* $X = \{X_{\mathbf{j}}\}_{\mathbf{j} \in \mathbb{Z}^d}$ *sagen wir, dass* X *die* <u>*endliche Suszeptibilitätsbedingung*</u> *erfüllt, wenn*

$$\sigma^2 := \sum_{\mathbf{j} \in \mathbb{Z}^d} \mathrm{Cov}(X_{\underline{\mathbf{0}}}, X_{\mathbf{j}}) < \infty. \tag{1.3}$$

Wie Bulinski und Shashkin (2007) anmerken, lässt sich zeigen, dass σ^2 aus (1.3) die asymptotische Varianz normalisierter Partialsummen und somit insbesondere nichtnegativ ist:

Satz 11. *Sei* $X = \{X_{\mathbf{j}}\}_{\mathbf{j}\in\mathbb{Z}^d}$ *ein schwach stationäres Zufallsfeld, für das die Reihe in* (1.3) *absolut konvergent ist. Weiter sei* $S_{\underline{\mathbf{n}}} := \sum\limits_{\underline{\mathbf{1}}\leq\mathbf{j}\leq\underline{\mathbf{n}}} X_{\mathbf{j}}$ $(n \in \mathbb{N})$*. Dann gilt:*

$$\frac{\mathrm{Var}(S_{\underline{\mathbf{n}}})}{n^d} \to \sigma^2,\ n \to \infty \tag{1.4}$$

Beweis. Satz 11 ist eine Einschränkung von Theorem 3.1.8 von Bulinski und Shashkin (2007) auf die speziellen Mengen $U_n = \{1, \dots, n\}^d = (\underline{\mathbf{0}}, \underline{\mathbf{n}}] \cap \mathbb{Z}^d$. Diese wachsen gemäß Lemma 3.1.5 von Bulinski und Shashkin (2007) regulär im Sinne ihrer Definition 3.1.4. □

Zum Schluss dieses Abschnitts betrachten wir noch ein Lemma, das wir zum Beweis der Stetigkeit von abschnittsweise definierten Funktionen benutzen werden:

Lemma 12. $(X, \mathcal{T}_X)$ *und* $(Y, \mathcal{T}_Y)$ *seien topologische Räume. Für eine Teilmenge* $M \subset X$ *bezeichnen wir mit* $\mathcal{T}_{X,M} := \{A \cap M : A \in \mathcal{T}_X\}$ *die Spurtopologie von* M*. Für ein* $n \in \mathbb{N}$ *und abgeschlossene Mengen* $A_1, \dots, A_n \subset X$ *gelte* $X = \bigcup\limits_{k=1}^{n} A_k$*. Weiter seien* $f_k : A_k \to Y$ *stetige Funktionen bezüglich der Spurtopologien* $\mathcal{T}_{X,A_k}$ $(k = 1, \dots, n)$ *und es gelte* $f_k(x) = f_l(x)$ *für* $x \in A_k \cap A_l$ $(k, l \in \{1, \dots, n\})$*. Dann ist die Funktion* $f : X \to Y$ *mit* $f(x) = f_k(x)$ *für* $x \in A_k$ *bezüglich* $\mathcal{T}_X$ *stetig.*

Beweis. siehe Anhang □

1.3 Das epidemische Changepoint-Problem

In dieser Arbeit wird das epidemische Changepoint-Problem für stochastische Prozesse mit ein- und mehrdimensionalem Parameterraum betrachtet. Gegeben beobachtete Realisationen der n^d Zufallsvariablen $X_{\mathbf{j}}$, $\mathbf{j} \in \{1, \dots, n\}^d$, $n, d \in \mathbb{N}$, muss entschieden werden, ob

$$H_0: \ X_{\mathbf{k}} = Y_{\mathbf{k}} + \mu \quad \forall\, \mathbf{k} \in \{1, \dots, n\}^d$$

oder

$$H_A: \ \exists\, \underline{\mathbf{1}} \leq \mathbf{k}_0 < \mathbf{m}_0 \leq \underline{\mathbf{n}}: \ X_{\mathbf{k}} = Y_{\mathbf{k}} + \mu + \delta\, I_{\{\mathbf{k}_0 < \mathbf{k} \leq \mathbf{m}_0\}} \quad \forall\, \mathbf{k} \in \{1, \dots, n\}^d,$$

wobei μ, δ unbekannte reelle Zahlen und $\{Y_{\mathbf{k}}\}_{\mathbf{k}\in\mathbb{Z}^d}$ ein zentriertes, schwach stationäres Zufallsfeld sei. Die Punkte $\mathbf{k}_0$ und $\mathbf{m}_0$ bezeichnen wir als Changepoints, da sie den Rand der Menge charakterisieren, auf der das Verhalten des Prozesses verändert ist. δ ist die Höhe der Veränderung im Erwartungswert. Wir gehen davon aus, dass die Beobachtungen in ihrem „Normalzustand“ einen (unbekannten) Erwartungswert μ haben, und sind daran interessiert, retrospektiv festzustellen, ob und wo es ein Rechteck $(\mathbf{k}_0, \mathbf{m}_0] \subset \{1, \dots, n\}^d$ gibt, in dem die Daten einen veränderten Erwartungswert $\mu + \delta$ haben. Dabei wird die Annahme, dass kein Strukturbruch stattgefunden hat, als Nullhypothese gewählt, um die Wahrscheinlichkeit, dass homogene Daten durch zufällige Fluktuationen in den Beobachtungen fälschlicherweise als inhomogen angenommen werden, zu kontrollieren.

Bemerkung 13. *Die Größen* $\mathbf{k}_0$*,* $\mathbf{m}_0$ *und* δ *können von der Stichprobengröße* n *abhängen, um die Aussagen lesbarer zu machen wird dies aber in der Notation oft unterdrückt.*

Den Begriff der epidemischen Veränderung führen Levin und Kline (1985) in einer Untersuchung über den Zusammenhang zwischen chromosomalen Veränderungen und der Häufigkeit, mit der Fehlgeburten eintreten, ein. Diese medizinische Anwendung motiviert die Wahl der Bezeichnung „epidemisch“, weil die Datenfluktuation einer Epidemie entspricht. Retrospektiv betrachtet zeichnet sich eine Epidemie nämlich durch einen Anstieg und ein späteres Abfallen der Patientenzahl aus. Bezieht man - anders als Levin und Kline (1985) - neben Zeit auch Ort sowie Patientendaten ein, gerät man in die Situation, einen mehrdimensionalen Parameterraum betrachten zu wollen, wie dies in der vorliegenden Arbeit getan wird. Changepoint-Probleme finden außerhalb der Medizin auch in der Signaltransmission und bei der Prüfung der Homogenität hergestellter Textilien Anwendung.

Einen guten Einblick in die Forschung liefern die Bücher von Csörgő und Horváth (1997) und Brodsky und Darkhovsky (1993), sowie die Artikel von Antoch und Hušková (1996) und Hušková (1995), die sich speziell mit epidemischen Strukturbrüchen für unabhängige Beobachtungen mit eindimensionalem Parameterraum beschäftigen und einige Teststatistiken und Schätzer für die Changepoints betrachten. Für unabhängige, normalverteilte Beobachtungen betrachtete Yao (1993) das Tail-Verhalten einiger Teststatistiken. Jarušková (2011) beschäftigte sich mit dem Testproblem für den Fall, dass der Erwartungswert im epidemischen Intervall nicht konstant ist, sondern sich linear verändert. Der hier verfolgte Ansatz entspricht für das Testproblem dem Artikel von Jarušková und Piterbarg (2011), die als Beobachtungen unabhängige Zufallsvektoren mit mehrdimensionalem Parameterraum betrachten und eine Veränderung im Erwartungswert untersuchen. Dabei betrachten wir statt Zufallsvektoren reelle Zufallsvariablen und ersetzen die Unabhängigkeitsannahme durch die Annahme eines schwachen Invarianzprinzips. Für das Schätzproblem betrachten wir eine mehrdimensionale Verallgemeinerung des in Aston und Kirch (2012) untersuchten Schätzers.

Im Folgenden betrachten wir den Prozess $\{S_n(\mathbf{t})\}_{\mathbf{t}\in[0,1]^d}$ mit

$$S_n(\mathbf{t}) := \hat{\sigma}_n^{-1} n^{-d/2} \sum_{\mathbf{1}\leq\mathbf{j}\leq\lfloor n\mathbf{t}\rfloor} Y_{\mathbf{j}}.$$

Wir werden stets voraussetzen, dass folgende Annahme erfüllt ist:
Annahme 1.

$$\{S_n(\mathbf{t})\}_{\mathbf{t}\in[0,1]^d} \xrightarrow{D[0,1]^d} \{W(\mathbf{t})\}_{\mathbf{t}\in[0,1]^d},$$

wobei $\{W(\mathbf{t})\}_{\mathbf{t}\in[0,1]^d}$ *ein Wiener Feld und* $\hat{\sigma}_n^2$ *ein (schwach) konsistenter Schätzer für die asymptotische Varianz* $\sigma^2 > 0$ *sei.*

Diese Annahme ist erfüllt, wenn $\{Y_{\mathbf{k}}\}_{\mathbf{k}\in\mathbb{Z}^d}$ das schwache Invarianzprinzip (1.2) mit $\sigma^2 > 0$ erfüllt, und $\hat{\sigma}_n$ ein konsistenter Schätzer für σ ist (d. h. $\hat{\sigma}_n - \sigma = o_P(1)$).

Dann gilt nämlich

$$\frac{\sigma}{\hat{\sigma}_n} \xrightarrow{P} 1,$$

denn für beliebige $\gamma, \varepsilon > 0$ existieren $n_0 \in \mathbb{N}$ und $C > 0$, so dass für alle $n \geq n_0$ $P(|\hat{\sigma}_n| > C) < \varepsilon/2$ und $P(|\sigma - \hat{\sigma}_n| > C\gamma) < \varepsilon/2$, also

$$\begin{aligned} P\left(\left|\frac{\sigma}{\hat{\sigma}_n} - 1\right| > \gamma\right) &= P\left(\left|\frac{\sigma - \hat{\sigma}_n}{\hat{\sigma}_n}\right| > \gamma\right) \\ &\leq P(|\sigma - \hat{\sigma}_n| > C\gamma, |\hat{\sigma}_n| \leq C) + P\left(\left|\frac{\sigma - \hat{\sigma}_n}{\hat{\sigma}_n}\right| > \gamma, |\hat{\sigma}_n| > C\right) \\ &\leq P(|\sigma - \hat{\sigma}_n| > C\gamma) + P(|\hat{\sigma}_n| > C) < \varepsilon. \end{aligned}$$

Wegen

$$\frac{1}{\hat{\sigma}_n} n^{-d/2} \sum_{\mathbf{1} \leq \mathbf{j} \leq \lfloor n\mathbf{t} \rfloor} Y_{\mathbf{j}} = \frac{\sigma}{\hat{\sigma}_n} \frac{1}{\sigma} n^{-d/2} \sum_{\mathbf{1} \leq \mathbf{j} \leq \lfloor n\mathbf{t} \rfloor} Y_{\mathbf{j}},$$

gilt somit nach Lemma 34:

$$\left\{ \hat{\sigma}_n^{-1} n^{-d/2} \sum_{\mathbf{1} \leq \mathbf{j} \leq \lfloor n\mathbf{t} \rfloor} Y_{\mathbf{j}} \right\}_{\mathbf{t} \in [0,1]^d} \xrightarrow{D[0,1]^d} \{W(\mathbf{t})\}_{\mathbf{t} \in [0,1]^d}$$

Bemerkung 14. *Wir werden uns in dieser Arbeit nicht mit der konkreten Form solcher Schätzer für σ^2 beschäftigen und lediglich im Zusammenhang mit den Beispielen in Kapitel 5 auf spezielle Schätzer verweisen.*

2 Exkurs: Der Raum $D[0,1]^d$

2.1 Der Raum $D[0,1]^d$

In diesem Abschnitt wollen wir den Funktionenraum $D[0,1]^d$ ($d \in \mathbb{N}$), so wie er in Bickel und Wichura (1971) (S. 1662 ff.), Neuhaus (1969, 1971) und Straf (1972) betrachtet wird, einführen. Da das Hauptaugenmerk dieser Arbeit auf der Changepoint-Analyse liegt, werden wir die meisten Aussagen dieses Abschnitts lediglich zitieren. Eine ausführlichere Einführung sowie Beweise finden sich in den genannten Quellen sowie bei Billingsley (1968) und Jacod und Shiryaev (2003). Die hier beschriebenen Räume wurden als Verallgemeinerung des in Billingsley (1968) beschriebenen Ansatzes so gewählt, dass sie für $d = 1$ mit dem bekannten Skorohod-Raum (vgl. Billingsley (1968), Chapter 3) übereinstimmen. Wir wählen die Herangehensweise von Bickel und Wichura (Bickel und Wichura (1971), S. 1662) und definieren:

Definition 15. *Sei $T := [0,1]^d$. Eine Funktion $x : T \to \mathbb{R}$ heißt Treppenfunktion, wenn x eine Linearkombination von Funktionen der Form $t \mapsto I_{E_1 \times E_2 \times \dots \times E_d}(t)$ ist, wobei jedes E_d entweder ein links offenes, rechts abgeschlossenes Teilintervall von $[0,1]$, oder die einpunktige Menge $\{1\}$ ist. Sei V der Vektorraum der Treppenfunktionen. Dann bezeichnen wir mit $D[0,1]^d$ den Abschluss von V bezüglich der Supremumsnorm im Raum der beschränkten Funktionen von T nach $\mathbb{R}$.*

Alternativ lässt sich der Raum $D[0,1]^d$ auch wie folgt charakterisieren: Für $\mathbf{t} \in T$ und $R_p \in \{<, \geq\}$ ($1 \leq p \leq d$) definiere den Quadranten

$$Q_{R_1,\dots,R_d}(\mathbf{t}) := \{(s_1, \dots, s_d) \in T : s_p R_p t_p, 1 \leq p \leq d\}.$$

Für einen solchen Quadranten $Q = Q_{R_1,\dots,R_d}(\mathbf{t})$ und eine Funktion $x : T \to \mathbb{R}$ sei weiter

$$x_Q(\mathbf{t}) := \lim_{\mathbf{s} \to \mathbf{t}, \mathbf{s} \in Q} x(\mathbf{s}),$$

wenn der Grenzwert existiert. Nun gilt (vgl. Neuhaus (1971), S. 1288, Neuhaus (1969), Definition 1.2, S. 9, 18):

Lemma 16. *Eine Funktion $x : T \to \mathbb{R}$ ist genau dann in $D[0,1]^d$, wenn für alle $\mathbf{t} \in T$ und alle 2^d Quadranten $Q = Q_{R_1,\dots,R_d}(\mathbf{t})$ der Grenzwert $x_Q(\mathbf{t})$ existiert und $x = x_{Q_{\geq,\dots,\geq}}$ ist.*

Funktionen in $D[0,1]^d$ sind also „von oben stetig mit Grenzwerten von unten“.

Bemerkung 17. *Aufgrund der Eigenschaften des Grenzwertes sind für $x, y \in D[0,1]^d$ auch $x + y$ und $x \cdot y$ in $D[0,1]^d$.*

Beweis. siehe Anhang □

Nun führen wir eine metrische Topologie auf $D[0,1]^d$ ein. Diese wird so gewählt, dass sie für $d = 1$ mit der bekannten von Skorohod eingeführten J_1-Topologie (vgl. Billingsley (1968)) übereinstimmt. Dazu sei Λ der Raum der Funktionen $\lambda : T \to T$, mit $\lambda(t_1, \dots, t_d) = (\lambda_1(t_1), \dots, \lambda_d(t_d))$, wobei für alle $1 \le p \le d$ $\lambda_p : [0,1] \to [0,1]$ stetig und streng monoton wachsend mit $\lambda_p(0) = 0$ und $\lambda_p(1) = 1$ sei. Für Funktionen $x, y \in D[0,1]^d$ definieren wir den Abstand

$$d_S(x,y) := \inf\{\max(\|x - y\lambda\|_\infty, \|\lambda - id\|_\infty)\} : \lambda \in \Lambda\},$$

wobei $\|f\|_\infty := \sup\{|f(\mathbf{t})| : \mathbf{t} \in T\}$ für Funktionen $f : T \to \mathbb{R}$ und id die Identität auf T seien. Die zugehörige Topologie bezeichnen wir als <u>*S-Topologie*</u>. Dann gilt (vgl. Neuhaus (1971), S. 1289):

Lemma 18. *Der so definierte Abstand d_S ist eine Metrik auf $D[0,1]^d$. Für $x, x_1, \dots \in D[0,1]^d$ gilt:*

$$d_S(x_n, x) \overset{n\to\infty}{\longrightarrow} 0 \;\Leftrightarrow\; \exists(\lambda_n)_{n\in\mathbb{N}} \subset \Lambda : \;\|\lambda_n - id\|_\infty \overset{n\to\infty}{\longrightarrow} 0 \text{ und } \|x - x_n\lambda_n\|_\infty \overset{n\to\infty}{\longrightarrow} 0$$

Bemerkung 19. *Für $\lambda \in \Lambda$ existiert die Umkehrabbildung λ^{-1} und es gilt $\lambda^{-1} \in \Lambda$.*

Beweis. siehe Anhang □

Definition 20. *Mit $C[0,1]^d \subset D[0,1]^d$ bezeichnen wir die Menge der stetigen Funktionen auf $[0,1]^d$. Wenn nicht explizit anders angegeben, so fassen wir $C[0,1]^d$ immer als normierten Raum mit der Supremumsnorm*

$$\|x\|_\infty := \sup_{\mathbf{t}\in[0,1]^d} |x(\mathbf{t})|$$

auf. Die Borel σ-Algebra auf $C[0,1]^d$ (mit der durch die Supremumsnorm induzierten Topologie) bezeichnen wir mit $\mathcal{C}_d$.

Folgendes Lemma wird für unsere weiteren Überlegungen zur Verteilungskonvergenz von Prozessen in $D[0,1]^d$ sehr nützlich sein. Es besagt, dass Konvergenz bzgl. der Metrik d_S i. Allg. schwächer als Konvergenz bzgl. der Supremumsnorm ist, dass aber auf $C[0,1]^d$ beide Konvergenzarten äquivalent sind:

Lemma 21. *Für $x, x_1, \dots \in D[0,1]^d$ gilt:*

a) $\|x_n - x\|_\infty \overset{n\to\infty}{\longrightarrow} 0 \;\Rightarrow\; d_S(x_n, x) \overset{n\to\infty}{\longrightarrow} 0,$

b) Ist $x \in C[0,1]^d$, so folgt aus $d_S(x_n, x) \overset{n\to\infty}{\longrightarrow} 0$ auch $\|x_n - x\|_\infty \overset{n\to\infty}{\longrightarrow} 0$.

Beweis. Der Beweis lässt sich analog zum in Billingsley (1968) auf den Seiten 111 und 112 gelieferten Beweis führen, siehe Anhang. □

In der klassischen Theorie der Verteilungskonvergenz, wie sie zum Beispiel in Billingsley (1968) und Prokhorov (1956) behandelt wird, werden meist polnische Räume betrachtet. Das ist auch hier der Fall (vgl. Neuhaus (1971), S. 1289, Neuhaus (1969), Lemma 2.2, Sätze 2.4 und 2.5):

Lemma 22. *$D[0,1]^d$ ist mit der S-Topologie polnisch.*

Im Folgenden sei mit $D[0,1]^d$, wenn nicht explizit anders gesagt, immer der hier eingeführte topologische Raum mit der S-Topologie gemeint.
Definition 23. *Mit $\mathcal{D}_d$ bezeichnen wir die Borel σ-Algebra auf $D[0,1]^d$.*

2.2 Verteilungskonvergenz

2.2.1 Verteilungskonvergenz in $D[0,1]^d$

Nun beschäftigen wir uns mit der Verteilungskonvergenz in $D[0,1]^d$. Da wir im Folgenden stochastische Prozesse betrachten werden, die bzgl. $\mathcal{D}_d$ messbar sind, wollen wir diese σ-Algebra zunächst mit Hilfe der Projektionen näher betrachten:
Definition 24. *(vgl. Bickel und Wichura (1971), S. 1663) Für eine endliche Teilmenge $S \subset T$ sei $\pi_S : D[0,1]^d \to \mathbb{R}^S$ mit $\pi_S(x) := (x(\mathbf{s}))_{\mathbf{s}\in S}$ die Projektion auf S. Ist $S = \{\mathbf{s}\}$, so schreiben wir $\pi_{\mathbf{s}} := \pi_{\{\mathbf{s}\}}$.*

Die Projektionen sind messbar und charakterisieren die Borel σ-Algebra (siehe Bickel und Wichura (1971), S. 1662 und Neuhaus (1969), Satz 3.2, S. 28):
Lemma 25. *Für alle endlichen Mengen $S \subset T$ ist π_S $\mathcal{D}_d$-messbar. Für jede dichte Teilmenge $\tilde{T} \subset T$ ist $\mathcal{D}_d$ die kleinste σ-Algebra, für die alle Projektionen $\pi_{\mathbf{t}}$, $\mathbf{t} \in \tilde{T}$, messbar sind.*

Die folgenden Überlegungen stammen aus Neuhaus (1969) (S. 35). Lemma 25 liefert einen Zusammenhang zwischen $D[0,1]^d$-wertigen Zufallsvariablen und stochastischen Prozessen mit Pfaden in $D[0,1]^d$. Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Für eine $D[0,1]^d$-wertige Zufallsvariable X auf $(\Omega, \mathcal{A}, P)$ (d. h. eine Abbildung $X : (\Omega, \mathcal{A}) \to (D[0,1]^d, \mathcal{D}_d)$) ist $X_{\mathbf{t}} = \pi_{\mathbf{t}} \circ X$ nach Lemma 25 $\mathcal{A}-\mathcal{B}(\mathbb{R})$-messbar. Also ist $X_{\mathbf{t}}$ für alle $\mathbf{t} \in T$ eine reellwertige Zufallsvariable. Da $\mathcal{D}_d$ von den Projektionen erzeugt wird, folgt aus der $\mathcal{A}-\mathcal{B}(\mathbb{R})$-Messbarkeit von $X_{\mathbf{t}}$ für alle $\mathbf{t} \in T$ umgekehrt auch die $\mathcal{A}-\mathcal{D}_d$-Messbarkeit einer Abbildung $X : \Omega \to D[0,1]^d$. Eine $D[0,1]^d$-wertige Zufallsvariable X lässt sich als stochastischer Prozess gemäß Definition 1 auffassen. Genauer: Für X lässt sich die Familie $\tilde{X} := \{X(\mathbf{t})\}_{\mathbf{t}\in[0,1]^d}$ von reellen Zufallsvariablen auffassen als Funktion

$$\tilde{X} : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T), P^{\tilde{X}}),$$

wobei $\mathcal{B}(\mathbb{R}^T) \cap D[0,1]^d = \mathcal{D}_d$ und $P^{\tilde{X}}|_{\mathcal{D}_d} = P^X$. Andererseits lässt sich jeder stochastische Prozess $\tilde{X} := \{\tilde{X}_{\mathbf{t}}\}_{\mathbf{t}\in T}$, für den gilt, dass für fast alle $\omega \in \Omega$ die Abbildung $\mathbf{t} \mapsto \tilde{X}(\mathbf{t}, \omega)$ in $D[0,1]^d$ liegt, als $D[0,1]^d$-wertige Zufallsvariable X mit $P^X = P^{\tilde{X}}|_{\mathcal{D}_d}$ auffassen. Deshalb werden wir im Folgenden nicht zwischen $D[0,1]^d$-wertigen Zufallsvariablen und stochastischen Prozessen mit Pfaden in $D[0,1]^d$ unterscheiden und von $D[0,1]^d$-wertigen stochastischen Prozessen reden.

Wenn im Folgenden von einem stochastischen Prozess die Rede ist, ist immer ein bezüglich der Borel σ-Algebra messbarer Prozess gemeint.

Definition 26. *Seien (E, d_E) ein metrischer Raum und $X, X_1, X_2, \ldots$ E-wertige Zufallsvariablen. Wir sagen, $\underline{(X_n)_{n\in\mathbb{N}} \text{ konvergiert schwach gegen } X}$, geschrieben als*

$$X_n \xrightarrow{E} X,$$

wenn

$$E[f(X_n)] \to E[f(X)], \ n \to \infty$$

für alle d_E-stetigen beschränkten Funktionen $f : E \to \mathbb{R}$.

Definition 27. *Ist $E = \mathbb{R}^d$ $(d \in \mathbb{N})$ mit der euklidischen Metrik, so schreiben wir $X_n \xrightarrow{\mathcal{D}} X$ statt $X_n \xrightarrow{\mathbb{R}^d} X$.*

Wie in Bickel und Wichura (1971) (S. 1662) angemerkt, lässt sich die so definierte Verteilungskonvergenz auch wie folgt charakterisieren:

Lemma 28. *Seien (E, d_E) ein metrischer Raum mit Borel σ-Algebra $\mathcal{E}$ und $X, X_1, X_2, \ldots$ E-wertige Zufallsvariablen. Dann gilt*

$$X_n \xrightarrow{E} X$$

genau dann, wenn

$$f(X_n) \xrightarrow{\mathcal{D}} f(X)$$

für alle $\mathcal{E}$-messbaren Funktionen $f : E \to \mathbb{R}$, die stetig sind.

Beweis.

zu $\Rightarrow$: Dies folgt direkt aus dem Continuous Mapping Theorem (vgl. Klenke (2006), S. 245, Satz 13.25).

zu $\Leftarrow$: Zu zeigen ist, dass für alle stetigen, beschränkten Funktionen $f : E \to \mathbb{R}$ gilt $E[f(X_n)] \to E[f(X)]$. Sei f eine solche Funktion. Da f stetig ist, ist f auch bezüglich der Borel σ-Algebra messbar. Folglich gilt nach Voraussetzung $f(X_n) \xrightarrow{\mathcal{D}} f(X)$. Aufgrund der Beschränktheit von f gibt es ein $C > 0$, so dass für alle $n \in \mathbb{N}$ gilt $|f(X_n)| \leq C$. Somit ist die Folge $\{f(X_n)\}_{n\in\mathbb{N}}$ gleichgradig P-integrierbar und es folgt $E[f(X_n)] \to E[f(X)]$ (vgl. Billingsley (1995), S. 338, Theorem 25.12). $\square$

Bemerkung 29. *Da das Continuous Mapping Theorem (siehe Klenke (2006), S. 245, Satz 13.25) nur die fast sichere Stetigkeit fordert, könnte man in Lemma 28 statt zu fordern, dass $f : E \to \mathbb{R}$ stetig ist, auch fordern, dass f P_X-fast sicher stetig ist.*

Bemerkung 30. *In Wichura (1969) (Definition 1, S. 682) wird eine etwas andere Definition von Verteilungskonvergenz in $D[0,1]^d$ gegeben, als die hier benutzte Definition 26. (Die Definition in Wichura (1969) bezieht sich auf Wahrscheinlichkeitsmaße. Um den Kontrast zur hier verwendeten Definition 26 besser zu sehen, wurde sie im Folgenden für Zufallsvariablen umformuliert.) Wichura (1969) betrachtet den Raum $D[0,1]^d$ mit der von den Projektionen erzeugten σ-Algebra, welche nach Lemma 25 identisch mit der hier betrachteten σ-Algebra $\mathcal{D}_d$ ist. Anders als in Definition 26, bezeichnet Wichura eine Folge von $D[0,1]^d$-wertigen Zufallsvariablen $\{X_n\}_{n\in\mathbb{N}}$ als „in der U-Topologie“ konvergent gegen eine $D[0,1]^d$-wertige*

Zufallsvariable X, wenn für alle $\mathcal{D}_d - \mathcal{B}(\mathbb{R})$-messbaren, bezüglich der Supremumsnorm auf $D[0,1]^d$ stetigen Funktionen $f : D[0,1]^d \to \mathbb{R}$ gilt $f(X_n) \xrightarrow{\mathcal{D}} f(X)$. Dies ermöglicht es, die Einführung einer neuen Metrik auf $D[0,1]^d$ zu umgehen. Eine „in der U-Topologie" konvergente Folge ist auch bezüglich der hier betrachteten Definition konvergent: Nach Lemma 28 reicht es zu zeigen, dass für alle $\mathcal{D}_d - \mathcal{B}(\mathbb{R})$-messbaren Funktionen $f : D[0,1]^d \to \mathbb{R}$, die bezüglich der Metrik d_S stetig sind, gilt $f(X_n) \xrightarrow{\mathcal{D}} f(X)$. Eine bezüglich d_S stetige Funktion ist aber wegen Lemma 21 insbesondere auch bezüglich der Supremumsnorm stetig, so dass die Konvergenz von $f(X_n)$ gegen $f(X)$ aus der Konvergenz „in U-Topologie" folgt.

Satz 31. *(E, d_E) sei ein metrischer Raum und $X, X_1, \ldots$ seien $D[0,1]^d$-wertige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$, für die $X_n \xrightarrow{D[0,1]^d} X$ gilt. Weiter sei $f : D[0,1]^d \to E$ eine Funktion, für die $f(X), f(X_1), \ldots$ E-wertige Zufallsvariablen sind. f sei bezüglich der Supremumsnorm auf $D[0,1]^d$ stetig. Gilt $P(X \in C[0,1]^d) = 1$, so ist f auch bezüglich der S-Metrik d_S P_X-fast sicher stetig und es folgt:*

$$f(X_n) \xrightarrow{E} f(X) \tag{2.1}$$

Beweis. Es sei

$$A := \{\omega \in \Omega : X(\omega) \in C[0,1]^d\}.$$

Dann gilt für $\omega \in A$: Ist $(x_n)_{n\in\mathbb{N}}$ eine Folge in $D[0,1]^d$ mit $d_S(x_n, X(\omega)) \to 0$, so folgt wegen $\omega \in A$ nach Lemma 21 auch

$$\|x_n - X(\omega)\|_\infty \to 0.$$

Da f bei $X(\omega)$ stetig bezüglich der Supremumsnorm ist, folgt:

$$d_E(f(x_n), f(X(\omega))) \to 0$$

Somit ist f auch bezüglich d_S stetig in $X(\omega)$. Folglich gilt:

$$P_X(f \text{ ist stetig bezüglich } d_S) = P(f \text{ ist stetig in } X \text{ bezüglich } d_S) \geq P(A) = 1$$

(2.1) folgt nun aus Satz 32. □

Üblicherweise wird für das Continuous Mapping Theorem vorausgesetzt, dass es sich bei der betrachteten Abbildung h um eine messbare Abbildung handelt (vgl. Billingsley (2008), Klenke (2006)). Diese Voraussetzung lässt sich auf die Forderung, dass $h(X), h(X_1), \ldots$ Zufallsvariablen sind, abschwächen:

Satz 32. *Es seien (E, d_E) und $(E', d_{E'})$ metrische Räume mit Borel σ-Algebren $\mathcal{E}$ und $\mathcal{E}'$. $X, X_1, X_2, \ldots$ seien E-wertige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ (d. h. $X, X_n : (\Omega, \mathcal{A}) \to (E, \mathcal{E})$, $n \in \mathbb{N}$). Weiter sei $h : E \to E'$ eine P_X-fast sicher stetige Funktion (d. h. $P(X \in D_h) = 0$, wobei D_h die Menge der Unstetigkeitsstellen von h sei), für die $h(X), h(X_1), h(X_2), \ldots$ Zufallsvariablen auf $(\Omega, \mathcal{A}, P)$ sind. Dann folgt aus der Verteilungskonvergenz $X_n \xrightarrow{E} X$, dass auch*

$$h(X_n) \xrightarrow{E'} h(X).$$

Beweis. Billingsley (2008) (Theorem 2.7, S. 21) beweist dieses Theorem unter der Voraussetzung, dass die Funktion h messbar ist. Um zu zeigen, dass diese Voraussetzung auf die Messbarkeit von $h(X), h(X_1), \ldots$ abgeschwächt werden kann, führen wir hier den Beweis, so wie er in Billingsley (2008) geführt wird, und ergänzen Anmerkungen zur Messbarkeit der betrachteten Ereignisse. Nach dem Portemanteau Theorem (siehe Klenke (2006), Satz 13.16, S. 242) genügt es zu zeigen:

$$\limsup_{n\to\infty} P(h(X_n)^{-1}(F)) \leq P(h(X)^{-1}(F)) \ \forall F \subset E' \text{ abgeschlossen}$$

Sei $F \subset E'$ abgeschlossen. Sei D_h^c das Komplement von D_h und $x \in \overline{h^{-1}(F)} \cap D_h^c$. Dann gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ in $h^{-1}(F)$ mit $x_n \to x$. Dann sind $h(x_n) \in F$ für alle $n \in \mathbb{N}$ und wegen der Stetigkeit von h auf D_h^c konvergieren die $h(x_n)$ gegen $h(x)$. Nach Definition des Abschlusses ist folglich $h(x) \in \overline{F} = F$. Somit gilt:

$$\overline{h^{-1}(F)} \cap D_h^c \subset h^{-1}(F)$$

Wegen $X_n \overset{E}{\to} X$ und $P(X \in D_h^c) = 1$ gilt also:

$$\begin{aligned}
\limsup_{n\to\infty} P(h(X_n) \in F) &\leq \limsup_{n\to\infty} P(X_n \in \overline{h^{-1}(F)}) \\
&\leq P(X \in \overline{h^{-1}(F)}) \\
&= P(X \in \overline{h^{-1}(F)} \cap D_h^c) \\
&\leq P(h(X) \in F)
\end{aligned}$$

Anmerkung zur Messbarkeit:

- F ist als abgeschlossene Menge eine Borelmenge, also sind $h(X_n)^{-1}(F)$ und $h(X)^{-1}(F)$ messbar.
- $\overline{h^{-1}(F)}$ ist abgeschlossen in E, also Borel-messbar. Somit sind $X_n^{-1}(\overline{h^{-1}(F)})$ und $X^{-1}(\overline{h^{-1}(F)})$ messbar. □

Bemerkung 33. *Seien (E, d_E) und $(E', d_{E'})$ metrische Räume. Dann ist für jede Funktion $f : E \to E'$ die Menge der Stetigkeitsstellen Borel-messbar.*
Beweis. siehe Anhang □

Das folgende Lemma wurde aus Klenke (2006) entnommen:
Lemma 34. *(siehe Klenke (2006), Satz 13.18, S.243)*
Seien $X, X_1, X_2, \ldots$ und $Y_1, Y_2, \ldots$ Zufallsvariablen mit Werten in einem metrischen Raum (E, d_E) und $X_n \overset{E}{\to} X$, sowie $d_E(X_n, Y_n) \overset{n\to\infty}{\longrightarrow} 0$ stochastisch. Dann gilt $Y_n \overset{E}{\to} X$.

2.2.2 Konvergenzkriterium von Bickel und Wichura

Im Folgenden wollen wir ein nützliches Kriterium zum Beweis der Verteilungskonvergenz eines stochastischen Prozesses in $D[0,1]^d$ vorstellen. Dazu zunächst einige

Notationen und Sprechweisen (vgl. Bickel und Wichura (1971), S. 1658 und 1663): Ein *Block* B in $T = [0,1]^d$ ist eine Teilmenge von T der Form $(\mathbf{s}, \mathbf{t}]$ mit $\mathbf{s}, \mathbf{t} \in T$. Die *p-te Seite* von B ist $\prod\limits_{\tilde{p} \neq p} (s_{\tilde{p}}, t_{\tilde{p}}]$. Disjunkte Blöcke B und C heißen *p-Nachbarn*, wenn sie an einander grenzen und die gleiche p-te Seite haben, und *Nachbarn*, wenn sie p-Nachbarn für ein $p \in \{1, \ldots, d\}$ sind. Für einen Block $B = (\mathbf{s}, \mathbf{t}]$ sei der *Zuwachs* eines Prozesses X auf B definiert als

$$X(B) := \sum_{\boldsymbol{\varepsilon} \in \{0,1\}^d} (-1)^{d - \sum_{i=1}^d \varepsilon_i} X(\mathbf{s} + \boldsymbol{\varepsilon}(\mathbf{t} - \mathbf{s})). \tag{2.2}$$

Für zwei benachbarte Blöcke B und C sei

$$m(B, C) := \min\{|X(B)|, |X(C)|\}.$$

Seien $\beta > 1$, $\gamma > 0$ und μ ein endliches, nichtnegatives Maß auf T. Wir sagen, dass (X, μ) Bedingung (β, γ) erfüllt (geschrieben als $(X, \mu) \in \mathcal{C}(\beta, \gamma)$), falls für alle $\lambda > 0$ und alle benachbarten Blöcke $B, C \subset T$ gilt:

$$P(m(B, C) \geq \lambda) \leq \lambda^{-\gamma} (\mu(B \cup C))^\beta \tag{2.3}$$

Für eine Abbildung $x \in D[0,1]^d$ und $1 \leq p \leq d$ sei für $t \in [0,1]$ die Abbildung $x_t^{(p)} : [0,1]^{d-1} \to \mathbb{R}$ definiert als

$$x_t^{(p)}(\mathbf{s}) := x(s_1, \ldots, s_{p-1}, t, s_{p+1}, \ldots, s_d).$$

Wir nennen x *stetig am oberen Rand von T*, wenn für alle $t \in [0,1]$ und $p \in \{1, \ldots, d\}$ gilt:

$$\lim_{t \nearrow 1} x_t^{(p)} = x_1^{(p)}$$

Für eine Teilmenge $S = S_1 \times \cdots \times S_d$ von $[0,1]^d$ nennen wir

$$\bigcup_{1 \leq p \leq d} (S_1 \times \cdots \times S_{p-1} \times \{0\} \times S_{p+1} \times \cdots \times S_d)$$

den *unteren Rand* von S. $\mathcal{T}$ sei die Menge der Teilmengen von T, die die Form $U_1 \times \cdots \times U_d$ haben, wobei jedes U_p Null und Eins enthält und ein abzählbares Komplement hat.
Wie in Lemma 22 angemerkt, ist $D[0,1]^d$ polnisch, so dass das bekannte Prokhorov Theorem (siehe Billingsley (1968), Theorem 6.1 und Theorem 6.2) anwendbar ist: Dieses besagt, dass eine notwendige und hinreichende Bedingung für die Verteilungskonvergenz einer Folge von Prozessen $\{X_n\}_{n \in \mathbb{N}}$ mit $X_n = \{X_n(\mathbf{t})\}_{\mathbf{t} \in [0,1]^d}$ in $D[0,1]^d$ ist, dass die Folge der zugehörigen Wahrscheinlichkeitsmaße straff in $D[0,1]^d$ ist, und dass die endlich-dimensionalen Verteilungen der Prozesse gegen die eines Grenzprozesses konvergieren. Dabei bedeutet Straffheit, dass es für alle $\varepsilon > 0$ eine kompakte Teilmenge $K \subset D[0,1]^d$ gibt, so dass $P(X_n \in K) \geq 1 - \varepsilon$ für alle $n \in \mathbb{N}$ gilt. Die Straffheit wird bei Bickel und Wichura (1971) ähnlich wie bei

Billingsley (1968) mit Hilfe eines Stetigkeitsmoduls bewiesen. Das Stetigkeitsmodul auf $D[0,1]^d$ ist eine Abbildung $\omega_\delta^{''} : D[0,1]^d \to \mathbb{R}$ mit

$$\omega_\delta^{''}(x) := \max_{1 \le p \le d} \omega_\delta^{''(p)}(x),$$

wobei $\delta > 0$ und

$$\omega_\delta^{''(p)}(x) := \sup\{\min(\|x_t^{(p)} - x_s^{(p)}\|_\infty, \|x_u^{(p)} - x_t^{(p)}\|_\infty) : \; s \le t \le u, u - s \le \delta\}$$

ist (siehe Bickel und Wichura (1971), S. 1663). Der folgende Satz ist eine Mischung aus zwei Sätzen aus Bickel und Wichura (1971) (Theorem 3 und ein Korollar zu Theorem 2 aus Bickel und Wichura (1971), S. 1664, 1665):

Satz 35. *Es seien X_n, $n \ge 1$, und X $D[0,1]^d$-wertige stochastische Prozesse. Außerdem sei X fast sicher stetig auf dem oberen und $X_n = 0$ auf dem unteren Rand von T. Dann konvergiert $X_n \overset{D[0,1]^d}{\longrightarrow} X$, wenn es eine Menge $\tau \in \mathcal{T}$ gibt, so dass*

i) $\pi_S(X_n) \to \pi_S(X)$, *für alle endlichen Teilmengen S von τ und*

ii) $\lim\limits_{\delta \to 0} \limsup\limits_{n \to \infty} \omega_\delta^{''}(X_n) = 0.$

(ii) ist erfüllt, wenn es Konstanten $\beta > 1$, $\gamma > 0$ und ein endliches nichtnegatives Maß μ auf T mit stetigen Randverteilungen gibt, so dass $(X_n, \mu) \in \mathcal{C}(\beta, \gamma)$ für alle $n \in \mathbb{N}$.

Bemerkung 36. *Benutzt man statt Definition 26 die in Bemerkung 30 eingeführte Definition der Verteilungskonvergenz „in der U-Topologie", so liefert Theorem 2 von Wichura (1969) (siehe S. 683) eine analoge Aussage zu Satz 35 für den Fall, dass die Grenzvariable fast sicher in $C[0,1]^d$ liegt. Dabei wird Bedingung (ii) ersetzt durch die entsprechende Bedingung für die Abbildung*

$$\omega_\delta(x) := \sup_{\mathbf{s},\mathbf{t} \in [0,1]^d, \; \|\mathbf{t}-\mathbf{s}\| < \delta} |x(\mathbf{t}) - x(\mathbf{s})|, \; x \in D[0,1]^d.$$

Bickel und Wichura (1971) schreiben in einer Bemerkung zu ihrem Theorem 3, dass sich die Bedingung $(X_n, \mu) \in \mathcal{C}(\beta, \gamma)$ aus Satz 35 auch abschwächen lässt:

Bemerkung 37. *Statt $(X_n, \mu) \in \mathcal{C}(\beta, \gamma)$ reicht es, wenn es für jedes $n \in \mathbb{N}$ eine Teilmenge $T^n = T_1^n \times \cdots \times T_d^n$ von T mit folgenden Eigenschaften gibt:*

1.) $0, 1 \in T_p^n$ *für jedes $n \in \mathbb{N}$ und $1 \le p \le d$,*

2.) $\omega_\delta^{''}(X_n)$ *kann mit T^n (statt T) als Zeitparametermenge berechnet werden,*

3.) T^n *wird dicht in T für $n \to \infty$, und*

4.) Bedingung (β, γ) ist für alle Blöcke mit Endpunkten in T^n erfüllt.

Hieraus lässt sich folgendes Lemma ableiten (vgl. Lavancier (2005), Korollar 3, S. 13):

Lemma 38. *Seien* $\{X_{\mathbf{j}}\}_{\mathbf{j}\in\mathbb{Z}^d}$ *ein stochastischer Prozess auf einem Wahrscheinlichkeitsraum* $(\Omega, \mathcal{A}, P)$ *und* $S_{\mathbf{k}} := \sum_{\underline{\mathbf{1}}\leq\mathbf{j}\leq\mathbf{k}} X_{\mathbf{j}}$ *die zugehörigen Partialsummen. Für* $n \in \mathbb{N}$ *sei der Prozess* $\{\tilde{S}_n(\mathbf{t})\}_{\mathbf{t}\in[0,1]^d}$ *mit*

$$\tilde{S}_n(\mathbf{t}) := \frac{1}{n^{d/2}} \sum_{\underline{\mathbf{1}}\leq\mathbf{j}\leq\lfloor n\mathbf{t}\rfloor} X_{\mathbf{j}}$$

definiert. Weiter gebe es $p > 2$*, so dass für alle* $\mathbf{k}, \mathbf{m} \in \mathbb{N}^d$ *mit* $\mathbf{k} < \mathbf{m} \leq \underline{\mathbf{n}}$

$$E\left[\left|\frac{1}{n^{d/2}} \sum_{\mathbf{k}<\mathbf{j}\leq\mathbf{m}} X_{\mathbf{j}}\right|^p\right] \leq C\left(\frac{[\mathbf{m}-\mathbf{k}]}{n^d}\right)^{p/2} \tag{2.4}$$

gilt, wobei $C > 0$ *unabhängig von* n*,* $\mathbf{k}$ *und* $\mathbf{m}$ *gewählt werden kann. Dann erfüllen* $T^n := \left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\right\}^d \subset [0,1]^d$ *und* $\{\tilde{S}_n(\mathbf{t})\}_{\mathbf{t}\in[0,1]^d}$ *die Bedingungen aus Bemerkung 37 für* $\mu := C^{2/p}\lambda^d$*,* $\beta := p/2$ *und* $\gamma := p$*, wobei* λ^d *das Lebesgue-Maß auf* $[0,1]^d$ *sei.*

Beweis. $\mu := C^{2/p}\lambda^d$ ist ein endliches nichtnegatives Maß auf $[0,1]^d$ und T^n $(n \in \mathbb{N})$ erfüllen offensichtlich die Bedingungen 1.) und 3.) aus Bemerkung 37. Da $\tilde{S}_n(\mathbf{t}) = \tilde{S}_n(\frac{\mathbf{i}}{n})$ für $\mathbf{t} \in \left[\frac{\mathbf{i}}{n}, \frac{\mathbf{i}+\underline{\mathbf{1}}}{n}\right) \cap [0,1]^d$, $\mathbf{i} \in \{0, \ldots, n\}^d$, ist auch Bedingung 2.) erfüllt. Zu prüfen ist also noch Bedingung 4.). Seien $n \in \mathbb{N}$ und B_n ein Block mit Endpunkten in T_n, also

$$B_n = (\mathbf{s}, \mathbf{t}] = \left(\frac{\mathbf{k}}{n}, \frac{\mathbf{m}}{n}\right] \subset [0,1]^d$$

mit $\mathbf{k}, \mathbf{m} \in \{0, \ldots, n\}^d$, $\mathbf{k} < \mathbf{m}$. Für den Zuwachs gilt (vgl. Bemerkung 44):

$$\tilde{S}_n(B_n) = \sum_{\boldsymbol{\varepsilon}\in\{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} \tilde{S}_n(\mathbf{s} + \boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s})) = \frac{1}{n^{d/2}} \sum_{\mathbf{k}<\mathbf{j}\leq\mathbf{m}} X_{\mathbf{j}}$$

(2.4) liefert nun für ein $p > 2$:

$$E\left[\left|\tilde{S}_n(B_n)\right|^p\right] \leq C\left(\frac{[\mathbf{m}-\mathbf{k}]}{n^d}\right)^{p/2} = \mu(B_n)^{p/2}$$

Seien $B_n = (\frac{\mathbf{k}}{n}, \frac{\mathbf{m}}{n}]$ und $F_n = (\frac{\tilde{\mathbf{k}}}{n}, \frac{\tilde{\mathbf{m}}}{n}]$ zwei benachbarte Blöcke in $[0,1]^d$ mit Endpunkten in T_n. Dann liefert obige Ungleichung nach einer Anwendung der Markovund Cauchy-Schwarz-Ungleichungen für $\lambda > 0$:

$$\begin{aligned}
&P(\min\{|\tilde{S}_n(B_n)|, |\tilde{S}_n(F_n)|\} \geq \lambda) \\
\leq&\lambda^{-p} E[\min\{|\tilde{S}_n(B_n)|, |\tilde{S}_n(F_n)|\}^p] \\
\overset{\star}{\leq}&\lambda^{-p} \left\{ E\left[|\tilde{S}_n(B_n)|^p\right] E\left[|\tilde{S}_n(F_n)|^p\right]\right\}^{1/2} \\
\leq&\lambda^{-p} \ \left\{\mu(B_n)^{p/2}\mu(F_n)^{p/2}\right\}^{1/2} \\
=&\lambda^{-p} \ \left[\{\mu(B_n)\mu(F_n)\}^{1/2}\right]^{p/2} \\
\leq&\lambda^{-p} \ [\mu(B_n) + \mu(F_n)]^{p/2} \\
=&\lambda^{-p} \ \mu(B_n \cup F_n)^{p/2},
\end{aligned}$$

wobei für $\star$ die Cauchy-Schwarz-Ungleichung und die Tatsache, dass für $x, y \geq 0$ gilt

$$\min\{x, y\} \leq \sqrt{x}\sqrt{y},$$

benutzt wurde. Somit ist die Bedingung 4.) erfüllt. □

3 Testverfahren

3.1 Die Teststatistik

Nun wollen wir die von uns betrachtete Statistik für den Test von H_0 gegen H_A einführen. Diese ist eine Variante für reellwertige Zufallsfelder der von Jarušková und Piterbarg (2011) benutzten Statistik. Die Statistik ergibt sich aus folgender Heuristik: Zunächst nehmen wir an, die $Y_{\mathbf{j}}$ ($\mathbf{j} \in \mathbb{Z}^d$) wären i. i. d. und $\mathbf{k} < \mathbf{m}$ die (bekannten) Changepoints. Sei $R_{\mathbf{k},\mathbf{m}} := (\mathbf{k}, \mathbf{m}]$. Wir wollen als Teststatistik diejenige Statistik wählen, die den Pseudo-Log-Likelihood-Quotienten maximiert. Da wir die Verteilung der $Y_{\mathbf{j}}$ ($\mathbf{j} \in \mathbb{Z}^d$) nicht kennen, bestimmen wir die Statistik für standardnormalverteilte Zufallsvariablen $Y_{\mathbf{j}}$. Nun kann man die Beobachtungen in zwei Klassen aufteilen, die sich nur in ihrem Erwartungswert unterscheiden: Beobachtungen mit Index in $R_{\mathbf{k},\mathbf{m}}$ und der Rest. Definiere für $\mathbf{k}_1, \mathbf{k}_2 \in \mathbb{N}^d$ die Größen $S_{\mathbf{k}_1} := \sum_{\underline{\mathbf{1}} \le \mathbf{j} \le \mathbf{k}_1} X_{\mathbf{j}}$ und $S_{\mathbf{k}_1,\mathbf{k}_2} := \sum_{\mathbf{k}_1 < \mathbf{j} \le \mathbf{k}_2} X_{\mathbf{j}}$, und damit $\bar{X} := n^{-d} S_{\underline{\mathbf{n}}}$, $\bar{X}_{\mathbf{k},\mathbf{m}} := \frac{1}{[\mathbf{m}-\mathbf{k}]} S_{\mathbf{k},\mathbf{m}}$ und$\bar{X}^0_{\mathbf{k},\mathbf{m}} = \frac{1}{n^d - [\mathbf{m}-\mathbf{k}]} (S_{\underline{\mathbf{n}}} - S_{\mathbf{k},\mathbf{m}})$. Da die $Y_{\mathbf{j}}$ nach Annahme i. i. d normalverteilt sind, ist die gemeinsame Dichte von $X_{\mathbf{j}_1}, \ldots, X_{\mathbf{j}_p}$ ($p \in \mathbb{N}, \mathbf{j}_1, \ldots, \mathbf{j}_p \in \mathbb{Z}$) das Produkt von Normalverteilungsdichten. Elementare Umformungen liefern für den Log-Likelihood-Quotienten:

$$\begin{aligned}
\Lambda_{\mathbf{k},\mathbf{m}} &= \log \frac{\sup\limits_{a,b \in \mathbb{R}} \prod\limits_{\mathbf{j} \in R_{\mathbf{k},\mathbf{m}}} \Phi(X_{\mathbf{j}} - a) \prod\limits_{\mathbf{j} \notin R_{\mathbf{k},\mathbf{m}}} \Phi(X_{\mathbf{j}} - b)}{\sup\limits_{a \in \mathbb{R}} \prod\limits_{\underline{\mathbf{1}} \le \mathbf{j} \le \underline{\mathbf{n}}} \Phi(X_{\mathbf{j}} - a)} \\
&= \log \frac{\prod\limits_{\mathbf{j} \in R_{\mathbf{k},\mathbf{m}}} \Phi(X_{\mathbf{j}} - \bar{X}_{\mathbf{k},\mathbf{m}}) \prod\limits_{\mathbf{j} \notin R_{\mathbf{k},\mathbf{m}}} \Phi(X_{\mathbf{j}} - \bar{X}^0_{\mathbf{k},\mathbf{m}})}{\prod\limits_{\underline{\mathbf{1}} \le \mathbf{j} \le \underline{\mathbf{n}}} \Phi(X_{\mathbf{j}} - \bar{X})} \\
&= \frac{1}{2} \frac{n^d}{[\mathbf{m}-\mathbf{k}](n^d - [\mathbf{m}-\mathbf{k}])} \left(S_{\mathbf{k},\mathbf{m}} - \frac{[\mathbf{m}-\mathbf{k}]}{n^d} S_{\underline{\mathbf{n}}} \right)^2 \\
&= \frac{1}{2} \left(n^{-d/2} \frac{\left| \sum\limits_{\mathbf{k} < \mathbf{j} \le \mathbf{m}} \left(X_{\mathbf{j}} - \bar{X} \right) \right|}{\sqrt{\frac{[\mathbf{m}-\mathbf{k}]}{n^d} \left(1 - \frac{[\mathbf{m}-\mathbf{k}]}{n^d} \right)}} \right)^2
\end{aligned}$$

Dabei wurde ausgenutzt, dass $\bar{X}_n$, $\bar{X}_{\mathbf{k},\mathbf{m}}$ und $\bar{X}^0_{\mathbf{k},\mathbf{m}}$ die Maximum-Likelihood-Schätzer für die Erwartungswerte, gegeben i. i. d. normalverteilte Beobachtungen

$\{X_{\mathbf{j}} : \mathbf{j} \in \{1, \ldots, n\}^d\}$, $\{X_{\mathbf{j}} : \mathbf{j} \in R_{\mathbf{k},\mathbf{m}}\}$ bzw. $\{X_{\mathbf{j}} : \mathbf{j} \in \{1, \ldots, n\}^d \setminus R_{\mathbf{k},\mathbf{m}}\}$ sind. Sind die Changepoints $\mathbf{k}, \mathbf{m}$ nicht bekannt, so liegt es nahe, die Pseudo-Maximum-Likelihood Statistik über alle in Frage kommenden Rechtecke $R_{\mathbf{k},\mathbf{m}} = (\mathbf{k}, \mathbf{m}] \subset \{1, \ldots, n\}^d$ zu maximieren. Die Maximierung über alle möglichen Rechtecke führt aber bereits für $d = 1$ und i. i. d. Zufallsvariablen zu Problemen bei sehr kleinen oder großen Rechtecken: Wie in Antoch und Hušková (1996) angemerkt, divergiert die Teststatistik in diesem Fall unter der Nullhypothese stochastisch gegen Unendlich. Wir betrachten daher im Folgenden die „getrimmte“ Teststatistik

$$T_n := T_n(\alpha, \beta) := \hat{\sigma}_n^{-1} n^{-d/2} \max_{\substack{\mathbf{0} \leq \mathbf{k} < \mathbf{m} \leq \mathbf{n} \\ \lfloor \alpha n^d \rfloor \leq [\mathbf{m}-\mathbf{k}] \leq \lfloor (1-\beta) n^d \rfloor}} \frac{\left| \sum_{\mathbf{k} < \mathbf{j} \leq \mathbf{m}} \left(X_{\mathbf{j}} - \bar{X} \right) \right|}{\sqrt{\frac{[\mathbf{m}-\mathbf{k}]}{n^d} \left(1 - \frac{[\mathbf{m}-\mathbf{k}]}{n^d} \right)}},$$

mit $0 < \alpha < 1 - \beta < 1$. Große Werte dieser Statistik sprechen für die Existenz eines epidemischen Strukturbruchs, kleine Werte für die Nullhypothese. Dieses Vorgehen hat den offensichtlichen Nachteil, dass Strukturbrüche außerhalb des betrachteten Bereichs nicht entdeckt werden können und die Parameter α und β subjektiv gewählt werden müssen. Andererseits kann auf diese Weise mit Hilfe des schwachen Invarianzprinzips die Grenzverteilung der Statistik als Supremum eines Funktionals des Wiener Feldes auf einem kompakten Bereich hergeleitet werden, um die asymptotischen kritischen Werte des Tests zu bestimmen. Daher setzen wir voraus, dass mindestens ein bestimmter Anteil (αn^d) der Beobachtungen im epidemischen Bereich und mindestens ein bestimmter Anteil (βn^d) außerhalb liegt, und betrachten die eingeschränkte Alternative

$$\begin{aligned} H_{\alpha,\beta} : \exists \underline{\mathbf{1}} \leq \mathbf{k}_0 < \mathbf{m}_0 < \underline{\mathbf{n}},\ \lfloor \alpha n^d \rfloor \leq [\mathbf{k}_0 - \mathbf{m}_0] \leq \lfloor (1-\beta) n^d \rfloor : \\ X_{\mathbf{k}} = Y_{\mathbf{k}} + \mu + \delta I_{\{\mathbf{k}_0 < \mathbf{k} \leq \mathbf{m}_0\}}\ \forall \mathbf{k} \in \{1, \ldots, n\}^d \end{aligned}$$

Bemerkung 39. *Die Statistik ist unabhängig von μ.*
Beweis. siehe Anhang □

3.2 Verhalten unter der Nullhypothese

Nun untersuchen wir das Verhalten von $T_n = T_n(\alpha, \beta)$ unter der Nullhypothese. In diesem Fall ist

$$X_{\mathbf{k}} = Y_{\mathbf{k}} + \mu$$

für alle $\mathbf{k} \in \{1, \ldots, n\}^d$ ($d \in \mathbb{N}$). Da die Statistik unabhängig von μ ist, gehen wir im Folgenden o. E. davon aus, dass $\mu = 0$ ist.

Ziel dieses Abschnittes ist der Beweis des folgenden Satzes:

Satz 40. *Ist Annahme 1 erfüllt, so gilt unter der Nullhypothese für $n \to \infty$:*

$$T_n \xrightarrow{\mathcal{D}} \sup_{\underline{\mathbf{0}} \leq \mathbf{s} < \mathbf{t} \leq \underline{\mathbf{1}},\ \alpha \leq [\mathbf{t}-\mathbf{s}] \leq 1-\beta} \frac{\left| \sum_{\boldsymbol{\varepsilon} \in \{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} W(\mathbf{s} + \boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s})) - [\mathbf{t}-\mathbf{s}] W(\underline{\mathbf{1}}) \right|}{\sqrt{[\mathbf{t}-\mathbf{s}](1-[\mathbf{t}-\mathbf{s}])}},$$

wobei $\{W(\mathbf{t})\}_{\mathbf{t} \in [0,1]^d}$ ein Wiener Feld sei.

Für $d = 1, 2$ und unabhängig und identisch verteilte $\{Y_\mathbf{k}\}_{\mathbf{k} \in \mathbb{Z}^d}$ findet sich eine entsprechende Aussage im Aufsatz von Jarušková und Piterbarg (2011). Dort wird ohne detaillierten Beweis als Begründung für die Konvergenz auf ein Kriterium von Bickel und Wichura (1971) (vgl. Satz 35) verwiesen. Dies wurde für diese Arbeit als Anlass genommen, um die *i. i. d.*-Annahmen fallen zu lassen und lediglich die Verteilungskonvergenz des Partialsummenprozesses zu fordern.
Im Folgenden betrachten wir zunächst einige Hilfslemmata. Satz 40 wird sich dann als Anwendung der Lemmata ergeben.
Lemma 41. *Es gilt:*

$$\sup_{\mathbf{s},\mathbf{t} \in [0,1]^d} \left| \frac{[\lfloor n\mathbf{t} \rfloor - \lfloor n\mathbf{s} \rfloor]}{n^d} - [\mathbf{t}-\mathbf{s}] \right| \overset{n\to\infty}{\longrightarrow} 0$$

Beweis. siehe Anhang □
Lemma 42. *Sei $x : [0,1]^d \to \mathbb{R}$ eine stetige Funktion und*

$$K_d := \{(\mathbf{s},\mathbf{t}) \in [0,1]^{2d} :\ \underline{\mathbf{0}} \leq \mathbf{s} \leq \mathbf{t} \leq \underline{\mathbf{1}}\}.$$

Dann ist

$$g_x(\mathbf{s},\mathbf{t}) := \left(\sum_{\boldsymbol{\varepsilon} \in \{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} x(\mathbf{s} + \boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s})) - [\mathbf{t}-\mathbf{s}] x(\underline{\mathbf{1}}) \right) I_{K_d}(\mathbf{s},\mathbf{t}) \qquad (3.1)$$

stetig auf $[0,1]^{2d}$.
Beweis. g_x ist auf K_d als Verknüpfung von stetigen Abbildungen stetig: Die Abbildungen

$$h_{1,\boldsymbol{\varepsilon}}(\mathbf{s},\mathbf{t}) := \mathbf{s} + \boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s})$$

und

$$h_2(\mathbf{s},\mathbf{t}) := [\mathbf{t}-\mathbf{s}] = \prod_{i=1}^d (t_i - s_i)$$

sind offensichtlich auf $[0,1]^{2d}$ stetig und g_x hat auf K_d die Form

$$g_x(\mathbf{s},\mathbf{t}) = \sum_{\boldsymbol{\varepsilon} \in \{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} x(h_{1,\boldsymbol{\varepsilon}}(\mathbf{s},\mathbf{t})) - h_2(\mathbf{s},\mathbf{t}) x(\underline{\mathbf{1}}).$$

Nun zeigen wir, dass g_x konstant gleich Null und somit stetig auf $(\mathbf{s},\mathbf{t}) \in \overline{[0,1]^{2d} \setminus K_d}$ ist: Für $(\mathbf{s},\mathbf{t}) \notin K_d$ ist $g_x(\mathbf{s},\mathbf{t}) = 0$ per definitionem. Es gilt

$$\overline{[0,1]^{2d} \setminus K_d} = \{(\mathbf{s},\mathbf{t}) \in [0,1]^{2d} : \exists i \in \{1,\dots,d\} :\ s_i \geq t_i\}$$

und somit

$$\begin{aligned}&K_d \cap \overline{[0,1]^{2d} \setminus K_d}\\ &= \{(\mathbf{s},\mathbf{t}) \in [0,1]^{2d} : (\forall i \in \{1,\dots,d\}: \ s_i \le t_i) \ \wedge \ (\exists i \in \{1,\dots,d\}: \ s_i = t_i)\}.\end{aligned}$$

Für $(\mathbf{s},\mathbf{t}) \in K_d \cap \overline{[0,1]^{2d} \setminus K_d}$ gilt also:

- Für $d = 1$: Es ist $s = t$ und somit:

$$\sum_{\boldsymbol{\varepsilon} \in \{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} x(s + \boldsymbol{\varepsilon}(t-s)) = x(t) - x(s) = 0$$

- Für $d > 1$: O. E. sei $s_1 = t_1$. Es gilt:

$$\begin{aligned}&\sum_{\boldsymbol{\varepsilon} \in \{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} x(\mathbf{s} + \boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s}))\\ &= \sum_{\boldsymbol{\varepsilon} \in \{0,1\}^{d-1}} (-1)^{d-\sum_{i=2}^d \varepsilon_i} \Big(x(s_1, s_2 + \varepsilon_2(t_2 - s_2), \dots, s_2 + \varepsilon_d(t_d - s_d))\\ &\qquad - x(t_1, s_2 + \varepsilon_2(t_2 - s_2), \dots, s_d + \varepsilon_d(t_d - s_d)) \Big)\\ &\overset{s_1 = t_1}{=} 0\end{aligned}$$

Folglich ist g_x auf den abgeschlossenen Mengen K_d und $\overline{[0,1]^{2d} \setminus K_d}$ stetig und im Schnitt von beiden passend definiert, also nach Lemma 12 stetig. □

Wir werden Satz 31 anwenden, um aus der Konvergenz des Partialsummenprozesses auf die Konvergenz der Statistik zu schließen. Dazu müssen wir letztere als Funktional des Partialsummenprozesses darstellen. Dies machen wir schrittweise, indem wir die Teststatistik durch eine Verknüpfung von Funktionen des folgenden Typs approximieren:

Lemma 43. *Es seien $h : [0,1]^{2d} \to \mathbb{R}$ eine stetige Funktion mit $h \ge \gamma$ für ein $\gamma > 0$, $A \subset [0,1]^{2d}$ und K_d wie in Lemma 42. Dann sind die Funktionen*

$$f_1 : D[0,1]^d \to D[0,1]^{2d}, \ x \mapsto f_1(x)$$

mit

$$f_1(x)(\mathbf{s},\mathbf{t}) := \left(\sum_{\boldsymbol{\varepsilon} \in \{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} x(\mathbf{s} + \boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s})) - [\mathbf{t}-\mathbf{s}] x(\underline{\mathbf{1}}) \right) I_{K_d}(\mathbf{s},\mathbf{t})$$

$$f_2 : D[0,1]^{2d} \to D[0,1]^{2d}, \ x \mapsto f_2(x) \ \textit{mit} \ f_2(x)(\mathbf{s},\mathbf{t}) := \frac{x(\mathbf{s},\mathbf{t})}{\sqrt{h(\mathbf{s},\mathbf{t})}}$$

und

$$f_{3,A} : D[0,1]^{2d} \to \mathbb{R},\ x \mapsto f_{3,A}(x) \textit{ mit } f_{3,A}(x) := \sup_{(\mathbf{s},\mathbf{t})\in A} |x(\mathbf{s},\mathbf{t})|$$

wohldefiniert und stetig bezüglich der Supremumsnorm weiter gilt $f_1(C[0,1]^d) \subset C[0,1]^{2d}$ *und* $f_2(C[0,1]^{2d}) \subset C[0,1]^{2d}$.

Beweis. Zunächst zeigen wir die Wohldefiniertheit von f_1 und f_2. Seien $x \in D[0,1]^d$, $(\mathbf{s},\mathbf{t}) \in [0,1]^{2d}$ und $((\mathbf{s}_n,\mathbf{t}_n))_{n\in\mathbb{N}} \subset Q$ eine Folge mit $(\mathbf{s}_n,\mathbf{t}_n) \to (\mathbf{s},\mathbf{t})$, die in einem Quadranten Q (siehe Abschnitt 2.1) enthalten ist. Wir zeigen, dass der Grenzwert $\lim\limits_{(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\to(\mathbf{s},\mathbf{t}),(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in Q} f_1(x)(\tilde{\mathbf{s}},\tilde{\mathbf{t}})$ existiert und für $Q = Q_{\geq,\dots,\geq}$ gleich $f_1(x)(\mathbf{s},\mathbf{t})$ ist.

Falls $\mathbf{s} < \mathbf{t}$ ist, gibt es ein $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$ gilt $\mathbf{s}_n < \mathbf{t}_n$, und somit:

$$\begin{aligned} f_1(x)(\mathbf{s}_n,\mathbf{t}_n) &= \sum_{\boldsymbol{\varepsilon}\in\{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} \underbrace{x(\mathbf{s}_n + \boldsymbol{\varepsilon}(\mathbf{t}_n - \mathbf{s}_n))}_{\to x_Q(\mathbf{s}+\boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s}))} - \underbrace{[\mathbf{t}_n - \mathbf{s}_n]}_{\to[\mathbf{t}-\mathbf{s}]} x(\underline{\mathbf{1}}) \\ &\overset{n\to\infty}{\longrightarrow} \sum_{\boldsymbol{\varepsilon}\in\{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} x_Q(\mathbf{s} + \boldsymbol{\varepsilon}(\mathbf{t} - \mathbf{s})) - [\mathbf{t} - \mathbf{s}] x(\underline{\mathbf{1}}) \end{aligned}$$

Da $x \in D[0,1]^d$ ist, existieren obige Grenzwerte und für $Q = Q_{\geq,\dots,\geq}$ gilt

$$\lim_{n\to\infty} f_1(x)(\mathbf{s}_n,\mathbf{t}_n) = f_1(x)(\mathbf{s},\mathbf{t}).$$

Falls es ein $i \in \{1,\dots,d\}$ gibt mit $s_i > t_i$, so gilt für n groß genug auch $s_{i,n} > t_{i,n}$ und somit ist $\lim\limits_{n\to\infty} f_1(x)(\mathbf{s}_n,\mathbf{t}_n) = 0 = f_1(x)(\mathbf{s},\mathbf{t})$. Schließlich betrachten wir noch den Fall, dass es ein $i \in \{1,\dots,d\}$ gibt, so dass $s_i = t_i$. Wir stellen fest (vgl. Beweis von Lemma 42), dass die Funktionen

$$y \mapsto \sum_{\boldsymbol{\varepsilon}\in\{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} y(\mathbf{s} + \boldsymbol{\varepsilon}(\mathbf{t} - \mathbf{s}))$$

und

$$y \mapsto [\mathbf{t} - \mathbf{s}] y(\underline{\mathbf{1}})$$

in diesem Fall konstant 0 sind. Folglich ist $f_1(x)(\mathbf{s},\mathbf{t}) = 0$. Es lässt sich zeigen, dass auch $\lim\limits_{n\to\infty} f_1(x)(\mathbf{s}_n,\mathbf{t}_n) = 0$ ist: Betrachte die Indexmenge

$$N := \{n_k \in \mathbb{N} :\ \forall j \in \{1,\dots,d\} : \mathbf{s}_{n_k} \leq \mathbf{t}_{n_k}\}.$$

Für alle $n \notin N$ ist $f_1(x)(\mathbf{s}_n,\mathbf{t}_n) = 0$, d.h. wenn N endlich ist, ist nichts mehr zu zeigen. O. E. sei N nicht endlich. Für eine Folge $(n_k)_{k\in\mathbb{N}} \subset N$ gilt $(\mathbf{s}_{n_k},\mathbf{t}_{n_k}) \to (\mathbf{s},\mathbf{t})$ und somit folgt aus obigen Überlegungen:

$$f_1(x)(\mathbf{s}_{n_k},\mathbf{t}_{n_k}) \overset{k\to\infty}{\longrightarrow} \sum_{\boldsymbol{\varepsilon}\in\{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} x_Q(\mathbf{s} + \boldsymbol{\varepsilon}(\mathbf{t} - \mathbf{s})) - [\mathbf{t} - \mathbf{s}] x(\underline{\mathbf{1}}) = 0$$

Nach Lemma 16 folgt $f_1(x) \in D[0,1]^{2d}$. h ist stetig mit $h \geq \gamma$ und somit ist auch $(\mathbf{s},\mathbf{t}) \mapsto \frac{1}{\sqrt{h(\mathbf{s},\mathbf{t})}}$ stetig. Also liegt $f_2(x)$ nach Bemerkung 17 für $x \in D[0,1]^{2d}$

in $D[0,1]^{2d}$. Da Funktionen in $D[0,1]^{2d}$ nach Korollar 1.12 von Neuhaus (1969) beschränkt sind, ist auch $f_{3,A}$ wohldefiniert.
Zur Stetigkeit: Für $x, x_1, x_2, \dots \in D[0,1]^d$ mit $\|x_n - x\|_\infty \overset{n\to\infty}{\longrightarrow} 0$ gilt:

$$\begin{aligned}
&\|f_1(x_n) - f_1(x)\|_\infty \\
=& \sup_{(\mathbf{s},\mathbf{t})\in K_d} \Bigg| \sum_{\boldsymbol{\varepsilon}\in\{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} (x_n(\mathbf{s}+\boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s})) - x(\mathbf{s}+\boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s}))) \\
&\qquad - [\mathbf{t}-\mathbf{s}](x_n(\underline{\mathbf{1}}) - x(\underline{\mathbf{1}}))\Bigg| \\
\leq& \sum_{\boldsymbol{\varepsilon}\in\{0,1\}^d} \underbrace{\sup_{(\mathbf{s},\mathbf{t})\in K_d} |x_n(\mathbf{s}+\boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s})) - x(\mathbf{s}+\boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s}))|}_{\leq \|x_n - x\|_\infty} \\
&+ \underbrace{\sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} |[\mathbf{t}-\mathbf{s}]|}_{=1} \underbrace{|x_n(\underline{\mathbf{1}}) - x(\underline{\mathbf{1}})|}_{\leq\|x_n - x\|_\infty} \\
\leq& (2^d+1)\|x_n - x\|_\infty \overset{n\to\infty}{\longrightarrow} 0
\end{aligned}$$

Nun seien $x, x_1, x_2, \dots \in D[0,1]^{2d}$ mit $\|x_n - x\|_\infty \overset{n\to\infty}{\longrightarrow} 0$. Es gilt:

$$\begin{aligned}
\|f_2(x_n) - f_2(x)\|_\infty =& \sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \left| \frac{1}{\sqrt{h(\mathbf{s},\mathbf{t})}} (x_n(\mathbf{s},\mathbf{t}) - x(\mathbf{s},\mathbf{t})) \right| \\
\leq& \frac{1}{\sqrt{\gamma}} \|x_n - x\|_\infty \overset{n\to\infty}{\longrightarrow} 0
\end{aligned}$$

und

$$\begin{aligned}
|f_{3,A}(x_n) - f_{3,A}(x)| =& \left| \sup_{(\mathbf{s},\mathbf{t})\in A} |x_n(\mathbf{s},\mathbf{t})| - \sup_{(\mathbf{s},\mathbf{t})\in A} |x(\mathbf{s},\mathbf{t})| \right| \\
\leq& \sup_{(\mathbf{s},\mathbf{t})\in A} |x_n(\mathbf{s},\mathbf{t}) - x(\mathbf{s},\mathbf{t})| \\
\leq& \|x_n - x\|_\infty \overset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Zu $f_1(C[0,1]^d) \subset C[0,1]^{2d}$: Für $x \in C[0,1]^d$ ist $f_1(x)$ gemäß Lemma 42 stetig. Zu $f_2(C[0,1]^{2d}) \subset C[0,1]^{2d}$: Sei $x \in C[0,1]^{2d}$. Da wir h streng positiv vorgegeben haben, weist $f_2(x)$ keine Pole auf und ist als Quotient zweier stetiger Funktionen selbst stetig. □

Wir definieren für $(\mathbf{s},\mathbf{t}) \in [0,1]^{2d}$ und $n \in \mathbb{N}$:

$$\tilde{Z}_n(\mathbf{s},\mathbf{t}) := \begin{cases} \sum_{\boldsymbol{\varepsilon}\in\{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} S_n(\mathbf{s}+\boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s})) - \frac{[\lfloor n\mathbf{t}\rfloor - \lfloor n\mathbf{s}\rfloor]}{n^d} S_n(\underline{\mathbf{1}}), & \mathbf{s}\leq\mathbf{t} \\ 0, & \text{sonst} \end{cases}$$

und

$$Z_n(\mathbf{s},\mathbf{t}) := \begin{cases} \sum_{\boldsymbol{\varepsilon}\in\{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} S_n(\mathbf{s}+\boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s})) - [\mathbf{t}-\mathbf{s}] S_n(\underline{\mathbf{1}}), & \mathbf{s}\leq\mathbf{t} \\ 0, & \text{sonst} \end{cases}$$

Seien $\tilde{\alpha}, \tilde{\beta} \in (0,1)$, $0 < \tilde{\alpha} < 1 - \tilde{\beta} < 1$. Für

$$h_n(\mathbf{s},\mathbf{t}) := \begin{cases} \tilde{\alpha}(1-\tilde{\alpha}), & 0 \leq |[\lfloor n\mathbf{t}\rfloor - \lfloor n\mathbf{s}\rfloor]| < n^d\tilde{\alpha} \\ \frac{|[\lfloor n\mathbf{t}\rfloor - \lfloor n\mathbf{s}\rfloor]|}{n^d}\left(1 - \frac{|[\lfloor n\mathbf{t}\rfloor - \lfloor n\mathbf{s}\rfloor]|}{n^d}\right), & n^d\tilde{\alpha} \leq |[\lfloor n\mathbf{t}\rfloor - \lfloor n\mathbf{s}\rfloor]| \leq n^d(1-\tilde{\beta}) \\ \tilde{\beta}(1-\tilde{\beta}), & n^d(1-\tilde{\beta}) < |[\lfloor n\mathbf{t}\rfloor - \lfloor n\mathbf{s}\rfloor]| \leq n^d \end{cases}$$

und

$$h(\mathbf{s},\mathbf{t}) := \begin{cases} \tilde{\alpha}(1-\tilde{\alpha}), & 0 \leq |[\mathbf{t}-\mathbf{s}]| < \tilde{\alpha} \\ |[\mathbf{t}-\mathbf{s}]|(1-|[\mathbf{t}-\mathbf{s}]|), & \tilde{\alpha} \leq |[\mathbf{t}-\mathbf{s}]| \leq 1-\tilde{\beta} \\ \tilde{\beta}(1-\tilde{\beta}), & 1-\tilde{\beta} < |[\mathbf{t}-\mathbf{s}]| \leq 1 \end{cases}$$

seien die Prozesse

$$\tilde{Y}_n(\mathbf{s},\mathbf{t}) := \frac{\tilde{Z}_n(\mathbf{s},\mathbf{t})}{\sqrt{h_n(\mathbf{s},\mathbf{t})}} \text{ und } Y_n(\mathbf{s},\mathbf{t}) := \frac{Z_n(\mathbf{s},\mathbf{t})}{\sqrt{h(\mathbf{s},\mathbf{t})}}$$

definiert.

Bemerkung 44. *(a) Es gilt für $\underline{\mathbf{0}} \leq \mathbf{s} \leq \mathbf{t} \leq \underline{\mathbf{1}}$:*

$$\begin{aligned} & n^{-d/2}\hat{\sigma}_n^{-1} \sum_{\lfloor n\mathbf{s}\rfloor < \mathbf{j} \leq \lfloor n\mathbf{t}\rfloor} (X_{\mathbf{j}} - \bar{X}) \\ &= \sum_{\boldsymbol{\varepsilon}\in\{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} S_n(\mathbf{s}+\boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s})) - \frac{[\lfloor n\mathbf{t}\rfloor - \lfloor n\mathbf{s}\rfloor]}{n^d} S_n(\underline{\mathbf{1}}) \end{aligned}$$

(b) Es ist $h_n(\mathbf{s},\mathbf{t}) = h(\frac{\lfloor n\mathbf{s}\rfloor}{n}, \frac{\lfloor n\mathbf{t}\rfloor}{n})$ und es gibt ein $\gamma > 0$ mit $h(\mathbf{s},\mathbf{t}) \geq \gamma$ und somit auch $h_n(\mathbf{s},\mathbf{t}) \geq \gamma$. Weiter ist h auf $[0,1]^{2d}$ gleichmäßig stetig.

Beweis. siehe Anhang □

Schließlich definieren wir noch die $D[0,1]^d$-wertigen Prozesse

$$Z(\mathbf{s},\mathbf{t}) := \begin{cases} \sum_{\boldsymbol{\varepsilon}\in\{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} W(\mathbf{s}+\boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s})) - [\mathbf{t}-\mathbf{s}] W(\underline{\mathbf{1}}), & \mathbf{s}\leq\mathbf{t} \\ 0, & \text{sonst} \end{cases} \tag{3.2}$$

und

$$Y(\mathbf{s},\mathbf{t}) := \frac{Z(\mathbf{s},\mathbf{t})}{\sqrt{h(\mathbf{s},\mathbf{t})}}. \tag{3.3}$$

Bemerkung 45. *Für feste $(\mathbf{s},\mathbf{t}) \in [0,1]^{2d}$ sind $Z_n(\mathbf{s},\mathbf{t})$, $\tilde{Z}_n(\mathbf{s},\mathbf{t})$, $Z(\mathbf{s},\mathbf{t})$, $Y_n(\mathbf{s},\mathbf{t})$, $\tilde{Y}_n(\mathbf{s},\mathbf{t})$ und $Y(\mathbf{s},\mathbf{t})$ offensichtlich reelle Zufallsvariablen. Analog zur Vorgehensweise im Beweis von Lemma 43 lässt sich zeigen, dass $\{Z_n(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}}$, $\{\tilde{Z}_n(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}}$, $\{Z(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}}$, $\{Y_n(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}}$, $\{\tilde{Y}_n(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}}$ und $\{Y(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}}$ $D[0,1]^{2d}$-wertige Prozesse sind, wobei wir o. E. voraussetzen, dass die Pfade von $\{W(\mathbf{t})\}_{\mathbf{t}\in[0,1]^d}$ in $D[0,1]^d$ liegen.*

Bemerkung 46. *Für die Einschränkung* $Y = \{Y(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in\tilde{A}}$ *auf*

$$\tilde{A} := \{(\mathbf{s},\mathbf{t}) \in [0,1]^{2d} : \ \underline{\mathbf{0}} \le \mathbf{s} \le \mathbf{t} \le \underline{\mathbf{1}}, \tilde{\alpha} \le [\mathbf{t}-\mathbf{s}] \le 1-\tilde{\beta}\}$$

und $0 < \tilde{\alpha} < 1-\tilde{\beta} < 1$ *gilt:*

a) Y *ist ein zentrierter Gauß'scher Prozess mit konstanter Varianz* 1.

b) Die Pfade von $\{Y(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^d}$ *sind fast sicher stetig.*

Beweis. siehe Anhang □

Nun untersuchen wir die Verteilungskonvergenz der eingeführten Prozesse. Zunächst stellen wir fest:

Satz 47. *Annahme 1 sei erfüllt. Dann gilt für* Z_n*,* Z*,* Y_n *und* Y *wie oben und* $A \subset [0,1]^{2d}$*:*

(a) $\{Z_n(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \stackrel{D[0,1]^{2d}}{\longrightarrow} \{Z(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}}$

(b) $\{Y_n(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \stackrel{D[0,1]^{2d}}{\longrightarrow} \{Y(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}}$

(c) $\sup\limits_{(\mathbf{s},\mathbf{t})\in A} |Y_n(\mathbf{s},\mathbf{t})| \stackrel{\mathcal{D}}{\to} \sup\limits_{(\mathbf{s},\mathbf{t})\in A} |Y(\mathbf{s},\mathbf{t})|$

(d) $Y \in C[0,1]^{2d}$ *P-fast sicher*

Beweis. $f_1, f_2, f_{3,A}$ seien die Funktionen aus Lemma 43. Offensichtlich sind $Z_n = f_1(S_n(\cdot))$, $Y_n = f_2(f_1(S_n(\cdot)))$, $Z = f_1(W(\cdot))$, $Y = f_2(f_1(W(\cdot)))$,

$$\sup_{(\mathbf{s},\mathbf{t})\in A} |Y_n(\mathbf{s},\mathbf{t})| = f_{3,A}(f_2(f_1(S_n(\cdot)))) \text{ und } \sup_{(\mathbf{s},\mathbf{t})\in A} |Y(\mathbf{s},\mathbf{t})| = f_{3,A}(f_2(f_1(W(\cdot)))).$$

Dabei sind $\sup\limits_{(\mathbf{s},\mathbf{t})\in A} |Y_n(\mathbf{s},\mathbf{t})|$ und $\sup\limits_{(\mathbf{s},\mathbf{t})\in A} |Y(\mathbf{s},\mathbf{t})|$ für alle $n \in \mathbb{N}$ reelle Zufallsvariablen, da nach Bemerkung 45 $f_2(f_1(S_n(\cdot)))$ und $f_2(f_1(W(\cdot)))$ $D[0,1]^{2d}$-wertige stochastische Prozesse sind und $f_{3,A}$ nach Lemma 2.22 von Döring (2007) $\mathcal{D}_q$–$\mathcal{B}(\mathbb{R})$-messbar ist. Da wir Annahme 1 voraussetzen und $P(W \in C[0,1]^d) = 1$ ist, folgt die Behauptung nun aus Lemma 43 und Satz 31:
Nach Lemma 43 gilt

$$P(f_2(f_1(W(\cdot))) \in C[0,1]^{2d}) \ge P(f_1(W(\cdot)) \in C[0,1]^{2d}) \ge P(W \in C[0,1]^d) = 1,$$

und f_1, f_2 und $f_{3,A}$ sind stetig bezüglich der Supremumsnorm. Satz 31 liefert somit die Behauptung. Da $Y = f_2(f_1(W(\cdot)))$ ist, liefert obige Ungleichungskette auch Behauptung (d). □

Lemma 48. *Für* $n \to \infty$ *gilt:*

$$\|\tilde{Y}_n - Y_n\|_\infty = o_P(1)$$

Beweis. Zunächst zeigen wir

$$\sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \left|\tilde{Z}_n(\mathbf{s},\mathbf{t}) - Z_n(\mathbf{s},\mathbf{t})\right| \stackrel{!}{=} o_P(1).$$

Betrachte dafür:

$$\begin{aligned}
&\sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \left|\tilde{Z}_n(\mathbf{s},\mathbf{t}) - Z_n(\mathbf{s},\mathbf{t})\right| \\
\leq &\sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \left|[\mathbf{t}-\mathbf{s}]S_n(\underline{\mathbf{1}}) - \frac{[\lfloor n\mathbf{t}\rfloor - \lfloor n\mathbf{s}\rfloor]}{n^d} S_n(\underline{\mathbf{1}})\right| \\
= &\underbrace{\sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \left|[\mathbf{t}-\mathbf{s}] - \frac{[\lfloor n\mathbf{t}\rfloor - \lfloor n\mathbf{s}\rfloor]}{n^d}\right|}_{=o(1),\text{ wg. Lemma 41}} \underbrace{|S_n(\underline{\mathbf{1}})|}_{=\mathcal{O}_P(1),\text{ wg. Annahme 1}} \\
= &o_P(1)
\end{aligned}$$

Nutzen wir dies und die Tatsache, dass nach Bemerkung 44 ein $\gamma > 0$ existiert, so dass $h_n(\mathbf{s},\mathbf{t}) \geq \gamma$ und $h(\mathbf{s},\mathbf{t}) \geq \gamma$ für alle $n \in \mathbb{N}$ und $(\mathbf{s},\mathbf{t}) \in [0,1]^{2d}$, so ergibt sich:

$$\begin{aligned}
&\|\tilde{Y}_n - Y_n\|_\infty \\
= &\sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} |\tilde{Y}_n(\mathbf{s},\mathbf{t}) - Y_n(\mathbf{s},\mathbf{t})| \\
\leq &\sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \left|\frac{\tilde{Z}_n(\mathbf{s},\mathbf{t}) - Z_n(\mathbf{s},\mathbf{t})}{\sqrt{h_n(\mathbf{s},\mathbf{t})}}\right| \\
&+ \sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \left|\left(\frac{1}{\sqrt{h_n(\mathbf{s},\mathbf{t})}} - \frac{1}{\sqrt{h(\mathbf{s},\mathbf{t})}}\right) Z_n(\mathbf{s},\mathbf{t})\right| \\
\leq &\frac{1}{\sqrt{\gamma}} \sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \left|\tilde{Z}_n(\mathbf{s},\mathbf{t}) - Z_n(\mathbf{s},\mathbf{t})\right| \\
&+ \frac{1}{\gamma} \sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \left|\left(\sqrt{h(\mathbf{s},\mathbf{t})} - \sqrt{h_n(\mathbf{s},\mathbf{t})}\right) Z_n(\mathbf{s},\mathbf{t})\right| \\
\leq &o_P(1) + \frac{1}{\gamma} \sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \left|\sqrt{h(\mathbf{s},\mathbf{t})} - \sqrt{h_n(\mathbf{s},\mathbf{t})}\right| \sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} |Z_n(\mathbf{s},\mathbf{t})|
\end{aligned}$$

Mit Lemma 43, Satz 47(a) und Satz 31 folgt

$$\sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} |Z_n(\mathbf{s},\mathbf{t})| \overset{\mathcal{D}}{\to} \sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} |Z(\mathbf{s},\mathbf{t})|,$$

also

$$\sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} |Z_n(\mathbf{s},\mathbf{t})| = \mathcal{O}_P(1).$$

Es bleibt noch zu zeigen:

$$\sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} |\sqrt{h(\mathbf{s},\mathbf{t})} - \sqrt{h_n(\mathbf{s},\mathbf{t})}| \overset{!}{=} o(1)$$

Sei $\varepsilon > 0$ beliebig. Aus Bemerkung 44 wissen wir, dass h und somit auch $\sqrt{h}$ gleichmäßig stetig sind. Folglich gibt es ein $\delta > 0$, so dass

$$\sup_{(\mathbf{s},\mathbf{t}),(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in[0,1]^{2d},\ \|(\mathbf{s},\mathbf{t})-(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\|<\delta} \left|\sqrt{h(\mathbf{s},\mathbf{t})} - \sqrt{h(\tilde{\mathbf{s}},\tilde{\mathbf{t}})}\right| < \varepsilon$$

Aufgrund der gleichmäßigen Konvergenz von $\lfloor nt \rfloor / n \to t$ auf $[0,1]$ gibt es ein $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$ gilt:

$$\sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \left\| (\mathbf{s},\mathbf{t}) - \left(\frac{\lfloor n\mathbf{s} \rfloor}{n}, \frac{\lfloor n\mathbf{t} \rfloor}{n} \right) \right\| \leq \sup_{t\in[0,1]} \left| t - \frac{\lfloor nt \rfloor}{n} \right| < \delta$$

Somit gilt für alle $n \geq n_0$:

$$\begin{aligned} &\sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \left| \sqrt{h(\mathbf{s},\mathbf{t})} - \sqrt{h_n(\mathbf{s},\mathbf{t})} \right| \\ = &\sup_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}} \left| \sqrt{h(\mathbf{s},\mathbf{t})} - \sqrt{h\left(\frac{\lfloor n\mathbf{s} \rfloor}{n}, \frac{\lfloor n\mathbf{t} \rfloor}{n} \right)} \right| \\ \leq &\sup_{(\mathbf{s},\mathbf{t}),(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in[0,1]^{2d},\ \|(\mathbf{s},\mathbf{t})-(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\|<\delta} \left| \sqrt{h(\mathbf{s},\mathbf{t})} - \sqrt{h(\tilde{\mathbf{s}},\tilde{\mathbf{t}})} \right| < \varepsilon \end{aligned}$$

□

Wir wollen aus den Konvergenz- und Approximationsaussagen in Satz 47 und Lemma 48 auf die Konvergenz der Statistik schließen. Für die Teststatistik wird das Supremum über den Bereich $\lfloor \alpha n^d \rfloor \leq [\lfloor n\mathbf{t} \rfloor - \lfloor n\mathbf{s} \rfloor] \leq \lfloor (1-\beta)n^d \rfloor$ und nicht über $\alpha \leq [\mathbf{t}-\mathbf{s}] \leq 1-\beta$ genommen. Folgendes Lemma zeigt, dass dies für die Verteilungskonvergenz unerheblich ist. Für den Beweis von Lemma 49 werden wir zwei Hilfslemmata benutzen, die wir im Anschluss an das Lemma angeben.

Lemma 49. *Seien $\{Y_n(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}}$ und $\{Y(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}}$ stochastische Prozesse mit Pfaden in $D[0,1]^{2d}$ für alle $n \in \mathbb{N}$ und $Y \in C[0,1]^{2d}$ fast sicher. Weiter seien für $\varepsilon > 0$ mit*

$$0 < \alpha - \varepsilon < \alpha + \varepsilon < 1 - \beta - \varepsilon < 1 - \beta + \varepsilon < 1 \tag{3.4}$$

die folgenden Teilmengen von $[0,1]^{2d}$ definiert:

$$A := \{(\mathbf{s},\mathbf{t}) \in [0,1]^{2d} :\ \underline{\mathbf{0}} \leq \mathbf{s} < \mathbf{t} \leq \underline{\mathbf{1}},\ \alpha \leq [\mathbf{t}-\mathbf{s}] \leq 1-\beta\},$$

$$A_n := \{(\mathbf{s},\mathbf{t}) \in [0,1]^{2d} :\ \underline{\mathbf{0}} \leq \mathbf{s} < \mathbf{t} \leq \underline{\mathbf{1}},\ \lfloor \alpha n^d \rfloor \leq [\lfloor \mathbf{t}n \rfloor - \lfloor \mathbf{s}n \rfloor] \leq \lfloor (1-\beta)n^d \rfloor\}$$

und

$$A_{\pm\varepsilon} := \{(\mathbf{s},\mathbf{t}) \in [0,1]^{2d} :\ \underline{\mathbf{0}} \leq \mathbf{s} < \mathbf{t} \leq \underline{\mathbf{1}},\ \alpha \pm \varepsilon \leq [\mathbf{t}-\mathbf{s}] \leq 1-\beta \mp \varepsilon\}$$

Auf einer kompakten Menge $K \supset A \cup A_{\pm\varepsilon}$ sei $\{Y(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in K}$ ein zentrierter Gauß'scher Prozess mit konstanter Varianz $\sigma^2 := E[Y(\mathbf{s},\mathbf{t})^2] > 0$. Für alle $\varepsilon > 0$, die (3.4) erfüllen, gelte:

$$T_{\pm\varepsilon,n} := \sup_{(\mathbf{s},\mathbf{t})\in A_{\pm\varepsilon}} |Y_n(\mathbf{s},\mathbf{t})| \xrightarrow{\mathcal{D}} T_{\pm\varepsilon} := \sup_{(\mathbf{s},\mathbf{t})\in A_{\pm\varepsilon}} |Y(\mathbf{s},\mathbf{t})| \quad (n \to \infty) \tag{3.5}$$

Dann gilt auch:

$$T_n := \sup_{(\mathbf{s},\mathbf{t})\in A_n} |Y_n(\mathbf{s},\mathbf{t})| \xrightarrow{\mathcal{D}} T := \sup_{(\mathbf{s},\mathbf{t})\in A} |Y(\mathbf{s},\mathbf{t})| \quad (n \to \infty)$$

Beweis. 1.) Wir zeigen zunächst:

$$\begin{aligned}
&\forall \varepsilon > 0 \text{ mit (3.4) } \exists n_0 \in \mathbb{N}\ \forall n \geq n_0\ \forall \underline{\mathbf{0}} \leq \mathbf{s} < \mathbf{t} \leq \underline{\mathbf{1}}: \\
&\alpha + \varepsilon \leq [\mathbf{t} - \mathbf{s}] \leq 1 - \beta - \varepsilon \\
\Rightarrow &\lfloor \alpha n^d \rfloor \leq [\lfloor \mathbf{t} n \rfloor - \lfloor \mathbf{s} n \rfloor] \leq \lfloor (1-\beta) n^d \rfloor \\
\Rightarrow &\alpha - \varepsilon \leq [\mathbf{t} - \mathbf{s}] \leq 1 - \beta + \varepsilon
\end{aligned} \tag{3.6}$$

Sei ein $\varepsilon > 0$, das (3.4) erfüllt, gegeben. Als erstes stellen wir fest, dass es wegen Lemma 41 ein $n_0 \in \mathbb{N}$ gibt, so dass für alle $n \geq n_0$ gilt:

$$\sup_{(\mathbf{s},\mathbf{t}) \in [0,1]^{2d}} \left| [\mathbf{t} - \mathbf{s}] - \left[\frac{\lfloor n\mathbf{t} \rfloor - \lfloor n\mathbf{s} \rfloor}{n} \right] \right| \leq \frac{\varepsilon}{2} \tag{3.7}$$

und

$$\lfloor \alpha n^d \rfloor \geq \alpha n^d - \frac{\varepsilon}{2} n^d \tag{3.8}$$

Aus (3.7) folgt für $\underline{\mathbf{0}} \leq \mathbf{s} < \mathbf{t} \leq \underline{\mathbf{1}}$:

$$[\mathbf{t} - \mathbf{s}] n^d - \frac{\varepsilon}{2} n^d \leq [\lfloor n\mathbf{t} \rfloor - \lfloor n\mathbf{s} \rfloor] \leq [\mathbf{t} - \mathbf{s}] n^d + \frac{\varepsilon}{2} n^d$$

Somit folgt aus $\alpha + \varepsilon \leq [\mathbf{t} - \mathbf{s}] \leq 1 - \beta - \varepsilon$:

$$\begin{aligned}
\lfloor \alpha n^d \rfloor &\leq \alpha n^d \\
&= (\alpha + \varepsilon) n^d - \varepsilon n^d \\
&\leq [\mathbf{t} - \mathbf{s}] n^d - \varepsilon n^d \\
&\leq [\mathbf{t} - \mathbf{s}] n^d - \frac{\varepsilon}{2} n^d \\
&\leq [\lfloor n\mathbf{t} \rfloor - \lfloor n\mathbf{s} \rfloor] \\
&\leq [\mathbf{t} - \mathbf{s}] n^d + \frac{\varepsilon}{2} n^d \\
&\leq (1-\beta) n^d - \frac{\varepsilon}{2} n^d \\
&\leq (1-\beta) n^d
\end{aligned}$$

Da $[\lfloor n\mathbf{t} \rfloor - \lfloor n\mathbf{s} \rfloor] \in \mathbb{N}$, muss somit auch $[\lfloor n\mathbf{t} \rfloor - \lfloor n\mathbf{s} \rfloor] \leq \lfloor (1-\beta) n^d \rfloor$ gelten.

Für $\underline{\mathbf{0}} \leq \mathbf{s} < \mathbf{t} \leq \underline{\mathbf{1}}$ mit $\lfloor \alpha n^d \rfloor \leq [\lfloor \mathbf{t} n \rfloor - \lfloor \mathbf{s} n \rfloor] \leq \lfloor (1-\beta) n^d \rfloor$ gilt:

$$\begin{aligned}
\alpha - \varepsilon &\overset{(3.8)}{\leq} \frac{\lfloor \alpha n^d \rfloor}{n^d} - \frac{\varepsilon}{2} \\
&\leq \frac{[\lfloor n\mathbf{t} \rfloor - \lfloor n\mathbf{s} \rfloor]}{n^d} - \frac{\varepsilon}{2} \\
&\leq [\mathbf{t} - \mathbf{s}] \\
&\leq \frac{[\lfloor n\mathbf{t} \rfloor - \lfloor n\mathbf{s} \rfloor]}{n^d} + \frac{\varepsilon}{2} \\
&\leq \frac{\lfloor (1-\beta) n^d \rfloor}{n^d} + \frac{\varepsilon}{2} \\
&\leq 1 - \beta + \varepsilon
\end{aligned}$$

Folglich gilt (3.6).

2.) Aus 1. folgt:

$$\forall \varepsilon > 0 \text{ mit } (3.4)\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0 :\ T_{+\varepsilon,n} \leq T_n \leq T_{-\varepsilon,n} \tag{3.9}$$

Wir gehen im Folgenden o. E. davon aus, dass $\varepsilon > 0$ so klein ist, dass (3.4) erfüllt ist. Dann sind $T_{\pm\varepsilon}$ nach Lemma 52 stetig verteilt. Mit (3.5) folgt nun für $x \in \mathbb{R}$:

$$\begin{aligned}
P(T_{-\varepsilon} \leq x) &\overset{(3.5)}{=} \lim_{n\to\infty} P(T_{-\varepsilon,n} \leq x) \\
&\overset{(3.9)}{\leq} \liminf_{n\to\infty} P(T_n \leq x) \\
&\leq \limsup_{n\to\infty} P(T_n \leq x) \\
&\overset{(3.9)}{\leq} \lim_{n\to\infty} P(T_{+\varepsilon,n} \leq x) \\
&\overset{(3.5)}{=} P(T_{+\varepsilon} \leq x)
\end{aligned}$$

3.) Nach Lemma 51 gilt $T_{\pm\varepsilon} \xrightarrow{\varepsilon\to 0} T$ fast sicher, also auch in Verteilung. Da T nach Lemma 52 stetig verteilt ist, gilt somit:

$$P(T_{\pm\varepsilon} \leq x) \xrightarrow{\varepsilon\to 0} P(T \leq x)$$

Damit folgt aus 2.:

$$\begin{aligned}
P(T \leq x) &= \lim_{\varepsilon\to 0} P(T_{-\varepsilon} \leq x) \\
&\leq \liminf_{n\to\infty} P(T_n \leq x) \\
&\leq \limsup_{n\to\infty} P(T_n \leq x) \\
&\leq \lim_{\varepsilon\to 0} P(T_{+\varepsilon} \leq x) \\
&= P(T \leq x)
\end{aligned}$$

Und schließlich ergibt sich:

$$\lim_{n\to\infty} P(T_n \leq x) = P(T \leq x)$$

□

Bemerkung 50. *Nach Bemerkung 46 erfüllt der Prozess* (3.3) *die Voraussetzungen von Lemma 49.*

Lemma 51. *Seien* $f : [0,1]^{2d} \to \mathbb{R}$ *eine stetige Funktion,* $0 < a < b < 1$ *und* $(a_n)_{n\in\mathbb{N}}$ *und* $(b_n)_{n\in\mathbb{N}}$ *reelle Zahlenfolgen mit*

(a) $a_n \nearrow a$ *und* $b_n \searrow b$

oder

(b) $a_n \searrow a$ *und* $b_n \nearrow b$

Dann gilt:

$$\sup_{(\mathbf{s},\mathbf{t})\in A_n} |f(\mathbf{s},\mathbf{t})| - \sup_{(\mathbf{s},\mathbf{t})\in A} |f(\mathbf{s},\mathbf{t})| \overset{n\to\infty}{\longrightarrow} 0,$$

wobei

$$A_n := \{(\mathbf{s},\mathbf{t}) \in [0,1]^{2d} : \ \underline{\mathbf{0}} \leq \mathbf{s} < \mathbf{t} \leq \underline{\mathbf{1}}, \ a_n \leq [\mathbf{t}-\mathbf{s}] \leq b_n\}$$

und

$$A := \{(\mathbf{s},\mathbf{t}) \in [0,1]^{2d} : \ \underline{\mathbf{0}} \leq \mathbf{s} < \mathbf{t} \leq \underline{\mathbf{1}}, \ a \leq [\mathbf{t}-\mathbf{s}] \leq b\}.$$

Beweis. siehe Anhang □

Folgendes Lemma ist eine direkte Folgerung aus Proposition 2 von Lifshits (1984) (bzw. Tsirel'son (1975), Theorem 3):

Lemma 52. *Sei* $\{Y(\mathbf{t})\}_{\mathbf{t}\in T}$ *ein zentrierter Gauß'scher Prozess mit kompakter Parametermenge* $T \subset \mathbb{R}^d$ *und fast sicher stetigen Pfaden. Weiter sei* $\sigma^2 := E[Y^2(\mathbf{t})] > 0$ *für alle* $\mathbf{t} \in T$*. Sei* $\tilde{T} \subseteq T$ *ebenfalls eine kompakte Menge. Dann ist* $\sup_{\mathbf{t}\in\tilde{T}} |Y(\mathbf{t})|$ *stetig verteilt.*

Beweis. Zunächst stellen wir fest, dass es reicht, die Stetigkeit der Verteilungen von $G := \sup_{\mathbf{t}\in\tilde{T}} Y(\mathbf{t})$ und $G^\star := \sup_{\mathbf{t}\in\tilde{T}}(-Y(\mathbf{t}))$ zu zeigen. Für $x \in \mathbb{R}$ gilt nämlich:

$$\begin{aligned} P(\sup_{\mathbf{t}\in\tilde{T}} |Y(\mathbf{t})| = x) &= P(\sup_{\mathbf{t}\in\tilde{T}} Y(\mathbf{t}) = x \ \vee \ \sup_{\mathbf{t}\in\tilde{T}}(-Y(\mathbf{t})) = x) \\ &\leq P(\sup_{\mathbf{t}\in\tilde{T}} Y(\mathbf{t}) = x) + P(\sup_{\mathbf{t}\in\tilde{T}}(-Y(\mathbf{t})) = x) \end{aligned}$$

Da die Pfade von Y fast sicher stetig sind, ist Y ein fast sicher separabler Prozess. Weiter nimmt Y als stetige Funktion auf kompakten Mengen fast sicher sein Supremum an, so dass G fast sicher endlich ist. Sei A die Einsmenge, auf der Y stetig (und damit separabel) ist. Dann gilt für eine beliebige Borelmenge $B \subset \mathbb{R}$, dass $P(G \in B, A) = P(G \in B)$, weshalb wir im Folgenden o. E. $A = \Omega$ annehmen werden. Somit sind alle Voraussetzungen von Proposition 2 aus Lifshits (1984) erfüllt, und es folgt die Stetigkeit der Verteilung von G. Wie Y ist auch $-Y$ ein zentrierter Gauß'scher Prozess mit fast sicher stetigen Pfaden und konstanter, positiver Varianz, so dass die Stetigkeit von $G^\star$ analog folgt. □

Nun setzen wir unsere Erkenntnisse aus diesem Abschnitt zusammen, um Satz 40 zu beweisen:

Beweis von Satz 40. Bemerkung 44 liefert für $0 < \tilde{\alpha} < \alpha$, $0 < \tilde{\beta} < \beta$ und n

groß genug:

$$
\begin{aligned}
T_n(\alpha,\beta) &= \hat{\sigma}_n^{-1} n^{-d/2} \max_{\underline{\mathbf{0}}\le\mathbf{k}<\mathbf{m}\le\underline{\mathbf{n}},\ \lfloor\alpha n^d\rfloor\le[\mathbf{m}-\mathbf{k}]\le\lfloor(1-\beta)n^d\rfloor} \frac{\left|\sum\limits_{\mathbf{k}<\mathbf{j}\le\mathbf{m}}\left(X_{\mathbf{j}}-\bar{X}\right)\right|}{\sqrt{\frac{[\mathbf{m}-\mathbf{k}]}{n^d}\left(1-\frac{[\mathbf{m}-\mathbf{k}]}{n^d}\right)}} \\
&= \sup_{\underline{\mathbf{0}}\le\mathbf{s}<\mathbf{t}\le\underline{\mathbf{1}},\ \lfloor\alpha n^d\rfloor\le[\lfloor\mathbf{t}n\rfloor-\lfloor\mathbf{s}n\rfloor]\le\lfloor(1-\beta)n^d\rfloor} \frac{\hat{\sigma}_n^{-1} n^{-d/2}\left|\sum\limits_{\lfloor n\mathbf{s}\rfloor<\mathbf{j}\le\lfloor n\mathbf{t}\rfloor}(X_{\mathbf{j}}-\bar{X})\right|}{\sqrt{\left[\frac{\lfloor n\mathbf{t}\rfloor}{n}-\frac{\lfloor n\mathbf{s}\rfloor}{n}\right]\left(1-\left[\frac{\lfloor n\mathbf{t}\rfloor}{n}-\frac{\lfloor n\mathbf{s}\rfloor}{n}\right]\right)}} \\
&= \sup_{\underline{\mathbf{0}}\le\mathbf{s}<\mathbf{t}\le\underline{\mathbf{1}},\ \lfloor\alpha n^d\rfloor\le[\lfloor\mathbf{t}n\rfloor-\lfloor\mathbf{s}n\rfloor]\le\lfloor(1-\beta)n^d\rfloor} |\tilde{Y}_n(\mathbf{s},\mathbf{t})|
\end{aligned}
$$

Sei A_n wie in Lemma 49. Aus Lemma 48 folgt:

$$
\begin{aligned}
\left|\sup_{(\mathbf{s},\mathbf{t})\in A_n}|\tilde{Y}_n(\mathbf{s},\mathbf{t})| - \sup_{(\mathbf{s},\mathbf{t})\in A_n}|Y_n(\mathbf{s},\mathbf{t})|\right| &\le \sup_{(\mathbf{s},\mathbf{t})\in A_n}|\tilde{Y}_n(\mathbf{s},\mathbf{t})-Y_n(\mathbf{s},\mathbf{t})| \\
&\le \|\tilde{Y}_n - Y_n\|_\infty = o_P(1)
\end{aligned}
$$

Somit ist

$$
T_n(\alpha,\beta) = \sup_{(\mathbf{s},\mathbf{t})\in A_n}|Y_n(\mathbf{s},\mathbf{t})| + o_P(1).
$$

Satz 47 (c) angewandt auf die Mengen $A_{\pm\varepsilon}$ aus Lemma 49 liefert nun, dass (3.5) für $\{Y_n(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in[0,1]^{2d}}$ erfüllt ist, und somit folgt aus Lemma 49:

$$
\sup_{(\mathbf{s},\mathbf{t})\in A_n}|Y_n(\mathbf{s},\mathbf{t})| \xrightarrow{\mathcal{D}} \sup_{(\mathbf{s},\mathbf{t})\in A}|Y(\mathbf{s},\mathbf{t})|
$$

Schließlich folgt die Behauptung aus Lemma 34. □

3.3 Asymptotische kritische Werte

Um einen Test für das epidemische Changepoint-Problem zu konstruieren, haben wir in Abschnitt 3.1 die Teststatistik $T_n(\alpha,\beta)$ eingeführt, deren exakte Verteilung i. Allg. nicht bekannt ist. Wollen wir dennoch einen Test mit T_n konstruieren, der zumindest asymptotisch das Niveau γ hat, so ist ein Ansatz, die in Abschnitt 3.2 hergeleitete Grenzvariable

$$
Y = \sup_{\underline{\mathbf{0}}\le\mathbf{s}<\mathbf{t}\le\underline{\mathbf{1}},\ \alpha\le[\mathbf{t}-\mathbf{s}]\le1-\beta} \frac{\left|\sum\limits_{\boldsymbol{\varepsilon}\in\{0,1\}^d}(-1)^{d-\sum_{i=1}^d\varepsilon_i}W(\mathbf{s}+\boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s}))-[\mathbf{t}-\mathbf{s}]W(\underline{\mathbf{1}})\right|}{\sqrt{[\mathbf{t}-\mathbf{s}](1-[\mathbf{t}-\mathbf{s}])}}
$$

zu betrachten. Der Test hat dann die Form

$$\phi(x) = \begin{cases} 1, & T_n(x) \geq c_\gamma \\ 0, & T_n(x) < c_\gamma \end{cases},$$

wobei $x = (x_{\mathbf{i}})_{\mathbf{i} \in \{1,\ldots,n\}^d}$ die Beobachtungen sind und c_γ geeignet gewählt wird. Die Nullhypothese wird also verworfen, wenn $T_n(x) \geq c_\gamma$. Um den Fehler 1. Art abzuschätzen bzw. c_γ passend zu wählen, brauchen wir eine Approximation der Wahrscheinlichkeit von $T_n(X) \geq c_\gamma$ unter der Nullhypothese. Dafür betrachten wir statt $T_n(X) \geq c_\gamma$ das Ereignis $Y \geq c_\gamma$, dass die Grenzvariable den kritischen Wert überschreitet. Auch diese Wahrscheinlichkeit werden wir nicht exakt bestimmen, sondern durch Untersuchen des Verhaltens von $P(Y \geq u)$ für $u \to \infty$ approximieren. Wir beschränken uns dabei auf die Fälle $d = 1, 2$.

Wir definieren $\phi(c) := \frac{1}{\sqrt{2\pi}} e^{-c^2/2}$,

$$C_1(\alpha, \beta) := \frac{1}{4} \left(\log \frac{(1-\beta)(1-\alpha)}{\beta\alpha} + \frac{1}{\alpha} - \frac{1}{1-\beta} \right)$$

und

$$C_2(\alpha, \beta) := \int_{1-\beta}^{\alpha} \frac{-2(1+\xi) - (1+\xi)\log\xi}{16\xi^2(1-\xi)^4} d\xi.$$

Weiter seien

$$D := \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^{2d} : \ \underline{\mathbf{0}} \leq \mathbf{x} < \mathbf{y} \leq \underline{\mathbf{1}}, \ \alpha \leq [\mathbf{y} - \mathbf{x}] \leq 1 - \beta\}$$

und

$$D_\delta := \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^{2d} : \ \underline{\delta} \leq \mathbf{x} < \mathbf{y} \leq \underline{\mathbf{1}} - \underline{\delta}, \ \alpha \leq [\mathbf{y} - \mathbf{x}] \leq 1 - \beta\}$$

für $0 < \delta < \frac{1}{2}$ und $0 < \alpha < 1 - \beta < 1$. Ziel dieses Abschnittes ist nun der Beweis des folgenden Satzes:

Satz 53. *Für $u \to \infty$ und $d = 1, 2$ gilt:*

$$P\left(\sup_{(\mathbf{x},\mathbf{y}) \in D} |X(\mathbf{s}, \mathbf{t})| > u \right) \sim 2\, C_d(\alpha, \beta) u^{4d-1} \phi(u),$$

wobei $\{X(\mathbf{s}, \mathbf{t})\}_{(\mathbf{s},\mathbf{t}) \in D}$ ein Zufallsfeld der Form

$$X(\mathbf{s}, \mathbf{t}) := \frac{Z(\mathbf{s}, \mathbf{t})}{\sqrt{[\mathbf{t} - \mathbf{s}](1 - [\mathbf{t} - \mathbf{s}])}}$$

mit

$$Z(\mathbf{s}, \mathbf{t}) := \begin{cases} B(t) - B(s), & \text{für } d = 1 \\ W(\mathbf{t}) - W(t_1, s_2) - W(s_1, t_2) + W(\mathbf{s}) - [\mathbf{t} - \mathbf{s}] W(\underline{\mathbf{1}}), & \text{für } d = 2 \end{cases}$$

sei.

Der Beweis von Satz 53 verläuft in mehreren Schritten. Zunächst zeigen wir, dass die Kovarianzfunktion von $\{X(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in D}$ (für $(\mathbf{s},\mathbf{t}) \neq (\tilde{\mathbf{s}},\tilde{\mathbf{t}})$) kleiner als 1 und auf D_δ sogar betraglich kleiner als 1 ist. Dafür ist es hilfreich, den Prozess mit Hilfe eines Weißen Rauschens auszudrücken. Da für die Brownsche Brücke $B(t) \overset{\mathcal{D}}{=} W(t) - tW(1)$ gilt, lässt sich $Z(\mathbf{s},\mathbf{t})$ auch schreiben als

$$Z(\mathbf{s},\mathbf{t}) \overset{\mathcal{D}}{=} \sum_{\boldsymbol{\varepsilon}\in\{0,1\}^d} (-1)^{d-\sum_{i=1}^d \varepsilon_i} W(\mathbf{s}+\boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s})) - [\mathbf{t}-\mathbf{s}]W(\underline{\mathbf{1}})$$

für $d = 1,2$. Nach Bemerkung 7 kann man dies durch ein Weißes Rauschen $\{W(A) : A \in \mathcal{B}^q\}$ (siehe Definition 6) beschreiben:

$$Z(\mathbf{s},\mathbf{t}) \overset{\mathcal{D}}{=} W((\mathbf{s},\mathbf{t}]) - \lambda^d((\mathbf{s},\mathbf{t}])W((\underline{\mathbf{0}},\underline{\mathbf{1}}]),$$

wobei ausgenutzt wird, dass $[\mathbf{t}-\mathbf{s}] = \lambda^d((\mathbf{s},\mathbf{t}])$ für $\mathbf{s} \leq \mathbf{t}$. Dann gilt für $(\mathbf{s},\mathbf{t}) \neq (\tilde{\mathbf{s}},\tilde{\mathbf{t}})$ in D:

$$\begin{aligned}
&\operatorname{Cov}(X(\mathbf{s},\mathbf{t}), X(\tilde{\mathbf{s}},\tilde{\mathbf{t}})) \\
=&\frac{E[\{W(\mathbf{s},\mathbf{t}] - \lambda^d((\mathbf{s},\mathbf{t}])W((\underline{\mathbf{0}},\underline{\mathbf{1}}])\}\{W(\tilde{\mathbf{s}},\tilde{\mathbf{t}}] - \lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])W((\underline{\mathbf{0}},\underline{\mathbf{1}}])\}]}{\sqrt{\lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])(1-\lambda^d((\mathbf{s},\mathbf{t}]))(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}} \\
=&\frac{\lambda^d((\mathbf{s},\mathbf{t}] \cap (\tilde{\mathbf{s}},\tilde{\mathbf{t}}]) - \lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])}{\sqrt{\lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])(1-\lambda^d((\mathbf{s},\mathbf{t}]))(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}},
\end{aligned} \tag{3.10}$$

wobei auch hier $[\mathbf{t}-\mathbf{s}] = \lambda^d((\mathbf{s},\mathbf{t}])$ genutzt wurde. Mit Hilfe dieser Darstellung der Kovarianzfunktion können wir nun folgendes Lemma beweisen:

Lemma 54. *Für $(\mathbf{s},\mathbf{t}) \neq (\tilde{\mathbf{s}},\tilde{\mathbf{t}})$ in D gilt*

$$\operatorname{Cov}(X(\mathbf{s},\mathbf{t}), X(\tilde{\mathbf{s}},\tilde{\mathbf{t}})) < 1.$$

Ist zusätzlich $(\mathbf{s},\mathbf{t}), (\tilde{\mathbf{s}},\tilde{\mathbf{t}}) \in D_\delta$, so gilt auch

$$\operatorname{Cov}(X(\mathbf{s},\mathbf{t}), X(\tilde{\mathbf{s}},\tilde{\mathbf{t}})) > -1.$$

Beweis. Seien $(\mathbf{s},\mathbf{t}) \neq (\tilde{\mathbf{s}},\tilde{\mathbf{t}})$ in D gegeben.
<u>1. Fall:</u> $(\mathbf{s},\mathbf{t}] \cap (\tilde{\mathbf{s}},\tilde{\mathbf{t}}] = \emptyset$: Dann sieht man anhand von (3.10), dass $\operatorname{Cov}(X(\mathbf{s},\mathbf{t}), X(\tilde{\mathbf{s}},\tilde{\mathbf{t}})) < 0 < 1$ ist. Für $(\mathbf{s},\mathbf{t}) \neq (\tilde{\mathbf{s}},\tilde{\mathbf{t}})$ in D_δ gilt

$$\lambda^d((\mathbf{s},\mathbf{t}]) + \lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]) = \lambda^d((\mathbf{s},\mathbf{t}] \cup (\tilde{\mathbf{s}},\tilde{\mathbf{t}}]) < \lambda^d((0,1]^d) = 1,$$

und somit:

$$\operatorname{Cov}(X(\mathbf{s},\mathbf{t}), X(\tilde{\mathbf{s}},\tilde{\mathbf{t}})) = -\sqrt{\frac{\lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])}{(1-\lambda^d((\mathbf{s},\mathbf{t}]))(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}} > -1,$$

denn es gilt

$$\begin{aligned}
\lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]) &< \overbrace{1-\lambda^d((\mathbf{s},\mathbf{t}])-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])}^{>0} + \lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]) \\
&= (1-\lambda^d((\mathbf{s},\mathbf{t}]))(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])).
\end{aligned}$$

<u>2. Fall:</u> $(\mathbf{s},\mathbf{t}] \cap (\tilde{\mathbf{s}},\tilde{\mathbf{t}}] = (\mathbf{a},\mathbf{b}]$ mit $\mathbf{a} < \mathbf{b}$: Sei o. E.

$$\lambda^d((\mathbf{s},\mathbf{t}]) = \min\{\lambda^d((\mathbf{s},\mathbf{t}]), \lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])\}.$$

Falls $(\mathbf{a},\mathbf{b}]$ eine echte Teilmenge von $(\mathbf{s},\mathbf{t}]$ und $(\tilde{\mathbf{s}},\tilde{\mathbf{t}}]$ ist, gilt $\lambda^d((\mathbf{a},\mathbf{b}]) < \lambda^d((\mathbf{s},\mathbf{t}])$, und es folgt:

$$\begin{aligned}
&\operatorname{Cov}(X(\mathbf{s},\mathbf{t}), X(\tilde{\mathbf{s}},\tilde{\mathbf{t}})) \\
=&\frac{\lambda^d((\mathbf{a},\mathbf{b}]) - \lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])}{\sqrt{\lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])(1-\lambda^d((\mathbf{s},\mathbf{t}]))(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}} \\
<&\frac{\lambda^d((\mathbf{s},\mathbf{t}])(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}{\sqrt{\lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])(1-\lambda^d((\mathbf{s},\mathbf{t}]))(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}} \\
=&\underbrace{\sqrt{\frac{\lambda^d((\mathbf{s},\mathbf{t}])}{\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])}}}_{\leq 1}\underbrace{\sqrt{\frac{(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}{(1-\lambda^d((\mathbf{s},\mathbf{t}]))}}}_{\leq 1} \leq 1
\end{aligned}$$

Ist zusätzlich $\lambda^d((\mathbf{a},\mathbf{b}]) \geq \lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])$, so ist die Kovarianz nichtnegativ, also insbesondere größer als -1. Ist $\lambda^d((\mathbf{a},\mathbf{b}]) < \lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])$, so gilt:

$$\begin{aligned}
&\operatorname{Cov}(X(\mathbf{s},\mathbf{t}), X(\tilde{\mathbf{s}},\tilde{\mathbf{t}})) \\
&= -\sqrt{\frac{\big(\lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]) - \lambda^d((\mathbf{a},\mathbf{b}])\big)^2}{\lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])(1-\lambda^d((\mathbf{s},\mathbf{t}]))(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}}
\end{aligned}$$

Um die Übersichtlichkeit zu erhöhen, schreiben wir $L := \lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])$. Wegen $\lambda^d((\mathbf{a},\mathbf{b}]) < L$ ist nun

$$\big(L - \lambda^d((\mathbf{a},\mathbf{b}])\big)^2 = L^2 - 2\lambda^d((\mathbf{a},\mathbf{b}])L + \lambda^d((\mathbf{a},\mathbf{b}])^2 < L^2 - \lambda^d((\mathbf{a},\mathbf{b}])L,$$

und somit:

$$\begin{aligned}
&\operatorname{Cov}(X(\mathbf{s},\mathbf{t}), X(\tilde{\mathbf{s}},\tilde{\mathbf{t}})) \\
&= -\sqrt{\frac{(L-\lambda^d((\mathbf{a},\mathbf{b}]))^2}{L(1-\lambda^d((\mathbf{s},\mathbf{t}]))(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}} \\
&> -\sqrt{\frac{L^2-\lambda^d((\mathbf{a},\mathbf{b}])L}{L^2+L(1-\lambda^d((\mathbf{s},\mathbf{t}])-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}}
\end{aligned}$$

Der letzte Term ist größer oder gleich -1 genau dann, wenn

$$L^2 - \lambda^d((\mathbf{a},\mathbf{b}])L \leq L^2 + L(1 - \lambda^d((\mathbf{s},\mathbf{t}]) - \lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])),$$

und dies ist äquivalent zu

$$0 \leq L(1 - \lambda^d((\mathbf{s},\mathbf{t}]) - \lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]) + \lambda^d((\mathbf{a},\mathbf{b}])).$$

Da $L > 0$ ist, folgt die behauptete Abschätzung nun aus der Tatsache, dass

$$\lambda^d((\mathbf{s},\mathbf{t}]) + \lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]) - \lambda^d((\mathbf{a},\mathbf{b}]) = \lambda^d((\mathbf{s},\mathbf{t}] \cup (\tilde{\mathbf{s}},\tilde{\mathbf{t}}]) \leq \lambda^d((0,1]^d) = 1$$

gilt. Ist $(\mathbf{a},\mathbf{b}]$ keine echte Teilmenge von $(\mathbf{s},\mathbf{t}]$ oder $(\tilde{\mathbf{s}},\tilde{\mathbf{t}}]$, so muss

$$(\mathbf{a},\mathbf{b}] = (\mathbf{s},\mathbf{t}] \subset (\tilde{\mathbf{s}},\tilde{\mathbf{t}}]$$

gelten, und, da $(\mathbf{s},\mathbf{t}) \neq (\tilde{\mathbf{s}},\tilde{\mathbf{t}})$ muss dann auch gelten

$$\lambda^d((\mathbf{s},\mathbf{t}]) < \lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]).$$

Somit ergibt sich:

$$\begin{aligned}
&\operatorname{Cov}(X(\mathbf{s},\mathbf{t}), X(\tilde{\mathbf{s}},\tilde{\mathbf{t}})) \\
=&\frac{\lambda^d((\mathbf{a},\mathbf{b}]) - \lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])}{\sqrt{\lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])(1-\lambda^d((\mathbf{s},\mathbf{t}]))(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}} \\
=&\frac{\lambda^d((\mathbf{s},\mathbf{t}])(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}{\sqrt{\lambda^d((\mathbf{s},\mathbf{t}])\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])(1-\lambda^d((\mathbf{s},\mathbf{t}]))(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}} \\
=&\underbrace{\sqrt{\frac{\lambda^d((\mathbf{s},\mathbf{t}])}{\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}])}}}_{<1}\underbrace{\sqrt{\frac{(1-\lambda^d((\tilde{\mathbf{s}},\tilde{\mathbf{t}}]))}{(1-\lambda^d((\mathbf{s},\mathbf{t}]))}}}_{<1} < 1
\end{aligned}$$

Wie aus obiger Überlegung ersichtlich wird, ist die Kovarianz in diesem Fall größer als 0, also insbesondere größer als -1. □

Jarušková und Piterbarg (2011) verwenden in ihrem Artikel Theorem 7.1 aus Piterbarg (1996), um Aussagen über das Tail-Verhalten der von ihnen betrachteten Variablen - und damit über asymptotische kritische Werte für ihre Tests - zu erhalten. Wir werden nicht direkt auf Theorem 7.1, sondern auf Theorem A.1 von Jarušková (2011), ergänzt um die Voraussetzung (3.12), zurückgreifen, da dieses für unsere Belange direkter einsetzbar als Theorem 7.1 ist. Jarušková (2011) zeigt in ihrem Artikel, dass ein Prozess, dessen Kovarianzfunktion die Entwicklung (3.11) zulässt, lokal stationär ist, und folgert (3.13) als Korollar aus Theorem 7.1 von Piterbarg (1996).

Lemma 55. *(vgl. Jarušková (2011), Theorem A.1)* $\{X(\mathbf{t})\}_{\mathbf{t}\in K}$ *sei ein zentrierter, standardisierter Gauß'scher Prozess, der auf einer kompakten Menge* $K \subset \mathbb{R}^m$ *(*$m \in \mathbb{N}$*) definiert sei. Sei* $p \in \{1, \dots, m\}$*. Für* $\mathbf{x}, \mathbf{y} \in K$ *habe die Kovarianzfunktion*

$$r(\mathbf{x}, \mathbf{y}) = E[X(\mathbf{x})X(\mathbf{y})]$$

die Entwicklung

$$\begin{aligned} &r(x_1, \dots, x_m; x_1 + h_1, \dots, x_m + h_m) \\ =&1 - c_1(x_1, \dots, x_m)|h_1| - \dots - c_p(x_1, \dots, x_m)|h_p| \\ &- c_{p+1}(x_1, \dots, x_m)h_{p+1}^2 - \dots - c_m(x_1, \dots, x_m)h_m^2 \\ &+o(|h_1| + \dots + |h_p| + h_{p+1}^2 + \dots + h_m^2) \text{ für } h_1 \to 0, \dots, h_m \to 0, \end{aligned} \tag{3.11}$$

wobei $c_i(x_1, \dots, x_m)$ $(i = 1, \dots, m)$ *stetige Funktionen auf* K *seien. Weiter sei*

$$\{\mathbf{x} \in K : \ c_1(\mathbf{x}) = 0 \cup c_2(\mathbf{x}) = 0 \cup \dots \cup c_m(\mathbf{x}) = 0\}$$

eine Nullmenge bezüglich des Lebesgue-Maßes. Dann ist $\{X(\mathbf{t})\}_{\mathbf{t}\in K}$ *ein stochastischer Prozess mit stationärer lokaler Struktur. Ist zusätzlich*

$$r(\mathbf{s}, \mathbf{t}) < 1 \tag{3.12}$$

für alle $\mathbf{s}, \mathbf{t} \in K$ *mit* $\mathbf{s} \neq \mathbf{t}$*, so gilt:*

$$P(\sup_{\mathbf{t}\in K} X(\mathbf{t}) > u) = \frac{1}{\pi^{(m-p)/2}} I_K u^{m+p}(1 - \Phi(u))(1 + o(1)) \text{ für } u \to \infty, \tag{3.13}$$

wobei

$$I_K := \int \underset{K}{\cdots} \int c_1(\mathbf{x}) \cdots c_p(\mathbf{x})(c_{p+1}(\mathbf{x}))^{1/2} \cdots (c_m(\mathbf{x}))^{1/2} dx_1 \dots dx_m$$

und Φ *die Verteilungsfunktion der Standardnormalverteilung ist.*

Zunächst stellen wir fest, dass der von uns betrachtete Prozess auf D die Bedingung (3.11) erfüllt:

Lemma 56. *Unter den Voraussetzungen von Satz 53 gilt für* $(\mathbf{s}, \mathbf{t}), (\mathbf{s} + \mathbf{h}, \mathbf{t} + \mathbf{f}) \in D$ *und* $\mathbf{h}, \mathbf{f} \to 0$*:*

$$E[X(s,t)X(s + h, t + f)] = 1 - \frac{|h| + |f|}{2(t - s)(1 - (t - s))} + o(|h| + |f|)$$

für $d = 1$ *und*

$$\begin{aligned} &E[X(\mathbf{s}, \mathbf{t})X(\mathbf{s} + \mathbf{h}, \mathbf{t} + \mathbf{f})] \\ =&1 - \frac{(t_1 - s_1)(|h_2| + |f_2|) + (t_2 - s_2)(|h_1| + |f_1|)}{2[\mathbf{t} - \mathbf{s}](1 - [\mathbf{t} - \mathbf{s}])} + o(|h_1| + |h_2| + |f_1| + |f_2|) \end{aligned}$$

für $d = 2$.

Beweis. siehe Anhang □

Der folgende Satz ist eine leicht veränderte Version der Theoreme 1 und 2 aus Jarušková und Piterbarg (2011):

Satz 57. *Unter den Voraussetzungen von Satz 53 gilt für $u \to \infty$ und $d = 1, 2$:*

$$P\left(\sup_{(\mathbf{x},\mathbf{y})\in D} X(\mathbf{s},\mathbf{t}) > u\right) \sim u^{4d-1}\phi(u)\int_D H(\mathbf{s},\mathbf{t})d(\mathbf{s},\mathbf{t}),$$

mit

$$H(\mathbf{s},\mathbf{t}) := \begin{cases} \frac{1}{4(t-s)^2(1-(t-s))^2}, & d=1 \\ \frac{1}{16[\mathbf{t}-\mathbf{s}]^2(1-[\mathbf{t}-\mathbf{s}])^4}, & d=2 \end{cases} \tag{3.14}$$

und

$$\int_D H(\mathbf{s},\mathbf{t})d(\mathbf{s},\mathbf{t}) = C_d(\alpha,\beta). \tag{3.15}$$

Beweis. Der Beweis entspricht den Beweisen aus Jarušková und Piterbarg (2011) (vgl. S. 554, 556) in leicht abgewandelter und detaillierterer Form. Wir machen mehrere Schritte:

1. Schritt: Der Bereich D entspricht den Voraussetzungen aus Lemma 55:
$D \subset [\underline{\mathbf{0}}, \underline{\mathbf{1}}]^{2d}$ ist offensichtlich beschränkt. Da

$$\begin{aligned} D &= \{(\mathbf{s},\mathbf{t}) : \ \underline{\mathbf{0}} \le \mathbf{s} < \mathbf{t} \le \underline{\mathbf{1}}, \alpha \le [\mathbf{t}-\mathbf{s}] \le 1-\beta\} \\ &= \{(\mathbf{s},\mathbf{t}) : \ \underline{\mathbf{0}} \le \mathbf{s} \le \mathbf{t} \le \underline{\mathbf{1}}, \alpha \le [\mathbf{t}-\mathbf{s}] \le 1-\beta\} \end{aligned}$$

ist D offensichtlich auch abgeschlossen und somit kompakt.

2. Schritt: Wie im Beweis der Bemerkung 46 gezeigt, ist $X(\mathbf{s},\mathbf{t})$ für alle $(\mathbf{s},\mathbf{t})$ zentriert mit Varianz 1. Lemma 54 liefert die Eigenschaft (3.12). Nach Lemma 56 besitzt der Prozess die Eigenschaft (3.11) mit Funktionen

$$c_1(s,t) = c_2(s,t) = \frac{1}{2(t-s)(1-(t-s))},$$

für $d = 1$ und

$$c_1(\mathbf{s},\mathbf{t}) = c_3(\mathbf{s},\mathbf{t}) = \frac{t_2 - s_2}{2[\mathbf{t}-\mathbf{s}](1-[\mathbf{t}-\mathbf{s}])}$$

und

$$c_2(\mathbf{s},\mathbf{t}) = c_4(\mathbf{s},\mathbf{t}) = \frac{t_1 - s_1}{2[\mathbf{t}-\mathbf{s}](1-[\mathbf{t}-\mathbf{s}])},$$

für $d = 2$. $H(\mathbf{s},\mathbf{t})$ aus (3.14) hat dann die Form

$$H(\mathbf{s},\mathbf{t}) = \prod_{i=1}^{2d} c_i(\mathbf{s},\mathbf{t}).$$

3. Schritt: Wir benötigen nun die Werte $\int_D H(\mathbf{s},\mathbf{t})d(\mathbf{s},\mathbf{t})$. Zunächst für $d=1$: Der Transformationssatz liefert mit der Transformation

$$h_1(s,t) = (u,x) := (s, t-s):$$

$$\begin{aligned}
\int_D H(s,t)d(s,t) &= \int_0^1 \int_0^1 I_D(s,t)\frac{1}{4(t-s)^2(1-(t-s))^2}dsdt \\
&= \int_\alpha^{1-\beta} \int_0^{1-x} \frac{1}{4x^2(1-x)^2}dudx \\
&= \int_\alpha^{1-\beta} \frac{1}{4x^2(1-x)}dx \\
&= \frac{1}{4}\int_\alpha^{1-\beta} \left(\frac{1}{x^2}+\frac{1}{x}+\frac{1}{1-x}\right)dx \\
&= \frac{1}{4}\left[-\frac{1}{x}+\log(x)-\log(1-x)\right]_\alpha^{1-\beta} \\
&= \frac{1}{4}\left(\frac{1}{\alpha}-\frac{1}{1-\beta}+\log\frac{(1-\alpha)(1-\beta)}{\alpha\beta}\right) \\
&= C_1(\alpha,\beta)
\end{aligned}$$

Analog zum Fall $d=1$ verwenden wir für $d=2$ den Transformationssatz, jetzt mit der Transformation

$$h_2(s_1,s_2,t_1,t_2) = (u_1,u_2,x_1,x_2) := (s_1,s_2,t_1-s_1,t_2-s_2):$$

$$\begin{aligned}
&\int_D H(\mathbf{s},\mathbf{t})d(\mathbf{s},\mathbf{t}) \\
&= \int_D \frac{1}{16(t_1-s_1)^2(t_2-s_2)^2(1-(t_1-s_1)(t_2-s_2))^4}d(\mathbf{s},\mathbf{t}) \\
&= \int_0^1\int_0^1 I_{\{\alpha\le x_1x_2\le 1-\beta\}}(\mathbf{x})\int_0^{1-x_2}\int_0^{1-x_1}\frac{1}{16x_1^2x_2^2(1-x_1x_2)^4}du_1du_2dx_1dx_2 \\
&= \int_0^1\int_0^1 I_{\{\alpha\le x_1x_2\le 1-\beta\}}(\mathbf{x})\frac{(1-x_1)(1-x_2)}{16x_1^2x_2^2(1-x_1x_2)^4}dx_1dx_2 \\
&= \int_0^1\int_0^1 I_{\{\alpha\le x_1x_2\le 1-\beta\}}(\mathbf{x})\frac{(1-x_1)(x_1-x_1x_2)}{16x_1^2x_2^2(1-x_1x_2)^4x_1}dx_1dx_2
\end{aligned}$$

Eine zweite Anwendung des Transformationssatzes mit der Transformation

$$h_3(x_1,x_2) = (\eta,\xi) := (x_1, x_1x_2)$$

liefert:

$$\int_0^1 \int_0^1 I_{\{\alpha \le x_1 x_2 \le 1-\beta\}}(\mathbf{x}) \frac{(1-x_1)(x_1 - x_1x_2)}{16x_1^2x_2^2(1-x_1x_2)^4 x_1} dx_1 dx_2$$
$$= \int_\alpha^{1-\beta} \frac{1}{16\xi^2(1-\xi)^4} \int_\xi^1 \frac{(1-\eta)(\eta-\xi)}{\eta^2} d\eta d\xi$$

Mit partieller Integration erhält man schließlich für beliebiges $\alpha \le \xi \le 1-\beta$:

$$\begin{aligned}\int_\xi^1 \frac{(1-\eta)(\eta-\xi)}{\eta^2} d\eta &= \int_\xi^1 (\eta^2 + (-1-\xi)\eta + \xi) \left(-\frac{1}{\eta^2}\right) d\eta \\ &= [\eta - 1 - \xi + \xi/\eta]_\xi^1 - \int_\xi^1 \left(2 - \frac{1+\xi}{\eta}\right) d\eta \\ &= -[2\eta - (1+\xi)\log\eta]_\xi^1 \\ &= -2(1-\xi) - (1+\xi)\log\xi,\end{aligned}$$

also

$$\int_D H(\mathbf{s},\mathbf{t}) d(\mathbf{s},\mathbf{t}) = C_2(\alpha,\beta).$$

Lemma 55 zusammen mit der Feststellung, dass

$$1 - \Phi(u) \sim \frac{1}{\sqrt{2\pi}u} \exp\left(-\frac{u^2}{2}\right) = \frac{1}{u}\phi(u)$$

ist, liefert nun die Behauptung. □

Nun können wir Satz 57 benutzen, um eine Aussage über das Supremum des Betrages des Prozesses zu machen. Dies machen wir nach einem Vorschlag von Dr. Zakhar Kabluchko (persönliche Korrespondenz, 28.11.2012). Für diese Aussage benötigen wir allerdings, dass die Kovarianzfunktion des Prozesses betraglich kleiner als 1 ist (vgl. Lemma 54), weshalb wir zunächst den Bereich, über den das Supremum genommen wird, etwas einschränken.

Korollar 58. *Unter den Voraussetzungen von Satz 57 gilt für beliebiges* $0 < \delta < \frac{1}{2}$*:*

$$P\left(\sup_{(\mathbf{x},\mathbf{y})\in D_\delta} |X(\mathbf{s},\mathbf{t})| > u\right) \sim 2\, u^{4d-1}\phi(u) \int_{D_\delta} H(\mathbf{s},\mathbf{t}) d(\mathbf{s},\mathbf{t})$$

Beweis. Seien $D_\delta^{(1)} := D_\delta$ und

$$D_\delta^{(2)} = \{(\mathbf{s},\mathbf{t}) + \underline{\mathbf{2}} :\ (\mathbf{s},\mathbf{t}) \in D_\delta^{(1)}\}.$$

Wir betrachten das Zufallsfeld $\{Y(\mathbf{s},\mathbf{t})\}_{(\mathbf{s},\mathbf{t})\in D_\delta^{(1)}\cup D_\delta^{(2)}}$ mit

$$Y(\mathbf{s},\mathbf{t}) := \begin{cases} X(\mathbf{s},\mathbf{t}), & (\mathbf{s},\mathbf{t}) \in D_\delta^{(1)} \\ -X(\mathbf{s}-\underline{\mathbf{2}},\mathbf{t}-\underline{\mathbf{2}}), & (\mathbf{s},\mathbf{t}) \in D_\delta^{(2)}.\end{cases}$$

Dann gilt

$$P\left(\sup_{(\mathbf{s},\mathbf{t})\in D_\delta^{(1)}\cup D_\delta^{(1)}} Y(\mathbf{s},\mathbf{t})>u\right)=P\left(\sup_{(\mathbf{s},\mathbf{t})\in D_\delta}|X(\mathbf{s},\mathbf{t})|>u\right)$$

und somit reicht es zu zeigen, dass

$$P\left(\sup_{(\mathbf{s},\mathbf{t})\in D_\delta^{(1)}\cup D_\delta^{(2)}} Y(\mathbf{s},\mathbf{t})>u\right)\sim 2\,u^{4d-1}\phi(u)\int\limits_{D_\delta^{(1)}} H(\mathbf{s},\mathbf{t}).$$

Dies ergibt sich aus Satz 57: Für $(\mathbf{s},\mathbf{t}),(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in D_\delta^{(1)}\cup D_\delta^{(2)}$ ist

$$\begin{aligned}
&\operatorname{Cov}(Y(\mathbf{s},\mathbf{t}),Y(\tilde{\mathbf{s}},\tilde{\mathbf{t}}))\\
&=\begin{cases}\operatorname{Cov}(X(\mathbf{s},\mathbf{t}),X(\tilde{\mathbf{s}},\tilde{\mathbf{t}})), & (\mathbf{s},\mathbf{t}),(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in D_\delta^{(1)}\\
\operatorname{Cov}(X(\mathbf{s}-\underline{\mathbf{2}},\mathbf{t}-\underline{\mathbf{2}}),X(\tilde{\mathbf{s}}-\underline{\mathbf{2}},\tilde{\mathbf{t}}-\underline{\mathbf{2}})), & (\mathbf{s},\mathbf{t}),(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in D_\delta^{(2)}\\
-\operatorname{Cov}(X(\mathbf{s},\mathbf{t}),X(\tilde{\mathbf{s}}-\underline{\mathbf{2}},\tilde{\mathbf{t}}-\underline{\mathbf{2}})), & (\mathbf{s},\mathbf{t})\in D_\delta^{(1)},(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in D_\delta^{(2)}\\
-\operatorname{Cov}(X(\mathbf{s}-\underline{\mathbf{2}},\mathbf{t}-\underline{\mathbf{2}}),X(\tilde{\mathbf{s}},\tilde{\mathbf{t}})), & (\mathbf{s},\mathbf{t})\in D_\delta^{(2)},(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in D_\delta^{(1)}.\end{cases}
\end{aligned}$$

Nach Lemma 54 ist somit $\operatorname{Cov}(Y(\mathbf{s},\mathbf{t}),Y(\tilde{\mathbf{s}},\tilde{\mathbf{t}}))<1$. Für $\|(\mathbf{s},\mathbf{t})-(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\|\to 0$ können wir o. E. davon ausgehen, dass $(\mathbf{s},\mathbf{t})$ und $(\tilde{\mathbf{s}},\tilde{\mathbf{t}})$ entweder beide in $D_\delta^{(1)}$ oder beide in $D_\delta^{(2)}$ liegen. Wie wir in Lemma 56 und im Beweis von Satz 57 gezeigt haben, gibt es für $(\mathbf{s},\mathbf{t}),(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in D_\delta^{(1)}$ mit $\|(\mathbf{s},\mathbf{t})-(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\|\to 0$ Funktionen $c_i(\mathbf{t}-\mathbf{s})$, $i=1,\dots,2d$, so dass

$$\begin{aligned}
&\operatorname{Cov}(X(\mathbf{s},\mathbf{t}),X(\tilde{\mathbf{s}},\tilde{\mathbf{t}}))\\
=&1-\sum_{i=1}^{d}c_i(\mathbf{t}-\mathbf{s})|\tilde{s}_i-s_i|-\sum_{i=d+1}^{2d}c_i(\mathbf{t}-\mathbf{s})|\tilde{t}_i-t_i|+o(\|\tilde{\mathbf{s}}-\mathbf{s}\|_1+\|\tilde{\mathbf{t}}-\mathbf{t}\|_1),
\end{aligned}$$

wobei $\|x\|_1:=\sum_{i=1}^{d}|x_i|$ für $x\in\mathbb{R}^d$ sei. Folglich gilt für $(\mathbf{s},\mathbf{t}),(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in D_\delta^{(1)}\cup D_\delta^{(2)}$ mit $\|(\mathbf{s},\mathbf{t})-(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\|\to 0$:

$$\begin{aligned}
&\operatorname{Cov}(Y(\mathbf{s},\mathbf{t}),Y(\tilde{\mathbf{s}},\tilde{\mathbf{t}}))\\
=&\begin{cases}\operatorname{Cov}(X(\mathbf{s},\mathbf{t}),X(\tilde{\mathbf{s}},\tilde{\mathbf{t}})), & (\mathbf{s},\mathbf{t}),(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in D_\delta^{(1)}\\
\operatorname{Cov}(X(\mathbf{s}-\underline{\mathbf{2}},\mathbf{t}-\underline{\mathbf{2}}),X(\tilde{\mathbf{s}}-\underline{\mathbf{2}},\tilde{\mathbf{t}}-\underline{\mathbf{2}})), & (\mathbf{s},\mathbf{t}),(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in D_\delta^{(2)}\end{cases}\\
=&1-\sum_{i=1}^{d}c_i(\mathbf{t}-\mathbf{s})|\tilde{s}_i-s_i|-\sum_{i=d+1}^{2d}c_i(\mathbf{t}-\mathbf{s})|\tilde{t}_i-t_i|+o(\|\tilde{\mathbf{s}}-\mathbf{s}\|_1+\|\tilde{\mathbf{t}}-\mathbf{t}\|_1)
\end{aligned}$$

Da $D_\delta^{(1)}\cup D_\delta^{(2)}$ als endliche Vereinigung kompakter Mengen kompakt ist, liefert Lemma 55:

$$P\left(\sup_{(\mathbf{s},\mathbf{t})\in D_\delta^{(1)}\cup D_\delta^{(2)}} Y(\mathbf{s},\mathbf{t})>u\right)\sim u^{4d-1}\phi(u)\int\limits_{D_\delta^{(1)}\cup D_\delta^{(2)}} H(\mathbf{s},\mathbf{t})d(\mathbf{s},\mathbf{t}),$$

wobei

$$H(\mathbf{s},\mathbf{t}) := \prod_{i=1}^{2d} c_i(\mathbf{t}-\mathbf{s})$$

wie im Beweis von Satz 57 ist. Die c_i hängen nur von der Differenz $\mathbf{t}-\mathbf{s}$ ab, die unter Translationen erhalten bleibt. Folglich gilt

$$\int\limits_{D_\delta^{(2)}} H(\mathbf{s},\mathbf{t})d(\mathbf{s},\mathbf{t}) = \int\limits_{D_\delta^{(1)}} H(\mathbf{s},\mathbf{t})d(\mathbf{s},\mathbf{t}),$$

und somit folgt die Behauptung. □

Im Beweis von Satz 53 werden wir Aussagen über den Limes superior und Limes inferior von Produkten beschränkter Folgen machen. Dort benötigen wir für Abschätzungen folgendes elementare Hilfsmittel aus der Analysis:

Lemma 59. *(vgl. Timmann (2006), S. 92) $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ seien beschränkte Folgen in $\mathbb{R}_+$. Dann gilt:*

(a) Ist $a_n \leq b_n$ für alle $n \in \mathbb{N}$, so gilt

$$\limsup_{n\to\infty} a_n \leq \limsup_{n\to\infty} b_n \quad \textit{und} \quad \liminf_{n\to\infty} a_n \leq \liminf_{n\to\infty} b_n.$$

(b) $\limsup_{n\to\infty} a_n b_n \leq \limsup_{n\to\infty} a_n \limsup_{n\to\infty} b_n$

(c) $\liminf_{n\to\infty} a_n b_n \geq \liminf_{n\to\infty} a_n \liminf_{n\to\infty} b_n$

Nun können wir Satz 53 beweisen:

Beweis von Satz 53. Es sei $H(\mathbf{s},\mathbf{t})$ gemäß (3.14) definiert und $C(u) := \phi(u)u^{4d-1}$ für $u \in \mathbb{R}$. Nach dem Satz von Lebesgue gilt dann

$$C_\delta := \int_{D_\delta} H(\mathbf{s},\mathbf{t})d(\mathbf{s},\mathbf{t}) \overset{\delta\to 0}{\longrightarrow} \int_D H(\mathbf{s},\mathbf{t})d(\mathbf{s},\mathbf{t}) =: C.$$

Für $u \to \infty$ liefert Satz 57

$$a_1(u) := P\left(\sup_{(\mathbf{s},\mathbf{t})\in D} X(\mathbf{s},\mathbf{t}) > u\right) \sim C(u)\int_D H(\mathbf{s},\mathbf{t})d(\mathbf{s},\mathbf{t}) =: \frac{1}{2}b(u) \tag{3.16}$$

und für alle $0 < \delta < \frac{1}{2}$ folgt aus Korollar 58

$$a(u,\delta) := P\left(\sup_{(\mathbf{s},\mathbf{t})\in D_\delta} |X(\mathbf{s},\mathbf{t})| > u\right) \sim 2C(u)\int_{D_\delta} H(\mathbf{s},\mathbf{t})d(\mathbf{s},\mathbf{t}) =: b(u,\delta). \tag{3.17}$$

Zu zeigen ist nun, dass für $u \to \infty$ gilt:

$$a(u) := P\left(\sup_{(\mathbf{s},\mathbf{t})\in D} |X(\mathbf{s},\mathbf{t})| > u\right) \sim 2C(u)\int_D H(\mathbf{s},\mathbf{t})d(\mathbf{s},\mathbf{t}) = b(u) \tag{3.18}$$

Dazu zeigen wir zunächst $\limsup\limits_{u\to\infty} \frac{a(u)}{b(u)} \leq 1$. Für $\delta \in (0, \frac{1}{2})$ gilt:

$$\begin{aligned}
\limsup_{u\to\infty} \frac{a(u)}{b(u)} &= \limsup_{u\to\infty} \left(\frac{a(u,\delta)}{b(u,\delta)} \frac{a(u)}{a(u,\delta)} \underbrace{\frac{b(u,\delta)}{b(u)}}_{=C_\delta/C} \right) \\
&\leq \limsup_{u\to\infty} \left(\frac{a(u,\delta)}{b(u,\delta)} \right) \limsup_{u\to\infty} \left(\frac{a(u)}{a(u,\delta)} \right) \frac{C_\delta}{C}
\end{aligned}$$

Nach (3.17) ist

$$\limsup_{u\to\infty} \left(\frac{a(u,\delta)}{b(u,\delta)} \right) = 1.$$

Weiter gilt wegen (3.16) und (3.17):

$$\begin{aligned}
\limsup_{u\to\infty} \left(\frac{a(u)}{a(u,\delta)} \right) &\leq \limsup_{u\to\infty} \left(\frac{2a_1(u)}{a(u,\delta)} \right) \\
&= \limsup_{u\to\infty} \left(\frac{2a_1(u)}{b(u)} \frac{b(u)}{b(u,\delta)} \frac{b(u,\delta)}{a(u,\delta)} \right) \\
&\leq \frac{C}{C_\delta} \limsup_{u\to\infty} \left(\frac{2a_1(u)}{b(u)} \right) \limsup_{u\to\infty} \left(\frac{b(u,\delta)}{a(u,\delta)} \right) \\
&= \frac{C}{C_\delta}
\end{aligned}$$

Eingesetzt liefert dies:

$$\limsup_{u\to\infty} \frac{a(u)}{b(u)} \leq 1 \cdot \frac{C}{C_\delta} \cdot \frac{C_\delta}{C} = 1$$

Es bleibt nur noch zu zeigen, dass $\liminf\limits_{u\to\infty} \frac{a(u)}{b(u)} \geq 1$:

$$\begin{aligned}
\liminf_{u\to\infty} \frac{a(u)}{b(u)} &= \liminf_{\delta\to 0} \left\{ \liminf_{u\to\infty} \frac{a(u)}{b(u)} \right\} \\
&= \liminf_{\delta\to 0} \left\{ \liminf_{u\to\infty} \left(\frac{a(u,\delta)}{b(u,\delta)} \underbrace{\frac{a(u)}{a(u,\delta)}}_{\geq 1} \underbrace{\frac{b(u,\delta)}{b(u)}}_{=C_\delta/C} \right) \right\} \\
&\geq \liminf_{\delta\to 0} \left\{ \frac{C_\delta}{C} \cdot \underbrace{\liminf_{u\to\infty} \frac{a(u,\delta)}{b(u,\delta)}}_{=1,\text{ wg. (3.17)}} \right\} \\
&= \liminf_{\delta\to 0} \frac{C_\delta}{C} = 1 \qquad \square
\end{aligned}$$

3.4 Verhalten unter Alternativen

Nun wollen wir das Verhalten der Teststatistik

$$T_n(\alpha,\beta) := \hat{\sigma}_n^{-1} n^{-d/2} \max_{\substack{\underline{\mathbf{0}}\leq\mathbf{k}<\mathbf{m}\leq\underline{\mathbf{n}} \\ \lfloor\alpha n^d\rfloor\leq[\mathbf{m}-\mathbf{k}]\leq\lfloor(1-\beta)n^d\rfloor}} \frac{\left|\sum_{\mathbf{k}<\mathbf{j}\leq\mathbf{m}} \left(X_{\mathbf{j}} - \bar{X}\right)\right|}{\sqrt{\frac{[\mathbf{m}-\mathbf{k}]}{n^d}\left(1-\frac{[\mathbf{m}-\mathbf{k}]}{n^d}\right)}}$$

unter Alternativen betrachten. Offensichtlich kann die Statistik keine Strukturbrüche, die außerhalb des betrachteten Bereichs

$$\lfloor\alpha n^d\rfloor \leq [\mathbf{k}-\mathbf{m}] \leq \lfloor(1-\beta)n^d\rfloor$$

liegen, erkennen. Daher betrachten wir nicht die allgemeine epidemische Alternative H_A, sondern die Einschränkung $H_{\alpha,\beta}$. Zunächst ein Hilfslemma:

Lemma 60. *(a) Es sei $X_n = a_n - Y_n$ $(n \in \mathbb{N})$, wobei $\{Y_n\}_{n\in\mathbb{N}}$ reelle Zufallsvariablen mit $Y_n = \mathcal{O}_P(1)$ und $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen mit $a_n \to \infty$ seien. Dann gilt $X_n \xrightarrow{P} \infty$.*

(b) Für zwei Folgen $\{A_n\}_{n\in\mathbb{N}}$ und $\{B_n\}_{n\in\mathbb{N}}$ reeller Zufallsvariablen mit $A_n \xrightarrow{P} 1$ und $B_n \xrightarrow{P} \infty$ sei $X_n = A_n B_n$. Dann gilt auch $X_n \xrightarrow{P} \infty$.

Beweis. siehe Anhang □

Satz 61. *Es gelte $H_{\alpha,\beta}$ mit $|\delta_n| n^{d/2} \to \infty$. Weiter gelte Annahme 1. Dann ist die Teststatistik konsistent:*

$$T_n(\alpha,\beta) \xrightarrow{P} \infty$$

Beweis. Definiere $\tilde{T}_n(\alpha,\beta) := \hat{\sigma}_n/\sigma T_n(\alpha,\beta)$. Für die Changepoints $\mathbf{k},\mathbf{m} \in \mathbb{Z}^d$ gilt nach Voraussetzung $\underline{\mathbf{1}} \leq \mathbf{k} < \mathbf{m} \leq \underline{\mathbf{n}}$ und $\lfloor\alpha n^d\rfloor \leq [\mathbf{m}-\mathbf{k}] \leq \lfloor(1-\beta)n^d\rfloor$. Dann gilt

$$\frac{[\mathbf{m}-\mathbf{k}]}{n^d} \leq \frac{\lfloor(1-\beta)n^d\rfloor}{n^d} \leq 1-\beta$$

und

$$\frac{[\mathbf{m}-\mathbf{k}]}{n^d} \geq \frac{\lfloor\alpha n^d\rfloor}{n^d} \geq \alpha - \frac{1}{n^d} \geq \alpha - 1$$

und somit:

$$\begin{aligned}
\tilde{T}_n(\alpha,\beta) &\geq \frac{\sigma^{-1}}{\sqrt{\frac{[\mathbf{m}-\mathbf{k}]}{n^d}\left(1-\frac{[\mathbf{m}-\mathbf{k}]}{n^d}\right)}} \frac{1}{n^{d/2}} \left|\sum_{\mathbf{k}<\mathbf{i}\leq\mathbf{m}} (X_{\mathbf{i}} - \bar{X})\right| \\
&\geq \frac{\sigma^{-1}}{\sqrt{(1-\beta)(2-\alpha)}} \frac{1}{n^{d/2}} \left|\sum_{\mathbf{k}<\mathbf{i}\leq\mathbf{m}} (X_{\mathbf{i}} - \bar{X})\right|
\end{aligned}$$

Somit reicht es zu zeigen, dass

$$\frac{1}{n^{d/2}} \left| \sum_{\mathbf{k}<\mathbf{i}\leq\mathbf{m}} (X_{\mathbf{i}} - \bar{X}) \right| \xrightarrow{P} \infty,$$

damit $\tilde{T}_n(\alpha, \beta) \xrightarrow{P} \infty$. Schreibt man $\bar{Y} := \frac{1}{n^d} \sum_{\mathbf{1}\leq\mathbf{k}\leq\mathbf{n}} Y_{\mathbf{k}}$, so gilt:

$$\begin{aligned}
&\frac{1}{n^{d/2}} \sum_{\mathbf{k}<\mathbf{i}\leq\mathbf{m}} (X_{\mathbf{i}} - \bar{X}) \\
=&\frac{1}{n^{d/2}} \sum_{\mathbf{k}<\mathbf{i}\leq\mathbf{m}} \left(Y_{\mathbf{i}} + \mu + \delta_n - \left(\bar{Y} + \mu + \frac{[\mathbf{m}-\mathbf{k}]}{n^d} \delta_n \right) \right) \\
=&\frac{1}{n^{d/2}} \sum_{\mathbf{k}<\mathbf{i}\leq\mathbf{m}} (Y_{\mathbf{i}} - \bar{Y}) + \delta_n \frac{[\mathbf{m}-\mathbf{k}]}{n^{d/2}} \left(1 - \frac{[\mathbf{m}-\mathbf{k}]}{n^d} \right)
\end{aligned}$$

Und somit:

$$\begin{aligned}
&\left| \frac{1}{n^{d/2}} \sum_{\mathbf{k}<\mathbf{i}\leq\mathbf{m}} (X_{\mathbf{i}} - \bar{X}) \right| \\
\geq&|\delta_n| n^{d/2} \underbrace{\frac{[\mathbf{m}-\mathbf{k}]}{n^d}}_{\geq\left(\alpha-\frac{1}{n^d}\right)} \underbrace{\left(1 - \frac{[\mathbf{m}-\mathbf{k}]}{n^d} \right)}_{\geq 1-(1-\beta)=\beta} - \left| \frac{1}{n^{d/2}} \sum_{\mathbf{k}<\mathbf{i}\leq\mathbf{m}} (Y_{\mathbf{i}} - \bar{Y}) \right| \\
\geq& \underbrace{|\delta_n| n^{d/2} \left(\alpha - \frac{1}{n^d} \right) \beta}_{\to\infty} - \left| \frac{1}{n^{d/2}} \sum_{\mathbf{k}<\mathbf{i}\leq\mathbf{m}} (Y_{\mathbf{i}} - \bar{Y}) \right|
\end{aligned}$$

Wie unter H_0 gezeigt, ist

$$\left| \frac{1}{n^{d/2}} \sum_{\mathbf{k}<\mathbf{i}\leq\mathbf{m}} (Y_{\mathbf{i}} - \bar{Y}) \right| = \mathcal{O}_P(1).$$

Lemma 60 (a) liefert nun $\tilde{T}_n(\alpha, \beta) \xrightarrow{P} \infty$. Wie in Abschnitt 1.3 gezeigt, gilt $\hat{\sigma}_n / \sigma \xrightarrow{P} 1$ wegen $\hat{\sigma}_n - \sigma \xrightarrow{P} 0$. Somit folgt die Behauptung aus Lemma 60 (b). $\square$

4 Schätzer

Nun betrachten wir wieder den Fall, dass ein Strukturbruch vorliegt. Es gelte die Alternative:

$$
\begin{aligned}
&H_{\boldsymbol{\vartheta},\boldsymbol{\gamma}} : \text{Es gibt } \underline{\mathbf{0}} < \boldsymbol{\vartheta} < \boldsymbol{\gamma} < \underline{\mathbf{1}}, \text{ so dass}\\
&\qquad X_{\mathbf{j}} = \begin{cases} \mu + \delta_n + Y_{\mathbf{j}} & \mathbf{j} \in R \\ \mu + Y_{\mathbf{j}} & \mathbf{j} \in \{1, \ldots, n\}^d \setminus R \end{cases}\\
&\qquad \text{für } R := (\lfloor n\boldsymbol{\vartheta} \rfloor, \lfloor n\boldsymbol{\gamma} \rfloor] := (\lfloor n\vartheta_1 \rfloor, \lfloor n\gamma_1 \rfloor] \times \cdots \times (\lfloor n\vartheta_d \rfloor, \lfloor n\gamma_d \rfloor],
\end{aligned}
$$

wobei wir für $d \geq 2$ zusätzlich annehmen, dass $\delta_n > 0$ ist. Nun ist es naheliegend, sich die Frage nach der Position des Rechtecks, in dem die Veränderung stattfindet, zu stellen. Dazu wollen wir die Randpunkte von R, bzw. die Parameter $\boldsymbol{\vartheta}$ und $\boldsymbol{\gamma}$, schätzen. Der Beweis der Konsistenz im vorherigen Abschnitt legt nahe, dass die Stellen, an denen die Summe

$$\sum_{\mathbf{k} < \mathbf{j} \leq \mathbf{m}} (X_{\mathbf{j}} - \bar{X})$$

maximiert wird, geeignete Schätzer liefern. Diesen Ansatz wollen wir im Folgenden verfolgen und nachweisen, dass so definierte Schätzer (Definition 63) (schwach) konsistent sind, d. h., dass sie stochastisch gegen $(\boldsymbol{\vartheta}, \boldsymbol{\gamma})$ konvergieren (Satz 64). Wir beschränken uns dabei auf die Fälle $d \in \{1, 2\}$.

Definition 62. *Für eine Funktion $Z : A \to \mathbb{R}$ ($A \subseteq [0,1]^d$, $d \in \mathbb{N}$) in $D[0,1]^d$ und eine kompakte Teilmenge $B \subset A$ sei die Menge*

$$\arg\max_B Z := \{\mathbf{a} \in B : \ Z(\mathbf{a}) = \max_{\mathbf{b} \in B} Z(\mathbf{b})\}.$$

Um die Punkte $\boldsymbol{\vartheta}$ und $\boldsymbol{\gamma}$ zu schätzen, benutzen wir den in Aston und Kirch (2012) (für den eindimensionalen Fall) betrachteten Schätzer, der als Maximalstelle einer Statistik, die der von uns betrachteten Teststatistik ähnelt, gewählt wird:

Definition 63. *Für $d \in \{1, 2\}$ seien*

$$K_d := \{(\mathbf{s}, \mathbf{t}) \in [0,1]^{2d} : \ \underline{\mathbf{0}} \leq \mathbf{s} \leq \mathbf{t} \leq \underline{\mathbf{1}}\}$$

und $G_{n,d} \in D[0,1]^{2d}$ mit

$$G_{n,d}(\mathbf{s}, \mathbf{t}) := \begin{cases} \frac{1}{n} \left| \sum\limits_{\lfloor ns \rfloor < i \leq \lfloor nt \rfloor} (X_i - \bar{X}) \right| I_{K_1}(s,t), & d = 1 \\ \frac{1}{n^d} \sum\limits_{\lfloor n\mathbf{s} \rfloor < \mathbf{i} \leq \lfloor n\mathbf{t} \rfloor} (X_{\mathbf{i}} - \bar{X}) I_{K_2}(\mathbf{s}, \mathbf{t}), & d = 2. \end{cases}$$

Die Schätzer $(\hat{\boldsymbol{\vartheta}}_n, \hat{\boldsymbol{\gamma}}_n)$ *seien beliebige Punkte aus* $\arg\max_{K_d} G_{n,d}$.

Ziel ist nun der Beweis des folgenden Satzes:
Satz 64. *Es gelte* $H_{\boldsymbol{\vartheta},\boldsymbol{\gamma}}$ *mit* $\delta_n = \delta$ *konstant. Weiter sei Annahme 1 erfüllt. Dann gilt für* $d \in \{1, 2\}$ *und* $n \to \infty$*:*

$$(\hat{\boldsymbol{\vartheta}}_n - \boldsymbol{\vartheta}, \hat{\boldsymbol{\gamma}}_n - \boldsymbol{\gamma}) = o_P(1)$$

Wie in den vorherigen Abschnitten beweisen wir zunächst einige Hilfslemmata, deren Anwendung den Beweis des Satzes 64 liefert. Die Idee des Beweises ist, zunächst zu zeigen, dass sich die Statistiken, deren Maximalstellen betrachtet werden, asymptotisch an eine Funktion annähern, die die gewünschten Parameter als eindeutige Maximalstelle hat und daraus die Konvergenz der Maximalstellen zu folgern. Den ersten Schritt dazu liefert folgendes Lemma:
Lemma 65. *Unter den Voraussetzungen von Satz 64 gilt für* $d = 1, 2$*:*

$$\sup_{\underline{\mathbf{0}} \le \mathbf{x} \le \mathbf{y} \le \underline{\mathbf{1}}} \left| \frac{1}{n^d} \sum_{\lfloor n\mathbf{x} \rfloor < \mathbf{i} \le \lfloor n\mathbf{y} \rfloor} (X_{\mathbf{i}} - \bar{X}) - \delta f(\mathbf{x}, \mathbf{y}) \right| = o_P(1),$$

wobei $f : [0, 1]^{2d} \to \mathbb{R}$ *wie folgt definiert ist:*

- $d = 1$*:*
$$f(x, y) := (g(y) - g(x)) I_{K_1}(x, y)$$
mit
$$g(a) := \begin{cases} -a(\gamma - \vartheta), & 0 \le a \le \vartheta \\ -(\vartheta - a(1 - (\gamma - \vartheta))), & \vartheta < a \le \gamma \\ (1 - a)(\gamma - \vartheta), & \gamma < a \le 1 \end{cases} \tag{4.1}$$

- $d = 2$*:*
$$f(\mathbf{x}, \mathbf{y}) := (g(\mathbf{y}) - g(x_1, y_2) - g(y_1, x_2) + g(\mathbf{x}) - [\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}]) I_{K_2}(\mathbf{x}, \mathbf{y})$$
mit
$$g(a, b) := \begin{cases} 0, & 0 \le a \le \vartheta_1 \text{ oder } 0 \le b \le \vartheta_2 \\ (a - \vartheta_1)(b - \vartheta_2), & \vartheta_1 < a \le \gamma_1 \text{ und } \vartheta_2 < b \le \gamma_2 \\ (a - \vartheta_1)(\gamma_2 - \vartheta_2), & \vartheta_1 < a \le \gamma_1 \text{ und } \gamma_2 < b \le 1 \\ (\gamma_1 - \vartheta_1)(b - \vartheta_2), & \gamma_1 < a \le 1 \text{ und } \vartheta_2 < b \le \gamma_2 \\ [\boldsymbol{\gamma} - \boldsymbol{\vartheta}], & \gamma_1 < a \le 1 \text{ und } \gamma_2 < b \le 1 \end{cases} \tag{4.2}$$

Beweis. Sei $\bar{Y} := \frac{1}{n^d} \sum_{\underline{\mathbf{1}} \le \mathbf{k} \le \underline{\mathbf{n}}} Y_{\mathbf{k}}$. Betrachte zunächst für beliebiges $d \in \mathbb{N}$ und

$\underline{\mathbf{0}} \leq \mathbf{x} \leq \mathbf{y} \leq \underline{\mathbf{1}}$:

$$
\begin{aligned}
&\frac{1}{n^d} \sum_{\lfloor n\mathbf{x}\rfloor < \mathbf{i} \leq \lfloor n\mathbf{y}\rfloor} (X_{\mathbf{i}} - \bar{X}) \\
=&\frac{1}{n^d} \sum_{\lfloor n\mathbf{x}\rfloor < \mathbf{i} \leq \lfloor n\mathbf{y}\rfloor} \left(\mu + Y_{\mathbf{i}} + \delta I_{\{\lfloor n\boldsymbol{\vartheta}\rfloor < \mathbf{i} \leq \lfloor n\boldsymbol{\gamma}\rfloor\}}(\mathbf{i}) - \left(\mu + \bar{Y} + \frac{[\lfloor n\boldsymbol{\gamma}\rfloor - \lfloor n\boldsymbol{\vartheta}\rfloor]}{n^d}\delta\right)\right) \\
=&\frac{1}{n^d} \sum_{\lfloor n\mathbf{x}\rfloor < \mathbf{i} \leq \lfloor n\mathbf{y}\rfloor} (Y_{\mathbf{i}} - \bar{Y}) \\
&+ \delta \underbrace{\left(\frac{1}{n^d} \sum_{\lfloor n\mathbf{x}\rfloor < \mathbf{i} \leq \lfloor n\mathbf{y}\rfloor} I_{\{\lfloor n\boldsymbol{\vartheta}\rfloor < \mathbf{i} \leq \lfloor n\boldsymbol{\gamma}\rfloor\}}(\mathbf{i}) - \frac{[\lfloor n\boldsymbol{\gamma}\rfloor - \lfloor n\boldsymbol{\vartheta}\rfloor]}{n^d} \frac{[\lfloor n\mathbf{y}\rfloor - \lfloor n\mathbf{x}\rfloor]}{n^d}\right)}_{=: f_n(\mathbf{x},\mathbf{y})}
\end{aligned}
$$

Folglich gilt:

$$
\begin{aligned}
&\sup_{\underline{\mathbf{0}} \leq \mathbf{x} \leq \mathbf{y} \leq \underline{\mathbf{1}}} \left| \frac{1}{n^d} \sum_{\lfloor n\mathbf{x}\rfloor < \mathbf{i} \leq \lfloor n\mathbf{y}\rfloor} (X_{\mathbf{i}} - \bar{X}) - \delta f(\mathbf{x},\mathbf{y}) \right| \\
\leq&\frac{1}{n^{d/2}} \sup_{\underline{\mathbf{0}} \leq \mathbf{x} \leq \mathbf{y} \leq \underline{\mathbf{1}}} \left| \frac{1}{n^{d/2}} \sum_{\lfloor n\mathbf{x}\rfloor < \mathbf{i} \leq \lfloor n\mathbf{y}\rfloor} (Y_{\mathbf{i}} - \bar{Y}) \right| + |\delta| \sup_{\underline{\mathbf{0}} \leq \mathbf{x} \leq \mathbf{y} \leq \underline{\mathbf{1}}} |f_n(\mathbf{x},\mathbf{y}) - f(\mathbf{x},\mathbf{y})|
\end{aligned}
$$

Wie unter H_0 gesehen, gilt

$$
\sup_{\underline{\mathbf{0}} \leq \mathbf{x} \leq \mathbf{y} \leq \underline{\mathbf{1}}} \left| \frac{1}{\sigma n^{d/2}} \sum_{\lfloor n\mathbf{x}\rfloor < \mathbf{i} \leq \lfloor n\mathbf{y}\rfloor} (Y_{\mathbf{i}} - \bar{Y}) \right| = \mathcal{O}_P(1),
$$

also

$$
\frac{\sigma}{n^{d/2}} \sup_{\underline{\mathbf{0}} \leq \mathbf{x} \leq \mathbf{y} \leq \underline{\mathbf{1}}} \left| \frac{1}{\sigma n^{d/2}} \sum_{\lfloor n\mathbf{x}\rfloor < \mathbf{i} \leq \lfloor n\mathbf{y}\rfloor} (Y_{\mathbf{i}} - \bar{Y}) \right| = o_P(1).
$$

Zu zeigen bleibt also nur:

$$
\sup_{\underline{\mathbf{0}} \leq \mathbf{x} \leq \mathbf{y} \leq \underline{\mathbf{1}}} |f_n(\mathbf{x},\mathbf{y}) - f(\mathbf{x},\mathbf{y})| \stackrel{!}{=} o(1)
$$

Wir schreiben $R := \{\mathbf{i} \in \mathbb{Z}^d : \ \lfloor n\boldsymbol{\vartheta}\rfloor < \mathbf{i} \leq \lfloor n\boldsymbol{\gamma}\rfloor\}$. Zunächst für $d = 1$:

$$
\begin{aligned}
&f_n(x, y) \\
=&\frac{1}{n} \sum_{i=\lfloor nx\rfloor+1}^{\lfloor ny\rfloor} I_R(i) - \frac{\lfloor ny\rfloor - \lfloor nx\rfloor}{n} \frac{\lfloor n\gamma\rfloor - \lfloor n\vartheta\rfloor}{n} \\
=&\left(\frac{1}{n} \sum_{i=1}^{\lfloor ny\rfloor} I_R(i) - \frac{\lfloor ny\rfloor}{n} \frac{\lfloor n\gamma\rfloor - \lfloor n\vartheta\rfloor}{n}\right) - \left(\frac{1}{n} \sum_{i=1}^{\lfloor nx\rfloor} I_R(i) - \frac{\lfloor nx\rfloor}{n} \frac{\lfloor n\gamma\rfloor - \lfloor n\vartheta\rfloor}{n}\right) \\
=:&g_n(y) - g_n(x)
\end{aligned}
$$

Durch Einsetzen der Definition von f folgt nun:

$$
\begin{aligned}
&\sup_{0\le x\le y\le 1} |f_n(x,y) - f(x,y)| \\
\le& 2 \sup_{0\le x\le 1} |g_n(x) - g(x)| \\
\le& 2 \left\{ \sup_{0\le x\le \vartheta} |g_n(x) + x(\gamma-\vartheta)| + \sup_{\vartheta< x\le \gamma} |g_n(x) + \vartheta - x(1-(\gamma-\vartheta))| \right. \\
&\left. + \sup_{\gamma< x\le 1} |g_n(x) - (1-x)(\gamma-\vartheta)| \right\} \\
=:& 2\{K_1^{(n)} + K_2^{(n)} + K_3^{(n)}\}
\end{aligned}
$$

Um zu zeigen, dass $K_i^{(n)} = o(1)$ $(i=1,2,3)$, betrachten wir $g_n(x)$ genauer. Es gilt:

$$
\begin{aligned}
g_n(x) &= \begin{cases} 0 - \frac{\lfloor nx\rfloor}{n}\frac{\lfloor n\gamma\rfloor - \lfloor n\vartheta\rfloor}{n}, & 0\le \lfloor nx\rfloor \le \lfloor n\vartheta\rfloor \\ \frac{\lfloor nx\rfloor - \lfloor n\vartheta\rfloor}{n} - \frac{\lfloor nx\rfloor}{n}\frac{\lfloor n\gamma\rfloor - \lfloor n\vartheta\rfloor}{n}, & \lfloor n\vartheta\rfloor < \lfloor nx\rfloor \le \lfloor n\gamma\rfloor \\ \frac{\lfloor n\gamma\rfloor - \lfloor n\vartheta\rfloor}{n} - \frac{\lfloor nx\rfloor}{n}\frac{\lfloor n\gamma\rfloor - \lfloor n\vartheta\rfloor}{n}, & \lfloor n\gamma\rfloor < \lfloor nx\rfloor \le n \end{cases} \\
&= \begin{cases} -\frac{\lfloor nx\rfloor}{n}\frac{\lfloor n\gamma\rfloor - \lfloor n\vartheta\rfloor}{n}, & 0\le \lfloor nx\rfloor \le \lfloor n\vartheta\rfloor \\ -\left(\frac{\lfloor n\vartheta\rfloor}{n} - \frac{\lfloor nx\rfloor}{n}\left(1-\frac{\lfloor n\gamma\rfloor - \lfloor n\vartheta\rfloor}{n}\right)\right), & \lfloor n\vartheta\rfloor < \lfloor nx\rfloor \le \lfloor n\gamma\rfloor \\ \left(1-\frac{\lfloor nx\rfloor}{n}\right)\frac{\lfloor n\gamma\rfloor - \lfloor n\vartheta\rfloor}{n}, & \lfloor n\gamma\rfloor < \lfloor nx\rfloor \le n \end{cases}
\end{aligned}
$$

Wie man sieht, unterscheiden sich g_n und g nur darin, dass g_n abhängig von $\frac{\lfloor n\cdot\rfloor}{n}$ ist. Die behauptete Konvergenz ergibt sich nun dadurch, dass $\frac{\lfloor nt\rfloor}{n}$ gleichmäßig gegen t konvergiert. Da die Rechnungen für $K_2^{(n)}$ und $K_3^{(n)}$ analog verlaufen, zeigen wir hier nur $K_1^{(n)} = o(1)$. Zunächst stellen wir fest, dass es für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt, so dass für alle $n \ge n_0$ gilt:

$$\sup_{0\le x\le 1}\left|x - \frac{\lfloor nx\rfloor}{n}\right| \le \frac{1}{n} \le \varepsilon$$

Da aus $0 \leq x \leq \vartheta$ folgt, dass $0 \leq \lfloor nx \rfloor \leq \lfloor n\vartheta \rfloor$, gilt somit für $n \geq n_0$:

$$
\begin{aligned}
K_1^{(n)} &\leq \sup_{0 \leq \lfloor nx \rfloor \leq \lfloor n\vartheta \rfloor} |g_n(x) + x(\gamma - \vartheta)| \\
&= \sup_{0 \leq \lfloor nx \rfloor \leq \lfloor n\vartheta \rfloor} \left| -\frac{\lfloor nx \rfloor}{n} \frac{\lfloor n\gamma \rfloor - \lfloor n\vartheta \rfloor}{n} + x(\gamma - \vartheta) \right| \\
&\leq \sup_{0 \leq \lfloor nx \rfloor \leq \lfloor n\vartheta \rfloor} \left| \overbrace{\frac{\lfloor nx \rfloor}{n}}^{\leq 1} \left(\frac{\lfloor n\gamma \rfloor - \lfloor n\vartheta \rfloor}{n} - (\gamma - \vartheta) \right) \right| \\
&+ \sup_{0 \leq \lfloor nx \rfloor \leq \lfloor n\vartheta \rfloor} \left| \underbrace{(\gamma - \vartheta)}_{\leq 1} \left(x - \frac{\lfloor nx \rfloor}{n} \right) \right| \\
&\leq 3 \sup_{0 \leq x \leq 1} \left| x - \frac{\lfloor nx \rfloor}{n} \right| \\
&\leq 3\varepsilon
\end{aligned}
$$

Folglich ist $K_1^{(n)} = o(1)$. Analog lässt sich zeigen, dass $K_2^{(n)} = o(1) = K_3^{(n)}$.

Nun zum Fall $d = 2$: Analog zu $d = 1$ bringen wir zunächst die Funktion f_n in eine der Funktion f ähnliche Form und nutzen dann die Konvergenz $\lfloor nx \rfloor / n \to x$.

$$
\begin{aligned}
f_n(\mathbf{x}, \mathbf{y}) =& \frac{1}{n^2} \sum_{i=\lfloor nx_1 \rfloor + 1}^{\lfloor ny_1 \rfloor} \sum_{j=\lfloor nx_2 \rfloor + 1}^{\lfloor ny_2 \rfloor} I_R(i,j) - \frac{[\lfloor n\mathbf{y} \rfloor - \lfloor n\mathbf{x} \rfloor]}{n^2} \frac{[\lfloor n\boldsymbol{\gamma} \rfloor - \lfloor n\boldsymbol{\vartheta} \rfloor]}{n^2} \\
=& \frac{1}{n^2} \sum_{\mathbf{1} \leq \mathbf{j} \leq \lfloor n\mathbf{y} \rfloor} I_R(\mathbf{j}) - \frac{1}{n^2} \sum_{i=1}^{\lfloor nx_1 \rfloor} \sum_{j=1}^{\lfloor ny_2 \rfloor} I_R(i,j) \\
&- \frac{1}{n^2} \sum_{i=1}^{\lfloor ny_1 \rfloor} \sum_{j=1}^{\lfloor nx_2 \rfloor} I_R(i,j) + \frac{1}{n^2} \sum_{\mathbf{1} \leq \mathbf{j} \leq \lfloor n\mathbf{x} \rfloor} I_R(\mathbf{j}) \\
&- \frac{[\lfloor n\mathbf{y} \rfloor - \lfloor n\mathbf{x} \rfloor]}{n^2} \frac{[\lfloor n\boldsymbol{\gamma} \rfloor - \lfloor n\boldsymbol{\vartheta} \rfloor]}{n^2} \\
=:& g_n(\mathbf{y}) - g_n(x_1, y_2) - g_n(y_1, x_2) + g_n(\mathbf{x}) - \frac{[\lfloor n\mathbf{y} \rfloor - \lfloor n\mathbf{x} \rfloor]}{n^2} \frac{[\lfloor n\boldsymbol{\gamma} \rfloor - \lfloor n\boldsymbol{\vartheta} \rfloor]}{n^2}
\end{aligned}
$$

Somit gilt:

$$
\begin{aligned}
&\sup_{\mathbf{0} \leq \mathbf{x} \leq \mathbf{y} \leq \mathbf{1}} |f_n(\mathbf{x}, \mathbf{y}) - f(\mathbf{x}, \mathbf{y})| \\
\leq& 4 \sup_{\mathbf{0} \leq \mathbf{x} \leq \mathbf{1}} |g_n(\mathbf{x}) - g(\mathbf{x})| \\
&+ \sup_{\mathbf{0} \leq \mathbf{x} \leq \mathbf{y} \leq \mathbf{1}} \left| \frac{[\lfloor n\mathbf{y} \rfloor - \lfloor n\mathbf{x} \rfloor]}{n^2} \frac{[\lfloor n\boldsymbol{\gamma} \rfloor - \lfloor n\boldsymbol{\vartheta} \rfloor]}{n^2} - [\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}] \right|
\end{aligned}
$$

Zunächst betrachten wir den letzten Summanden:

$$
\begin{aligned}
&\sup_{\underline{\mathbf{0}}\leq\mathbf{x}\leq\mathbf{y}\leq\underline{\mathbf{1}}}\left|\frac{[\lfloor n\mathbf{y}\rfloor-\lfloor n\mathbf{x}\rfloor]}{n^2}\frac{[\lfloor n\boldsymbol{\gamma}\rfloor-\lfloor n\boldsymbol{\vartheta}\rfloor]}{n^2}-[\mathbf{y}-\mathbf{x}][\boldsymbol{\gamma}-\boldsymbol{\vartheta}]\right|\\
\leq&\sup_{\underline{\mathbf{0}}\leq\mathbf{x}\leq\mathbf{y}\leq\underline{\mathbf{1}}}\left|\frac{[\lfloor n\mathbf{y}\rfloor-\lfloor n\mathbf{x}\rfloor]}{n^2}\left(\frac{[\lfloor n\boldsymbol{\gamma}\rfloor-\lfloor n\boldsymbol{\vartheta}\rfloor]}{n^2}-[\boldsymbol{\gamma}-\boldsymbol{\vartheta}]\right)\right|\\
&+\sup_{\underline{\mathbf{0}}\leq\mathbf{x}\leq\mathbf{y}\leq\underline{\mathbf{1}}}\left|\left(\frac{[\lfloor n\mathbf{y}\rfloor-\lfloor n\mathbf{x}\rfloor]}{n^2}-[\mathbf{y}-\mathbf{x}]\right)\underbrace{[\boldsymbol{\gamma}-\boldsymbol{\vartheta}]}_{\leq 1}\right|\\
\leq&\left|\frac{[\lfloor n\boldsymbol{\gamma}\rfloor-\lfloor n\boldsymbol{\vartheta}\rfloor]}{n^2}-[\boldsymbol{\gamma}-\boldsymbol{\vartheta}]\right|+\sup_{\underline{\mathbf{0}}\leq\mathbf{x}\leq\mathbf{y}\leq\underline{\mathbf{1}}}\left|\frac{[\lfloor n\mathbf{y}\rfloor-\lfloor n\mathbf{x}\rfloor]}{n^2}-[\mathbf{y}-\mathbf{x}]\right|\\
\leq&2\sup_{\underline{\mathbf{0}}\leq\mathbf{x}\leq\mathbf{y}\leq\underline{\mathbf{1}}}\left|\frac{[\lfloor n\mathbf{y}\rfloor-\lfloor n\mathbf{x}\rfloor]}{n^2}-[\mathbf{y}-\mathbf{x}]\right|\\
=&o(1),
\end{aligned}
$$

wobei im letzten Schritt Lemma 41 verwendet wurde. Nun betrachten wir g_n genauer:

$$
\begin{aligned}
&g_n(a,b)\\
=&\frac{1}{n^2}\sum_{i=1}^{\lfloor na\rfloor}\sum_{j=1}^{\lfloor nb\rfloor}I_R(i,j)\\
=&\begin{cases}
0, & \lfloor na\rfloor\leq\lfloor n\vartheta_1\rfloor \text{ oder } \lfloor nb\rfloor\leq\lfloor n\vartheta_2\rfloor\\
\frac{\lfloor na\rfloor-\lfloor n\vartheta_1\rfloor}{n}\frac{\lfloor nb\rfloor-\lfloor n\vartheta_2\rfloor}{n}, & \lfloor n\vartheta_1\rfloor<\lfloor na\rfloor\leq\lfloor n\gamma_1\rfloor,\ \lfloor n\vartheta_2\rfloor<\lfloor nb\rfloor\leq\lfloor n\gamma_2\rfloor\\
\frac{\lfloor na\rfloor-\lfloor n\vartheta_1\rfloor}{n}\frac{\lfloor n\gamma_2\rfloor-\lfloor n\vartheta_2\rfloor}{n}, & \lfloor n\vartheta_1\rfloor<\lfloor na\rfloor\leq\lfloor n\gamma_1\rfloor,\ \lfloor n\gamma_2\rfloor<\lfloor nb\rfloor\\
\frac{\lfloor n\gamma_1\rfloor-\lfloor n\vartheta_1\rfloor}{n}\frac{\lfloor nb\rfloor-\lfloor n\vartheta_2\rfloor}{n}, & \lfloor n\gamma_1\rfloor<\lfloor na\rfloor,\ \lfloor n\vartheta_2\rfloor<\lfloor nb\rfloor\leq\lfloor n\gamma_2\rfloor\\
\frac{[\lfloor n\boldsymbol{\gamma}\rfloor-\lfloor n\boldsymbol{\vartheta}\rfloor]}{n^2}, & \lfloor n\gamma_1\rfloor<\lfloor na\rfloor,\ \lfloor n\gamma_2\rfloor<\lfloor nb\rfloor
\end{cases}
\end{aligned}
$$

Analog zum Fall $d=1$ teilen wir den Bereich, über den das Supremum gebildet wird, passend zur Definition von g auf:

$$
\begin{aligned}
&\sup_{\underline{\mathbf{0}}\leq\mathbf{x}\leq\underline{\mathbf{1}}}|g_n(\mathbf{x})-g(\mathbf{x})|\\
\leq&\sup_{0\leq x_1\leq\vartheta_1\vee 0\leq x_2\leq\vartheta_2}|g_n(\mathbf{x})|+\sup_{\boldsymbol{\vartheta}<\mathbf{x}\leq\boldsymbol{\gamma}}|g_n(\mathbf{x})-[\mathbf{x}-\boldsymbol{\vartheta}]|\\
&+\sup_{\vartheta_1<x_1\leq\gamma_1,\gamma_2<x_2\leq 1}|g_n(\mathbf{x})-(x_1-\vartheta_1)(\gamma_2-\vartheta_2)|\\
&+\sup_{\gamma_1<x_1\leq 1,\vartheta_2<x_2\leq\gamma_2}|g_n(\mathbf{x})-(\gamma_1-\vartheta_1)(x_2-\vartheta_2)|\\
&+\sup_{\boldsymbol{\gamma}<\mathbf{x}\leq\underline{\mathbf{1}}}|g_n(\mathbf{x})-[\boldsymbol{\gamma}-\boldsymbol{\vartheta}]|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \sup_{0\leq\lfloor nx_1\rfloor\leq\lfloor n\vartheta_1\rfloor\vee 0\leq\lfloor nx_2\rfloor\leq\lfloor n\vartheta_2\rfloor} |g_n(\mathbf{x})| + \sup_{\lfloor n\boldsymbol{\vartheta}\rfloor\leq\lfloor n\mathbf{x}\rfloor\leq\lfloor n\boldsymbol{\gamma}\rfloor} |g_n(\mathbf{x}) - [\mathbf{x}-\boldsymbol{\vartheta}]| \\
&\quad + \sup_{\lfloor n\vartheta_1\rfloor\leq\lfloor nx_1\rfloor\leq\lfloor n\gamma_1\rfloor,\lfloor n\gamma_2\rfloor\leq\lfloor nx_2\rfloor\leq n} |g_n(\mathbf{x}) - (x_1-\vartheta_1)(\gamma_2-\vartheta_2)| \\
&\quad + \sup_{\lfloor n\gamma_1\rfloor\leq\lfloor nx_1\rfloor\leq n,\lfloor n\vartheta_2\rfloor\leq\lfloor nx_2\rfloor\leq\lfloor n\gamma_2\rfloor} |g_n(\mathbf{x}) - (\gamma_1-\vartheta_1)(x_2-\vartheta_2)| \\
&\quad + \sup_{\lfloor n\boldsymbol{\gamma}\rfloor\leq\lfloor n\mathbf{x}\rfloor\leq\underline{\mathbf{n}}} |g_n(\mathbf{x}) - [\boldsymbol{\gamma}-\boldsymbol{\vartheta}]| \\
&=: K_1^{(n)} + \cdots + K_5^{(n)}
\end{aligned}
$$

Nun kann man wieder die ähnliche Form von g_n, und die Konvergenz $\lfloor nt\rfloor/n \to t$ nutzen, um zu zeigen, dass $K_i^{(n)} = o(1)$ $(i = 1,\ldots,5)$. Da dabei jeweils sehr ähnliche Überlegungen einfließen, wie bisher, verzichten wir auf einen Beweis. $\square$

Um Satz 64 beweisen zu können, müssen wir zeigen, dass aus Lemma 65 die Konvergenz der Maximalstellen gegen die Maximalstelle der Grenzfunktion folgt. Dies zeigt folgendes Lemma, für den Fall, dass die Grenzfunktion stetig mit eindeutiger Maximalstelle ist. Es wurde in leicht veränderter Form (als Verallgemeinerung auf den d-dimensionalen Fall) aus Kirch (2008) (Lemma 2.15, S. 25) übernommen:

Lemma 66. *Seien $K \subset \mathbb{R}^d$ $(d \in \mathbb{N})$ eine kompakte Menge und $f : K \to \mathbb{R}$ eine stetige Funktion mit eindeutiger Maximalstelle $\mathbf{x}_0 \in K$ (d. h. $\{\mathbf{x}_0\} = \arg\max_K f$). Weiter seien $f_n : K \to \mathbb{R}$ Funktionen mit*

$$
\max_{\mathbf{x}\in K} |f_n(\mathbf{x}) - f(\mathbf{x})| \to 0 \tag{4.3}
$$

und (nicht unbedingt eindeutigen) Maximalstellen $\hat{\mathbf{x}}_n$ (d. h. $f_n(\hat{\mathbf{x}}_n) = \max_{\mathbf{x}\in K} f_n(x)$). Dann gilt:

$$
\hat{\mathbf{x}}_n \overset{n\to\infty}{\longrightarrow} \mathbf{x}_0
$$

Beweis. Wir führen einen Widerspruchsbeweis durch. Angenommen, $\hat{\mathbf{x}}_n \not\to \mathbf{x}_0$. Da K kompakt ist, existiert somit ein $\mathbf{x}_1 \in K$ mit $\mathbf{x}_1 \neq \mathbf{x}_0$ und eine Teilfolge $(\hat{\mathbf{x}}_{\alpha(n)})_{n\in\mathbb{N}} \subset K$ mit $\hat{\mathbf{x}}_{\alpha(n)} \to \mathbf{x}_1$. Dann ist auch

$$
\lim_{n\to\infty} |f_{\alpha(n)}(\hat{\mathbf{x}}_{\alpha(n)}) - f(\mathbf{x}_1)| = 0,
$$

denn es gilt:

$$
\begin{aligned}
&|f_{\alpha(n)}(\hat{\mathbf{x}}_{\alpha(n)}) - f(\mathbf{x}_1)| \\
&\leq \underbrace{|f_{\alpha(n)}(\hat{\mathbf{x}}_{\alpha(n)}) - f(\hat{\mathbf{x}}_{\alpha(n)})|}_{\leq \max_{\mathbf{x}\in K}|f_n(\mathbf{x})-f(\mathbf{x})|\to 0,\text{ wg. (4.3)}} + \underbrace{|f(\hat{\mathbf{x}}_{\alpha(n)}) - f(\mathbf{x}_1)|}_{\to 0,\text{ da } f \text{ stetig}} \to 0
\end{aligned}
$$

Da die Maximalstelle von f eindeutig ist, gilt $f(\mathbf{x}_0) > f(\mathbf{x}_1)$. Mit der Dreiecksungleichung folgt:

$$
\begin{aligned}
&\limsup_{n\to\infty} |f_{\alpha(n)}(\hat{\mathbf{x}}_{\alpha(n)}) - f(\mathbf{x}_0)| \\
\geq &\limsup_{n\to\infty} \left\{ |f(\mathbf{x}_1) - f(\mathbf{x}_0)| - |f_{\alpha(n)}(\hat{\mathbf{x}}_{\alpha(n)}) - f(\mathbf{x}_1)| \right\} \\
= &|f(\mathbf{x}_1) - f(\mathbf{x}_0)| - \underbrace{\liminf_{n\to\infty} |f_{\alpha(n)}(\hat{\mathbf{x}}_{\alpha(n)}) - f(\mathbf{x}_1)|}_{=\lim\limits_{n\to\infty} |f_{\alpha(n)}(\hat{\mathbf{x}}_{\alpha(n)}) - f(\mathbf{x}_1)| = 0} \\
= &f(\mathbf{x}_0) - f(\mathbf{x}_1) > 0
\end{aligned}
$$

Hieraus folgt

$$|f_n(\hat{\mathbf{x}}_n) - f(\mathbf{x}_0)| \not\to 0.$$

Aber aus Annahme (4.3) folgt auch

$$
\begin{aligned}
|f_n(\hat{\mathbf{x}}_n) - f(\mathbf{x}_0)| &= |\max_{\mathbf{x}\in K} f_n(\mathbf{x}) - \max_{\mathbf{x}\in K} f(\mathbf{x})| \\
&\leq \max_{\mathbf{x}\in K} |f_n(\mathbf{x}) - f(\mathbf{x})| \to 0,
\end{aligned}
$$

also ein Widerspruch. Somit gilt die Behauptung. □

Lemma 67. *Die Funktionen $f(\mathbf{x}, \mathbf{y})$ aus Lemma 65 sind stetig und haben die eindeutige Maximalstelle $(\boldsymbol{\vartheta}, \boldsymbol{\gamma})$. Für $d = 1$ hat auch $|f(x, y)|$ die eindeutige Maximalstelle (ϑ, γ).*

Beweis. Es sei K_d wie in Definition 63. Zunächst zeigen wir die Stetigkeit: Sei $d \in \{1, 2\}$. Die Funktion g ist nach Lemma 12 stetig auf $[0, 1]^d$ und f hat die Form (3.1). Somit liefert Lemma 42 die Stetigkeit von f. Nun zeigen wir zunächst die Eindeutigkeit für $d = 1$. Als erstes stellen wir fest, dass wegen $0 < \gamma - \vartheta < 1$ gilt:

$$|f(\vartheta, \gamma)| = f(\vartheta, \gamma) = g(\gamma) - g(\vartheta) = (1 - (\gamma - \vartheta))(\gamma - \vartheta) > 0$$

Folglich reicht es, die Maximalstelle für $(x, y) \in K_1$ zu bestimmen, da f außerhalb von K_1 gleich Null ist. Für $(x, y) \in K_1$ gilt:

$$
\begin{aligned}
|f(x, y)| &= |g(y) - g(x)| \\
&= \max\{ \overbrace{g(y)}^{\leq \max\limits_{z\in[0,1]} g(z)} - \overbrace{g(x)}^{\geq \min\limits_{z\in[0,1]} g(z)}, \overbrace{g(x)}^{\leq \max\limits_{z\in[0,1]} g(z)} - \overbrace{g(y)}^{\geq \min\limits_{z\in[0,1]} g(z)} \} \\
&\leq \max_{z\in[0,1]} g(z) - \min_{z\in[0,1]} g(z)
\end{aligned}
$$

$g(a)$ ist für $0 \leq a \leq \vartheta$ monoton fallend mit Minimum bei

$$g(\vartheta) = -\vartheta(\gamma - \vartheta) < 0,$$

für $\vartheta \leq a \leq \gamma$ monoton wachsend mit Maximum

$$g(\gamma) = (1 - \gamma)(\gamma - \vartheta)$$

und für $\gamma \leq a \leq 1$ monoton fallend mit Minimum $g(1) = 0 > g(\vartheta)$ und Maximum $g(\gamma)$. Also ist γ die eindeutige Maximalstelle und ϑ die eindeutige Minimalstelle von g auf $[0,1]$. Folglich ist (ϑ, γ) die eindeutige Maximalstelle von $|f(x,y)|$ auf K_1. Außerhalb von K_1 ist $f(x,y) = 0$, also ist (ϑ, γ) auch auf ganz $[0,1]^2$ eindeutige Maximalstelle. Da $f(\gamma, \vartheta) = |f(\gamma, \vartheta)|$ und $f(x,y) \leq |f(x,y)|$ hat auch $f(x,y)$ bei (ϑ, γ) seine Maximalstelle.
Für $d = 2$ stellen wir zunächst fest, dass wegen $0 < [\boldsymbol{\gamma} - \boldsymbol{\vartheta}] < 1$ gilt:

$$\begin{aligned} f(\boldsymbol{\vartheta}, \boldsymbol{\gamma}) &= g(\boldsymbol{\gamma}) - g(\vartheta_1, \gamma_2) - g(\gamma_1, \vartheta_2) + g(\boldsymbol{\vartheta}) - [\boldsymbol{\gamma} - \boldsymbol{\vartheta}]^2 \\ &= [\boldsymbol{\gamma} - \boldsymbol{\vartheta}] - 0 - 0 + 0 - [\boldsymbol{\gamma} - \boldsymbol{\vartheta}]^2 \\ &= (1 - [\boldsymbol{\gamma} - \boldsymbol{\vartheta}])[\boldsymbol{\gamma} - \boldsymbol{\vartheta}] \\ &> 0 \end{aligned}$$

Somit reicht es auch hier, $(\mathbf{x}, \mathbf{y}) \in K_2$ zu betrachten. Außerdem ist $f(\mathbf{x}, \mathbf{x}) = 0$, so dass wir o. E. $\mathbf{x} < \mathbf{y}$ betrachten können. Da (x_1, y_1) und (x_2, y_2) unabhängig von einander gewählt werden können, können für $\underline{\mathbf{0}} \leq \mathbf{x} < \mathbf{y} \leq \underline{\mathbf{1}}$ folgende Fälle auftreten:

(1) $0 \leq x_1 < y_1 \leq \vartheta_1$

(2) $0 \leq x_1 \leq \vartheta_1 < y_1 \leq \gamma_1$

(3) $0 \leq x_1 \leq \vartheta_1 < \gamma_1 \leq y_1 \leq 1$

(4) $\vartheta_1 \leq x_1 < y_1 \leq \gamma_1$

(5) $\vartheta_1 \leq x_1 \leq \gamma_1 \leq y_1 \leq 1$

(6) $\gamma_1 \leq x_1 \leq y_1 \leq 1$

in Kombination mit

(1') $0 \leq x_2 < y_2 \leq \vartheta_2$

(2') $0 \leq x_2 \leq \vartheta_2 < y_2 \leq \gamma_2$

(3') $0 \leq x_2 \leq \vartheta_2 < \gamma_2 \leq y_2 \leq 1$

(4') $\vartheta_2 \leq x_2 < y_2 \leq \gamma_2$

(5') $\vartheta_2 \leq x_2 \leq \gamma_2 \leq y_2 \leq 1$

(6') $\gamma_2 \leq x_2 \leq y_2 \leq 1$

Nun wollen wir die möglichen Kombinationen durchgehen und zeigen, dass f dort entweder nicht positiv ist, oder das Maximum bei $(\boldsymbol{\vartheta}, \boldsymbol{\gamma})$ erreicht wird. Aus Symmetriegründen lassen sich dabei Kombinationen der Form $(i) - (j')$ genauso behandeln wie $(j) - (i')$. Daher reicht es, folgende Kombinationen zu betrachten:

- (1) mit $(1'), (2'), \ldots, (6')$
- (6) mit $(2'), (3'), \ldots, (6')$

- (2) mit $(2'), (3'), \ldots, (5')$
- (3) mit $(3'), (4'), (5')$
- (4) mit $(4'), (5')$
- (5) mit $(5')$

Zunächst stellen wir fest, dass wegen $g(a,b) = 0$ für $a \leq \vartheta_1$ im Fall (1) unabhängig von der Wahl von (x_2, y_2) gilt:

$$f(\mathbf{x}, \mathbf{y}) = g(\mathbf{y}) - g(x_1, y_2) - g(y_1, x_2) + g(x_1, x_2) - [\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}] = -[\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}] \leq 0$$

Folglich wird das Maximum in diesem Fall nicht erreicht. Für die Kombination $(6) - (i')$ mit $i \in \{2, \ldots, 6\}$ gilt $\gamma_1 \leq x_1 \leq y_1 \leq 1$ und $y_2 > \vartheta_2$, also:

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) =& (\gamma_1 - \vartheta_1)(y_2 \wedge \gamma_2 - \vartheta_2) - (\gamma_1 - \vartheta_1)(y_2 \wedge \gamma_2 - \vartheta_2) \\ & - (\gamma_1 - \vartheta_1)(x_2 \wedge \gamma_2 - \vartheta_2) I_{\{x_2 > \vartheta_2\}} + (\gamma_1 - \vartheta_1)(x_2 \wedge \gamma_2 - \vartheta_2) I_{\{x_2 > \vartheta_2\}} \\ & - [\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}] \\ =& - [\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}] \leq 0 \end{aligned}$$

Folglich wird auch hier das Maximum nicht erreicht. In den Fällen $(2) - (2')$ und $(2) - (3')$ gilt $0 \leq x_1 \leq \vartheta_1 < y_1 \leq \gamma_1$ und $x_2 \leq \vartheta_2 < y_2$ und somit:

$$f(\mathbf{x}, \mathbf{y}) = (y_1 - \vartheta_1)(y_2 \wedge \gamma_2 - \vartheta_2) - 0 - 0 + 0 - [\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}]$$

Dies wird maximal, wenn der Subtrahend minimal, also $\mathbf{x}$ maximal - und somit gleich $\boldsymbol{\vartheta}$ - wird. Dann gilt:

$$\begin{aligned} f(\boldsymbol{\vartheta}, \mathbf{y}) &= (y_1 - \vartheta_1)(y_2 \wedge \gamma_2 - \vartheta_2) - [\mathbf{y} - \boldsymbol{\vartheta}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}] \\ &= (y_1 - \vartheta_1)\{y_2 \wedge \gamma_2 - \vartheta_2 - (\gamma_1 - \vartheta_1)(y_2 - \vartheta_2)(\gamma_2 - \vartheta_2)\} \\ &= (y_1 - \vartheta_1)\{y_2 \wedge \gamma_2 - \vartheta_2 - (\gamma_1 - \vartheta_1)(y_2 \wedge \gamma_2 - \vartheta_2)(y_2 \vee \gamma_2 - \vartheta_2)\} \\ &= (y_1 - \vartheta_1) \underbrace{(y_2 \wedge \gamma_2 - \vartheta_2)}_{>0} (1 - \underbrace{(\gamma_1 - \vartheta_1)(y_2 \vee \gamma_2 - \vartheta_2)}_{<1}), \end{aligned}$$

wobei $a \vee b := \max\{a, b\}$ ist. Dies wird maximal für maximales y_1. Ist $y_2 \leq \gamma_2$, so maximiert ein maximales y_2 die Funktion. Für $y_2 \geq \gamma_2$ wird die Funktion größer, je kleiner y_2 ist. Folglich wird das Maximum an der behaupteten Stelle angenommen. Analog gilt in den Fällen $(2) - (4')$ und $(2) - (5')$ $0 \leq x_1 \leq \vartheta_1 < y_1 \leq \gamma_1$ und $y_2 > x_2 \geq \vartheta_2$ und $x_2 \leq \gamma_2$ und somit:

$$f(\mathbf{x}, \mathbf{y}) = (y_1 - \vartheta_1)(y_2 \wedge \gamma_2 - \vartheta_2) - 0 - (y_1 - \vartheta_1)(x_2 - \vartheta_2) + 0 - [\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}]$$

Man erkennt, dass dies größer wird, je größer x_1 ist. Also können wir $x_1 = \vartheta_1$ einsetzen und erhalten:

$$\begin{aligned} f(\vartheta_1, x_2, y_1, y_2) &= (y_1 - \vartheta_1)\{y_2 \wedge \gamma_2 - x_2 - (y_2 - x_2)[\boldsymbol{\gamma} - \boldsymbol{\vartheta}]\} \\ &= (y_1 - \vartheta_1)\{y_2 \wedge \gamma_2 - y_2 \underbrace{[\boldsymbol{\gamma} - \boldsymbol{\vartheta}]}_{<1} - x_2(1 - [\boldsymbol{\gamma} - \boldsymbol{\vartheta}])\} \end{aligned}$$

Dies wird größer, je kleiner x_2 und dann je größer y_1 wird. Für $y_2 \leq \gamma_2$ wird die Funktion größer, je größer y_2 ist und für $y_2 \geq \gamma_2$ für kleines y_2. Folglich wird das Maximum auch in diesen Fällen bei $(\boldsymbol{\vartheta}, \boldsymbol{\gamma})$ angenommen. In den Fällen $(3) - (i')$, $i = 3, 4, 5$ gilt $0 \leq x_1 \leq \vartheta_1 < \gamma_1 \leq y_1$, $y_2 > \vartheta_2$ und $x_2 \leq \gamma_2$ und somit:

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= (\gamma_1 - \vartheta_1)(y_2 \wedge \gamma_2 - \vartheta_2) - 0 - (\gamma_1 - \vartheta_1)(x_2 - \vartheta_2)I_{\{x_2 \geq \vartheta_2\}} \\ &\quad + 0 - [\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}] \\ &= (\gamma_1 - \vartheta_1)\{y_2 \wedge \gamma_2 - \vartheta_2 - (x_2 - \vartheta_2)I_{\{x_2 \geq \vartheta_2\}} - [\mathbf{y} - \mathbf{x}](\gamma_2 - \vartheta_2)\} \end{aligned}$$

Man erkennt, dass die Funktion für y_1 minimal und x_1 maximal maximiert wird. Einsetzen von $y_1 = \gamma_1$ und $x_1 = \vartheta_1$ liefert:

$$\begin{aligned} f(\vartheta_1, x_2, \gamma_1, y_2) &= (\gamma_1 - \vartheta_1)\{y_2 \wedge \gamma_2 - \vartheta_2 - (x_2 - \vartheta_2)I_{\{x_2 \geq \vartheta_2\}} - (y_2 - x_2)[\boldsymbol{\gamma} - \boldsymbol{\vartheta}]\} \\ &= (\gamma_1 - \vartheta_1)\{y_2 \wedge \gamma_2 - x_2 \vee \vartheta_2 - (y_2 - x_2)[\boldsymbol{\gamma} - \boldsymbol{\vartheta}]\} \\ &= (\gamma_1 - \vartheta_1)\{y_2 \wedge \gamma_2 - y_2[\boldsymbol{\gamma} - \boldsymbol{\vartheta}] - x_2 \vee \vartheta_2 + x_2[\boldsymbol{\gamma} - \boldsymbol{\vartheta}]\} \end{aligned}$$

$y_2 \wedge \gamma_2 - y_2[\boldsymbol{\gamma} - \boldsymbol{\vartheta}]$ wird für $y_2 \leq \gamma_2$ größer, je größer y_2 ist und für $y_2 \geq \gamma_2$ wird der Term größer, je kleiner y_2 ist. Analog wird $-x_2 \vee \vartheta_2 + x_2[\boldsymbol{\gamma} - \boldsymbol{\vartheta}]$ für $x_2 \leq \vartheta_2$ größer, je größer x_2 ist, und für $x_2 \geq \vartheta_2$, je kleiner x_2 ist. Auch in diesen Fällen wird das Maximum also an der behaupteten Stelle angenommen. Für $(4) - (4')$ und $(4) - (5')$ gilt $\vartheta_1 \leq x_1 < y_1 \leq \gamma_1$, $\vartheta_2 \leq x_2 \leq \gamma_2$ und $\vartheta_2 > y_2$, also:

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= (y_1 - \vartheta_1)(y_2 \wedge \gamma_2 - \vartheta_2) - (x_1 - \vartheta_1)(y_2 \wedge \gamma_2 - \vartheta_2) - (y_1 - \vartheta_1)(x_2 - \vartheta_2) \\ &\quad + (x_1 - \vartheta_1)(x_2 - \vartheta_2) - [\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}] \\ &= (y_1 - x_1)(y_2 \wedge \gamma_2 - \vartheta_2) - (y_1 - x_1)(x_2 - \vartheta_2) - [\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}] \\ &= (y_1 - x_1)\{y_2 \wedge \gamma_2 - x_2 - (y_2 - x_2)[\boldsymbol{\gamma} - \boldsymbol{\vartheta}]\} \\ &= (y_1 - x_1)\{y_2 \wedge \gamma_2 - y_2[\boldsymbol{\gamma} - \boldsymbol{\vartheta}] - x_2(1 - [\boldsymbol{\gamma} - \boldsymbol{\vartheta}])\} \end{aligned}$$

Dies wird größer für minimales x_2 und dann für maximales y_1 und minimales x_1. Der Term $y_2 \wedge \gamma_2 - y_2[\boldsymbol{\gamma} - \boldsymbol{\vartheta}]$ wird für $y_2 \leq \gamma_2$ genauso wie für $y_2 \geq \gamma_2$ bei $y_2 = \gamma_2$ maximiert, also ergibt sich auch hier die Maximalstelle $(\boldsymbol{\vartheta}, \boldsymbol{\gamma})$. Schließlich gilt im Fall $(5) - (5')$, dass $\vartheta_i \leq x_i \leq \gamma_i \leq y_i$ $(i = 1, 2)$ und somit:

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= (\gamma_1 - \vartheta_1)(\gamma_2 - \vartheta_2) - (x_1 - \vartheta_1)(\gamma_2 - \vartheta_2) - (\gamma_1 - \vartheta_1)(x_2 - \vartheta_2) \\ &\quad + (x_1 - \vartheta_1)(x_2 - \vartheta_2) - [\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}] \\ &= (\gamma_1 - \vartheta_1)(\gamma_2 - x_2) - (x_1 - \vartheta_1)(\gamma_2 - x_2) - [\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}] \\ &= [\boldsymbol{\gamma} - \mathbf{x}] - [\mathbf{y} - \mathbf{x}][\boldsymbol{\gamma} - \boldsymbol{\vartheta}] \end{aligned}$$

Dies wird für minimales $\mathbf{y}$ - also für $\mathbf{y} = \boldsymbol{\gamma}$ - maximal. Einsetzen liefert:

$$f(\mathbf{x}, \boldsymbol{\gamma}) = [\boldsymbol{\gamma} - \mathbf{x}](1 - [\boldsymbol{\gamma} - \boldsymbol{\vartheta}])$$

Somit wird das Maximum bei minimalem $\mathbf{x}$ erreicht und es folgt die Behauptung.$\square$

Beweis von Satz 64. Wir betrachten K_d und $G_{n,d}$ aus Definition 63 und

$$F_d(\mathbf{x},\mathbf{y}) := \begin{cases} |\delta||f(x,y)|, & d=1 \\ \delta f(\mathbf{x},\mathbf{y}), & d=2 \end{cases}$$

Zu zeigen ist die stochastische Konvergenz von

$$(\hat{\boldsymbol{\vartheta}}_n, \hat{\boldsymbol{\gamma}}_n) \in \arg\max_{K_d} G_{n,d}$$

gegen

$$(\boldsymbol{\vartheta}, \boldsymbol{\gamma}) \in \arg\max_{K_d} F_d.$$

Dazu benutzen wir einen ähnlichen Ansatz wie Döring (2007) in seinem Beweis von Theorem 1.1 und zeigen, dass jede Teilfolge von $(\hat{\boldsymbol{\vartheta}}_n, \hat{\boldsymbol{\gamma}}_n)_{n\in\mathbb{N}}$ eine Teilteilfolge hat, die fast sicher gegen $(\boldsymbol{\vartheta}, \boldsymbol{\gamma})$ konvergiert. Nach Korollar 6.13 (S. 128) von Klenke (2006) folgt daraus die gewünschte stochastische Konvergenz. Sei $(\hat{\boldsymbol{\vartheta}}_{n'}, \hat{\boldsymbol{\gamma}}_{n'})$ eine Teilfolge von $(\hat{\boldsymbol{\vartheta}}_n, \hat{\boldsymbol{\gamma}}_n)_{n\in\mathbb{N}}$. Lemma 65 liefert:

$$\sup_{(\mathbf{x},\mathbf{y})\in K_d} \{|G_{n,d}(\mathbf{x},\mathbf{y}) - F_d(\mathbf{x},\mathbf{y})|\} = o_P(1)$$

Wir definieren:

$$h_{n,d} := \sup_{(\mathbf{x},\mathbf{y})\in K_d} \{|G_{n,d}(\mathbf{x},\mathbf{y}) - F_d(\mathbf{x},\mathbf{y})|\}$$

Da $h_{n,d} = o_P(1)$ ist, gibt es nach Korollar 6.13 von Klenke (2006) eine Teilteilfolge $(n'') \subset (n')$, für die $h_{n'',d}$ fast sicher gegen 0 konvergiert. Sei Ω_0 die Einsmenge, auf der $h_{n'',d}$ gegen 0 konvergiert. Aus Lemma 67 wissen wir, dass $|\delta||f(\cdot)| : K_1 \to \mathbb{R}$ bzw. $\delta f(\cdot) : K_2 \to \mathbb{R}$ stetig mit eindeutiger Maximalstelle $(\boldsymbol{\vartheta}, \boldsymbol{\gamma})$ sind. Somit liefert Lemma 66, dass für jedes $\omega \in \Omega_0$ gilt:

$$(\hat{\boldsymbol{\vartheta}}_{n''}(\omega), \hat{\boldsymbol{\gamma}}_{n''}(\omega)) \overset{n''\to\infty}{\longrightarrow} (\boldsymbol{\vartheta}, \boldsymbol{\gamma})$$

Da $P(\Omega_0) = 1$ ist, liefert dies die Behauptung. $\square$

5 Beispiele

5.1 Schwach $\mathcal{M}$-abhängige Prozesse

5.1.1 Starkes Invarianzprinzip

Zunächst betrachten wir Prozesse mit eindimensionalem Zeitparameter. Dabei konzentrieren wir uns auf sogenannte „in L^p schwach $\mathcal{M}$-abhängige" Prozesse. Wie wir sehen werden, bilden diese eine breite Klasse von stationären Prozessen und liefern daher einige Beispiele für die Anwendbarkeit des in dieser Arbeit betrachteten Invarianzprinzips. Eingeführt wurde diese Abhängigkeitsstruktur von Berkes et al. (2011), die ein starkes Invarianzprinzip für in L^p schwach $\mathcal{M}$-abhängige Prozesse bewiesen. Dieses wird im diesem Abschnitt kurz vorgestellt. Im Folgeabschnitt werden dann der Zentrale Grenzwertsatz, sowie das schwache Invarianzprinzip mit Hilfe von Satz 35 aus dem gegebenen starken Invarianzprinzip gefolgert. Schließlich werden einige Beispiele für in L^p schwach $\mathcal{M}$-abhängige Prozesse betrachtet.

Die in diesem Abschnitt vorgestellten Definitionen und Aussagen wurden aus Berkes et al. (2011) entnommen. Zunächst führen wir die Prozessklasse ein, die wir im Folgenden betrachten werden:

Definition 68. *(siehe Berkes et al. (2011), Definition 1) Sei $\{Y_k\}_{k\in\mathbb{Z}}$ ein stochastischer Prozess , $p \geq 1$ und $\delta : \mathbb{N} \to \mathbb{R}_+$ eine Funktion mit $\delta(m) \to 0$. $\{Y_k\}_{k\in\mathbb{Z}}$ heißt schwach $\mathcal{M}$-abhängig in L^p mit Ratenfunktion $\delta(\cdot)$, falls gilt:*

(A) Für alle $k \in \mathbb{Z}$, $m \in \mathbb{N}$ gibt es Zufallsvariablen $Y_k^{(m)} \in L^p$, so dass gilt:

$$\|Y_k - Y_k^{(m)}\|_p \leq \delta(m)$$

(B) Für $r \in \mathbb{N}$ und disjunkte Intervalle $I_1, \dots, I_r$ von ganzen Zahlen und $m_1, \dots, m_r \in \mathbb{N}$ so, dass $d(I_k, I_l) > \max\{m_k, m_l\}$ für $1 \leq k < l \leq r$, sind die Vektoren $\{Y_j^{(m_1)} : j \in I_1\}, \dots, \{Y_j^{(m_r)} : j \in I_r\}$ unabhängig.

In der Literatur gibt es eine Vielzahl Definitionen von schwacher Abhängigkeit, von denen einige der hier gewählten Approximierbarkeit ähneln. Oft wird aber eine bestimmte Darstellung für den Prozess gefordert, z. B. die Eigenschaft, dass es eine Familie $\{\varepsilon_k\}_{k\in\mathbb{Z}}$ von i. i. d. Zufallsvariablen und eine messbare Funktion f gibt, so dass Y_k für alle $k \in \mathbb{Z}$ die kausale Darstellung

$$Y_k = f(\varepsilon_k, \varepsilon_{k-1}, \dots) \tag{5.1}$$

(vgl. z. B. Aue et al. (2009), Wu (2007), Liu und Lin (2009)), oder die allgemeinere zweiseitige Darstellung

$$Y_k = f(\dots, \varepsilon_{k-1}, \varepsilon_k, \varepsilon_{k+1}, \dots) \tag{5.2}$$

(vgl. z. B. Berkes et al. (2009)) hat. Eine Abhängigkeitsstruktur für solche Prozesse kann dann mit Hilfe der geforderten Darstellung definiert werden. In einem solchen Modell können approximierende Variablen zu Y_k z. B. als messbare Funktionen der $\varepsilon_k, \dots, \varepsilon_{k-m}$ (siehe Aue et al. (2009)) oder als bedingte Erwartungswerte $E[Y_k|\varepsilon_{k-m}, \dots, \varepsilon_k]$ (vgl. Liu und Lin (2009)) definiert werden. Die Definition der schwachen $\mathcal{M}$-Abhängigkeit wurde zwar von solchen Fällen inspiriert, aber es wurde bewusst weder für den Prozess $\{Y_k\}_{k\in\mathbb{Z}}$ noch für die approximierenden Variablen $\{Y_k^{(m)} : m \in \mathbb{N}, k \in \mathbb{Z}\}$ eine bestimmte Darstellung gefordert. Damit wird später deutlich werden, dass die für das asymptotische Verhalten des Prozesses entscheidende Eigenschaft seine Approximierbarkeit durch m-abhängige Zufallsvariablen (siehe Abschnitt 5.1.3) ist.

Bemerkung 69.

1. *Obige Definition impliziert, dass die Y_k in L^p liegen:*

$$\|Y_k\|_p \leq \|Y_k - Y_k^{(m)}\|_p + \|Y_k^{(m)}\|_p < \infty$$

2. *Aufgrund von Eigenschaft (B) ist für festes $m \in \mathbb{N}$ die Familie $\{Y_k^{(m)}\}_{k\in\mathbb{Z}}$ ein m-abhängiger Prozess (vgl. Abschnitt 5.1.3). Eigenschaft (A) bedeutet also, dass ein schwach $\mathcal{M}$-abhängiger Prozess für jedes $m \in \mathbb{N}$ L^p-approximierbar durch m-abhängige Prozesse ist, wobei $\delta(m)$ den Approximationsfehler liefert.*

3. *Schwache $\mathcal{M}$-Abhängigkeit bleibt unter glatten Transformationen erhalten: Ist $\{Y_k\}_{k\in\mathbb{Z}}$ ein in L^p schwach $\mathcal{M}$-abhängiger Prozess mit Ratenfunktion $\delta(\cdot)$ und $h : \mathbb{R} \to \mathbb{R}$ eine Lipschitz-stetige Funktion der Ordnung α $(0 < \alpha \leq 1)$ mit Lipschitz-Konstante K, so ist $\{h(Y_k)\}_{k\in\mathbb{Z}}$ ebenfalls schwach $\mathcal{M}$-abhängig in L^p mit Ratenfunktion $\tilde{\delta}(\cdot) = K\delta(\cdot)$.*

Nun wollen wir das von Berkes et al. (2011) gezeigte starke Invarianzprinzip zitieren. Dabei verstehen wir unter einem *starken Invarianzprinzip* die Eigenschaft, dass die zu einem stochastischen Prozess $\{X_k\}_{k\in\mathbb{Z}}$ gehörige Partialsummenfolge $(S_n = \sum_{i=1}^n X_i)_{n\in\mathbb{N}}$ (gegebenenfalls nach Übergang auf einen anderen Wahrscheinlichkeitsraum) fast sicher durch einen (oder mehrere) Wiener-Prozess $\{W(t)\}_{t\geq 0}$ approximiert werden kann. Eine solche Approximation ist von großem Interesse in der Statistik und Wahrscheinlichkeitstheorie, da damit Aussagen wie der Zentrale Grenzwertsatz oder das Gesetz vom iterierten Logarithmus über das asymptotische Verhalten der Partialsummenfolge auf entsprechende Aussagen über Wiener-Prozesse zurückgeführt werden können. Komlós et al. (1975, 1976) zeigten für zentrierte und standardisierte i. i. d. X_k, die $E[e^{tX_k}] < \infty$ für $|t| < t_0$ und ein $t_0 > 0$ erfüllen, dass

$$S_n = W(n) + \mathcal{O}(\log n) \; P - f.s..$$

Außerdem zeigten sie

$$S_n = W(n) + o(n^{1/p})\ P - f.s.,$$

unter der Voraussetzung, dass statt $E[e^{tX_k}] < \infty$ nur $E[|X_k|^p] < \infty$ für ein $p > 2$ gilt. Setzt man für einen i. i. d. Prozess nur die Zentriertheit und $E[X_k^2] = 1$ voraus, so greift nach Strassen (1964) die Approximation

$$S_n = W(n) + o(\sqrt{n \log \log n})\ P - f.s.,$$

welche ohne die Annahme der Existenz weiterer Momente auch optimal ist (vgl. Major (1976)), aber nicht ausreicht, um Resultate wie den Zentralen Grenzwertsatz zu implizieren. Um dennoch eine stärkere Approximation zu erhalten, bewies Major (1979) unter den gleichen Voraussetzungen an $\{X_k\}_{k\in\mathbb{Z}}$ ein starkes Invarianzprinzip, bei dem nicht mehr $W(n)$, sondern $W(\tau_n)$ mit $\tau_n \sim n$ betrachtet wurde:

$$S_n = W(\tau_n) + o(n^{1/2})\ P - f.s.$$

Die Verallgemeinerung der für den i. i. d. Fall bekannten Aussagen auf schwach abhängige Prozesse ist Gegenstand aktueller Forschung. Wu (2007) zeigte für die von ihm betrachteten stationären Prozesse mit „physical dependence“ und $E|X_k|^p < \infty$ für $2 < p \leq 4$ ein starkes Invarianzprinzip der Form

$$S_n = W(n) + o(n^{1/p}(\log n)^\gamma)\ P - f.s.,$$

wobei $\gamma > 0$ ist. Liu und Lin (2009) zeigten, dass die Approximation sogar ohne den Term $(\log n)^\gamma$ möglich ist, sodass man für $2 < p \leq 4$ die Komlós-Major-Tusnády-Schranke erhält. Um für schwach abhängige Prozesse ein starkes Invarianzprinzip zu erhalten, das die Approximationsgüte $o(n^{1/p})$ auch für allgemeines $p > 2$ liefert, benutzen Berkes et al. (2011) zwei Wiener-Prozesse und lassen eine Perturbation des Zeitparamters n zu. Sie erhalten Resultate der Form

$$S_n = W_1(s_n^2) + W_2(t_n^2) + \mathcal{O}(n^{(1+\eta)/p})\ P - f.s..$$

Wie wir im nächsten Abschnitt sehen werden, sind dabei die Folgen $(s_n^2)_{n\in\mathbb{N}}$ und $(t_n^2)_{n\in\mathbb{N}}$ so gewählt, dass $W(t_n^2)$ im Vergleich zu $W(s_n^2)$ asymptotisch vernachlässigbar ist.

Satz 70. *(siehe Berkes et al. (2011), Theorem 1) Seien $p > 2$, $\eta > 0$ und $\{Y_k\}_{k\in\mathbb{Z}}$ ein zentrierter, stationärer, in L^p schwach $\mathcal{M}$-abhängiger stochastischer Prozess mit Ratenfunktion $\delta(\cdot)$, so dass*

$$\delta(m) \ll m^{-A}, \tag{5.3}$$

mit

$$A > \max\left\{\frac{p-2}{2\eta}\left(1 - \frac{1+\eta}{p}\right), 1\right\} \quad \textit{und} \quad \frac{1+\eta}{p} < \frac{1}{2}. \tag{5.4}$$

Dann ist die Reihe

$$\sigma^2 = \sum_{k\in\mathbb{Z}} EY_0Y_k \tag{5.5}$$

absolut konvergent und $\{Y_k\}_{k\in\mathbb{Z}}$ kann auf einem neuen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ zusammen mit zwei Wiener-Prozessen $\{W_1(t)\}_{t\geq 0}$ und $\{W_2(t)\}_{t\geq 0}$ definiert werden, so dass

$$\sum_{k=1}^{n} Y_k = W_1(s_n^2) + W_2(t_n^2) + \mathcal{O}(n^{(1+\eta)/p})\ P - f.s., \tag{5.6}$$

wobei $(s_n^2)_{n\in\mathbb{N}}$ und $(t_n^2)_{n\in\mathbb{N}}$ nichtfallende reelle Folgen sind, für die es $\gamma \in (0,1)$ und $c > 0$ gibt, so dass

$$s_n^2 \sim \sigma^2 n \quad \text{und} \quad t_n^2 \sim cn^\gamma. \tag{5.7}$$

Satz 71. *(siehe Berkes et al. (2011), Theorem 2) Seien $p > 2$, $\rho > 0$ und $\{Y_k\}_{k\in\mathbb{Z}}$ ein zentrierter, stationärer, in L^p schwach $\mathcal{M}$-abhängiger stochastischer Prozess mit Ratenfunktion $\delta(\cdot)$, so dass*

$$\delta(m) \ll \exp(-\rho m). \tag{5.8}$$

Dann ist die Reihe (5.5) absolut konvergent und $\{Y_k\}_{k\in\mathbb{Z}}$ kann auf einem neuen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ zusammen mit zwei Wiener-Prozessen $\{W_1(t)\}_{t\geq 0}$ und $\{W_2(t)\}_{t\geq 0}$ definiert werden, so dass

$$\sum_{k=1}^{n} Y_k = W_1(s_n^2) + W_2(t_n^2) + \mathcal{O}(n^{1/p}\log^2 n)\ P - f.s., \tag{5.9}$$

wobei $(s_n^2)_{n\in\mathbb{N}}$ und $(t_n^2)_{n\in\mathbb{N}}$ nichtfallende reelle Folgen sind, die

$$s_n^2 \sim \sigma^2 n, \quad t_n^2 \sim \sigma^2 n/\log n \tag{5.10}$$

erfüllen.

Bemerkung 72. *Die Folgen $(s_n^2)_{n\in\mathbb{N}}$ und $(t_n^2)_{n\in\mathbb{N}}$ werden mit Hilfe eines Blockbildungs-Arguments explizit konstruiert. Es lässt sich zeigen, dass*

$$\limsup_{n\to\infty}(s_{n+1}^2 - s_n^2) = \limsup_{n\to\infty}(t_{n+1}^2 - t_n^2) = \sigma^2. \tag{5.11}$$

Die Idee des Beweises dieser Aussagen besteht darin, zunächst die Partialsummen von Y durch Partialsummen über approximierende Variablen zu ersetzen und dann deren Abhängigkeitseigenschaften zu nutzen, um mit Hilfe von Blockbildung das Invarianzprinzip für diese approximierenden Partialsummen herzuleiten. Wir verzichten in dieser Arbeit auf den Beweis der Sätze 70 und 71 und verweisen interessierte Leser auf Berkes et al. (2011). Ab jetzt werden wir folgende Annahme voraussetzen:

Annahme 2. *$\{Y_k\}_{k\in\mathbb{Z}}$ sei ein stochastischer Prozess, der auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ zusammen mit zwei Wiener-Prozessen $\{W_1(t)\}_{t\geq 0}$, $\{W_2(t)\}_{t\geq 0}$ definiert werden kann, so dass*

$$\sum_{k=1}^{n} Y_k = W_1(s_n^2) + W_2(t_n^2) + \mathcal{O}(E_n)\ P - f.s., \tag{5.12}$$

für eine Folge $\{E_n\}_{n\in\mathbb{N}}$ von reellen Zahlen mit $E_n = o(\sqrt{n})$ und nichtfallende Folgen $\{s_n^2\}_{n\in\mathbb{N}}$ und $\{t_n^2\}_{n\in\mathbb{N}}$ positiver reeller Zahlen, die

$$s_n^2 \sim \sigma^2 n,\ t_n^2 = o(n),\ \limsup_{k\to\infty}(s_{k+1}^2 - s_k^2) = \limsup_{k\to\infty}(t_{k+1}^2 - t_k^2) = \sigma^2 \tag{5.13}$$

für ein $\sigma^2 \in (0,\infty)$ erfüllen.

Bemerkung 73. *Erfüllt ein Prozess $\{Y_k\}_{k\in\mathbb{Z}}$ die Voraussetzungen von Satz 70 oder 71, so erfüllt er Annahme 2 mit*

(i) $E_n = \mathcal{O}(n^{\frac{1+\eta}{p}})$ *oder*

(ii) $E_n = \mathcal{O}(n^{1/p}\log^2 n)$,

wobei $p > 2$ und $(1+\eta)/p < 1/2$ sind. Damit gilt auch $E_n = o(\sqrt{n})$, denn wegen $p > 2$ gilt $1/p - 1/2 < 0$, also $\mathcal{O}(n^{1/p-1/2}\log^2 n) = o(1)$, und wegen $(1+\eta)/p < 1/2$ gilt $\mathcal{O}(n^{(1+\eta)/p-1/2}) = o(1)$.

Nun betrachten wir ein Hilfslemma, welches den Übergang von $W(s_n^2)$ bzw. $W(t_n^2)$ zu $W(t)$ ermöglicht:

Lemma 74. *Sei $0 < \alpha < 1/2$. Für die Folgen $(s_n^2)_{n\in\mathbb{N}}$ und $(t_n^2)_{n\in\mathbb{N}}$ aus Annahme 2 und einen Wiener-Prozess $\{W(t)\}_{t\geq 0}$ gilt:*

(i) Es gibt eine Nullfolge $\{\varepsilon_n\}_{n\in\mathbb{N}}$, so dass:

$$\max_{1\leq i\leq n} |i - s_i^2/\sigma^2| \leq \varepsilon_n n$$

(ii) Für die Nullfolge aus (i) gilt:

$$\max_{1\leq i\leq n} |W(i) - W(s_i^2/\sigma^2)| = o_P(\sqrt{n}\varepsilon_n^\alpha)$$

Insbesondere gilt:

$$\max_{1\leq i\leq n} |W(i) - W(s_i^2/\sigma^2)| = o_P(\sqrt{n})$$

(iii) Definiere $\gamma_n := t_n^2/n$. Dann ist $\gamma_n = o(1)$ und es gilt:

$$\max_{1\leq i\leq n} |W(t_i^2/\sigma^2)| = o_P(\sqrt{n}\,\gamma_n^\alpha)$$

Insbesondere:

$$\max_{1\leq i\leq n} |W(t_i^2/\sigma^2)| = o_P(\sqrt{n})$$

Beweis. siehe Anhang □

Zum Schluss dieses Abschnitts betrachten wir noch eine Momentenungleichung aus Berkes et al. (2011) (Proposition 4), die wir im folgenden Abschnitt zum Beweis der Straffheit nutzen werden.

Satz 75. *Seien $p \geq 2$ und $\{Y_k\}_{k\in\mathbb{Z}}$ ein zentrierter, stationärer, in L^p schwach $\mathcal{M}$-abhängiger Prozess mit Ratenfunktion $\delta(\cdot)$. $\delta(\cdot)$ erfülle*

$$\sum_{m=0}^{\infty} \delta(m) < \infty. \tag{5.14}$$

Dann gilt für beliebige $n \in \mathbb{N}$, $b \in \mathbb{Z}$

$$E\left|\sum_{k=b+1}^{b+n} Y_k\right|^p \leq C_p\, n^{p/2}, \tag{5.15}$$

wobei $C_p > 0$ eine nur von p und $\{Y_k\}_{k\in\mathbb{Z}}$ abhängige reelle Zahl ist.

Bemerkung 76. *Die Bedingung* (5.14) *gilt für alle Prozesse, die die Voraussetzungen der Sätze 70 oder 71 erfüllen:*

1. Fall: $\{Y_k\}_{k\in\mathbb{Z}}$ *erfülle die Voraussetzung von Satz 70. Wegen $A > 1$ ist $\sum_{m=1}^{\infty} \frac{1}{m^A} < \infty$ und somit gibt es wegen $\delta(m) \ll m^{-A}$ ein $C > 0$ und $m_0 \in \mathbb{N}$, so dass*

$$\sum_{m=0}^{\infty} \delta(m) \leq \underbrace{\sum_{m=0}^{m_0} \delta(m)}_{<\infty} + C \underbrace{\sum_{m=m_0+1}^{\infty} \frac{1}{m^A}}_{<\infty} < \infty.$$

2. Fall: $\{Y_k\}_{k\in\mathbb{Z}}$ *erfülle die Voraussetzung von Satz 71. Wegen $\exp(-\rho) < 1$ ist $\sum_{m=0}^{\infty} \exp(-\rho m) < \infty$ und somit gibt es wegen $\delta(m) \ll \exp(-\rho m)$ ein $C > 0$ und $m_0 \in \mathbb{N}$, so dass*

$$\sum_{m=0}^{\infty} \delta(m) \leq \underbrace{\sum_{m=0}^{m_0} \delta(m)}_{<\infty} + C \underbrace{\sum_{m=m_0+1}^{\infty} \exp(-\rho m)}_{<\infty} < \infty.$$

5.1.2 Das schwache Invarianzprinzip

Im Folgenden zeigen wir die Verteilungskonvergenz des zu einem schwach $\mathcal{M}$-abhängigen Prozess gehörigen Partialsummenprozesses. Analog zum Vorgehen in Billingsley (1968) (vgl. Beweise von Theorem 10.1 und Theorem 16.1), zeigen wir dazu zuerst die Konvergenz der endlich-dimensionalen Verteilungen und dann die Straffheit des Prozesses, wobei wir für letzteres das Kriterium von Bickel und Wichura (1971) aus Abschnitt 2.2.2 benutzen. Da wir keine Darstellung des Prozesses der Form (5.1) oder (5.2) gefordert haben, unterscheidet sich der hier vorgestellte Konvergenzbeweis von bereits für ähnliche Abhängigkeitsstrukturen geführten Konvergenzbeweisen (vgl. Berkes et al. (2008) und Aue et al. (2009)), in denen die Darstellung des Prozesses ausgenutzt werden konnte, um Theorem 21.1 aus Billingsley (1968) anzuwenden. Zunächst zeigen wir den Zentralen Grenzwertsatz für schwach $\mathcal{M}$-abhängige Prozesse:

Satz 77. *$\{Y_k\}_{k\in\mathbb{Z}}$ sei ein zentrierter stochastischer Prozess in L^p ($p > 2$). Der Prozess erfülle Annahme 2. Dann gilt:*

$$\frac{1}{\sigma\sqrt{n}}\sum_{k=1}^{n} Y_k \xrightarrow{\mathcal{D}} N(0,1) \tag{5.16}$$

Beweis. Annahme 2 liefert:

$$\frac{1}{\sqrt{n}\sigma}\sum_{k=1}^{n} Y_k \stackrel{\mathcal{D}}{=} \frac{1}{\sqrt{n}}W_1(s_n^2/\sigma^2) + \frac{1}{\sqrt{n}}W_2(t_n^2/\sigma^2) + o_P(1) =: K_1 + K_2 + o_P(1)$$

Um die Notation zu vereinfachen sei im Folgenden o. E. $\sigma = 1$. Lemma 74 (iii) liefert $K_2 = o_P(1)$. Betrachte nun K_1. Es gilt:

$$K_1 = \frac{1}{\sqrt{n}}W_1(n) + \frac{1}{\sqrt{n}}(W_1(s_n^2) - W_1(n))$$

Aufgrund von Lemma 74 (*ii*) ist $\frac{1}{\sqrt{n}}(W_1(s_n^2) - W_1(n)) = o_P(1)$. Die Skalierungseigenschaft des Wiener-Prozesses liefert, dass $\frac{1}{\sqrt{n}}W_1(n) \stackrel{\mathcal{D}}{=} W_1(1)$ für alle $n \in \mathbb{N}$ standardnormalverteilt ist. Somit liefert Lemma 34 die Behauptung. □

Nun wenden wir uns dem schwachen Invarianzprinzip zu. Wir betrachten das Funktional

$$\tilde{S}_n(t) := \frac{1}{\sqrt{n}}\sum_{k=1}^{\lfloor nt\rfloor} Y_k, \quad 0 \le t \le 1. \tag{5.17}$$

Für zentrierte, unabhängig und identisch verteilte Zufallsvariablen mit endlichen zweiten Momenten liefert der Satz von Donsker, dass $\{\tilde{S}_n(t)\}_{t\in[0,1]}$ in $D[0,1]$ schwach konvergiert:

Satz 78. *(vgl. Billingsley (1968), Theorem 10.1, S. 68 und Theorem 16.1, S. 137) $\{Y_k\}_{k\in\mathbb{N}}$ seien unabhängig identisch verteilte Zufallsvariablen auf einem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Weiter seien $S_n = Y_1 + \cdots + Y_n$ und $EY_1 = 0$, $EY_1^2 = \sigma^2 < \infty$. Dann gilt für die Funktion $\tilde{S}_n(t)$ aus* (5.17)*:*

$$\left\{\frac{1}{\sigma}\tilde{S}_n(t)\right\}_{t\in[0,1]} \stackrel{D[0,1]}{\longrightarrow} \{W(t)\}_{t\in[0,1]} \tag{5.18}$$

Im Folgenden wollen wir das Theorem von Donsker auf stationäre Prozesse, die Annahme 2 erfüllen, erweitern, indem wir nachweisen, dass solche Prozesse die Voraussetzungen aus Satz 35 (bzw. Bemerkung 37) erfüllen. Dabei zeigen wir mit Hilfe des gegebenen starken Invarianzprinzips die Konvergenz der endlich-dimensionalen Verteilungen und prüfen dann mittels der Momentenungleichung (5.15) die Straffheit.

Zunächst zeigen wir die Konvergenz der endlich-dimensionalen Verteilungen:

Satz 79. $\{Y_k\}_{k\in\mathbb{Z}}$ *sei ein zentrierter stationärer Prozess in* L^p $(p \geq 2)$*, der Annahme 2 erfüllt.* $\{W(t)\}_{t\geq 0}$ *sei ein Wiener-Prozess. Dann gilt für*
$0 \leq t_1 < t_2 < \cdots < t_k \leq 1$*:*

$$\frac{1}{\sigma}(\tilde{S}_n(t_1), \ldots, \tilde{S}_n(t_n)) \xrightarrow{\mathcal{D}} (W(t_1), \ldots, W(t_n)) \quad (n \to \infty) \tag{5.19}$$

Beweis. Sei wieder o. E. $\sigma = 1$. Für $t = 0$ gilt $\tilde{S}_n(t) = 0 = W(0)$ $P - f.s.$ für alle $n \in \mathbb{N}$, also $\tilde{S}_n(t) \overset{\mathcal{D}}{\to} W(t)$. Für $t \in (0,1]$ und n groß genug ist $\lfloor nt \rfloor > 0$, so dass gilt:

$$\tilde{S}_n(t) = \sqrt{\frac{\lfloor nt \rfloor}{n}} \frac{1}{\sqrt{\lfloor nt \rfloor}} \sum_{k=1}^{\lfloor nt \rfloor} Y_k$$

Wegen

$$\sqrt{\frac{\lfloor nt \rfloor}{n}} = \left(t - \frac{nt - \lfloor nt \rfloor}{n}\right)^{1/2} \overset{n\to\infty}{\longrightarrow} \sqrt{t}$$

folgt aus Satz 77 nach Anwendung des Lemmas von Slutsky:

$$\tilde{S}_n(t) \overset{\mathcal{D}}{\to} \sqrt{t}W(1) \overset{\mathcal{D}}{=} W(t)$$

Insgesamt liefert dies $\tilde{S}_n(t) \overset{\mathcal{D}}{\to} W(t)$ für $t \in [0,1]$.
Seien nun $0 \leq t_1 < t_2 < \ldots t_k \leq 1$. Wir zeigen

$$(\tilde{S}_n(t_1), \tilde{S}_n(t_2), \ldots, \tilde{S}_n(t_k)) \overset{\mathcal{D}}{\to} (W(t_1), W(t_2), \ldots, W(t_k)) \tag{5.20}$$

mit Hilfe der Cramér-Wold Idee (Klenke (2006), Satz 15.55, S. 312), d. h. wir zeigen für $a_1, \ldots, a_k \in \mathbb{R}$:

$$\sum_{i=1}^{k} a_i \tilde{S}_n(t_i) \overset{\mathcal{D}}{\to} \sum_{i=1}^{k} a_i W(t_i) \tag{5.21}$$

Als erstes beobachten wir, dass dies im Falle $t_1 = 0$ äquivalent ist zu

$$\sum_{i=2}^{k} a_i \tilde{S}_n(t_i) \overset{\mathcal{D}}{\to} \sum_{i=2}^{k} a_i W(t_i),$$

mit $0 < t_2 < \cdots < t_k$, so dass wir o. E. annehmen können, dass
$0 < t_1 < \cdots < t_k \leq 1$ und n groß genug ist, so dass $\lfloor nt_k \rfloor > \cdots > \lfloor nt_1 \rfloor \geq 1$. Wir führen die Aussage zurück auf den i. i. d. Fall, indem wir zunächst

$$a_j \tilde{S}_n(t_j) = a_j \frac{1}{\sqrt{n}} W_1(\lfloor nt_j \rfloor) + o_P(1) \tag{5.22}$$

für $j = 1, \ldots, k$ zeigen. Dazu gehen wir analog zum Beweis des Zentralen Grenzwertsatzes (Satz 77) vor und nutzen Annahme 2 und Lemma 74:

$$\begin{aligned}
a_j\tilde{S}_n(t_j) &= a_j\frac{1}{\sqrt{n}}W_1(s^2_{\lfloor nt_j\rfloor}) + \overbrace{a_j\frac{1}{\sqrt{n}}W_1(t^2_{\lfloor nt_j\rfloor})}^{=o_P(1),\text{ wg. Le. 74 }(iii)} + o_P(1)\\
&= a_j\frac{1}{\sqrt{n}}W_1(\lfloor nt_j\rfloor) + \underbrace{a_j\frac{1}{\sqrt{n}}(W_1(s^2_{\lfloor nt_j\rfloor}) - W_1(\lfloor nt_j\rfloor))}_{=o_P(1),\text{ wg. Le. 74 }(ii)} + o_P(1)\\
&= a_j\frac{1}{\sqrt{n}}W_1(\lfloor nt_j\rfloor) + o_P(1)
\end{aligned}$$

(5.22) liefert nun:

$$\sum_{i=1}^{k} a_i\tilde{S}_n(t_i) = \sum_{i=1}^{k} a_i\frac{1}{\sqrt{n}}W_1(\lfloor nt_i\rfloor) + o_P(1)$$

Wir schreiben

$$W_1(\lfloor nt_i\rfloor) = \sum_{k=1}^{\lfloor nt_i\rfloor} Z_k$$

mit Zufallsvariablen $Z_k = W_1(k) - W_1(k-1)$, $k \in \mathbb{N}$. Da $\{W_1(t)\}_{t\geq 0}$ ein Wiener-Prozess ist, sind die Z_k i. i. d. $N(0,1)$-verteilt, so dass wir den Satz von Donsker (Satz 78) auf den Partialsummenprozess der Z_k anwenden können. Die Konvergenz in $D[0,1]$ impliziert die Konvergenz der endlich-dimensionalen Verteilungen, und somit liefert Satz 78:

$$\frac{1}{\sqrt{n}}\left(W_1(\lfloor nt_1\rfloor), W_1(\lfloor nt_2\rfloor), \ldots W_1(\lfloor nt_k\rfloor)\right) \xrightarrow{\mathcal{D}} \left(W(t_1), W(t_2), \ldots, W(t_k)\right),$$

beziehungsweise, nach Anwendung der Cramér-Wold-Idee:

$$\sum_{i=1}^{k} a_i\frac{1}{\sqrt{n}}W_1(\lfloor nt_i\rfloor) \xrightarrow{\mathcal{D}} \sum_{i=1}^{k} a_iW(t_i),$$

und folglich mit Lemma 34 die Behauptung. □

Satz 80. *Die Voraussetzungen von Satz 75 seien für $p > 2$ erfüllt. Weiter gelte die Annahme 2. Es seien $S_n := Y_1 + \cdots + Y_n$ und $\tilde{S}_n(t)$ wie in* (5.17)*. Dann gilt:*

$$\left\{\frac{1}{\sigma}\tilde{S}_n(t)\right\}_{t\in[0,1]} \xrightarrow{D[0,1]} \{W(t)\}_{t\in[0,1]} \tag{5.23}$$

Beweis. Wir wollen Satz 35 anwenden. Es ist $\tilde{S}_n(t) = 0$ für $t = 0$ und der Grenzprozess $\{W(t)\}_{t\in[0,1]}$ hat fast sicher stetige Pfade, ist also insbesondere stetig in 1. In Satz 79 haben wir bereits gezeigt, dass die endlich-dimensionalen Verteilungen des Prozesses gegen die eines Wiener-Prozesses konvergieren. Nun müssen wir noch das Straffheitskriterium nachprüfen. Da S_n nach Satz 75 die in Lemma 38 vorausgesetzte Momentenungleichung erfüllt, folgt aus Lemma 38 und Bemerkung 37 die Behauptung. □

5.1.3 Beispiele

In diesem Abschnitt stellen wir einige Arten von Prozessen vor, die gemäß Definition 68 schwach $\mathcal{M}$-abhängig in L^p sind. Hat die Ratenfunktion eine entsprechend hohe Konvergenzgeschwindigkeit, so erfüllen die zugehörigen Partialsummenprozesse nach Abschnitt 5.1.2 das von uns betrachtete schwache Invarianzprinzip.
Die hier betrachteten Beispiele stellen eine Auswahl aus den von Berkes et al. (2011) behandelten Beispielen dar.

***m*-abhängige Prozesse** Wir bezeichnen einen Prozess $\{Y_k\}_{k\in\mathbb{Z}}$ als <u>*m-abhängig*</u> für ein $m \in \mathbb{N}$, falls für alle $n \in \mathbb{Z}$ die Mengen $\{Y_k : k \leq n\}$ und $\{Y_k : k > n + m\}$ stochastisch unabhängig sind. m-abhängige Prozesse sind schwach $\mathcal{M}$-abhängig:
Satz 81. *Der Prozess $\{Y_k\}_{k\in\mathbb{Z}}$ sei m-abhängig für ein festes $m = m_0 \in \mathbb{N}$ und es gelte $K = \sup_{k\in\mathbb{Z}} \|Y_k\|_p < \infty$ für ein $p > 2$. (Letzteres ist z. B. gegeben, falls $\{Y_k\}_{k\in\mathbb{Z}}$ in L^p und stationär ist.) Dann ist $\{Y_k\}_{k\in\mathbb{Z}}$ $\mathcal{M}$-abhängig in L^p mit*

$$Y_k^{(m)} = \begin{cases} 0, & m < m_0 \\ Y_k, & m \geq m_0 \end{cases}$$

und Ratenfunktion $\delta(m) \ll \exp(-\rho m)$ für $\rho > 0$.

***NED*-Prozesse** Ein stochastischer Prozess $\{Y_k\}_{k\in\mathbb{Z}}$, der darstellbar ist als

$$Y_k = f(\ldots, \varepsilon_{k-1}, \varepsilon_k, \varepsilon_{k+1}, \ldots), \quad k \in \mathbb{Z},$$

mit einer Folge $\{\varepsilon_k\}_{k\in\mathbb{Z}}$ von Zufallsvariablen und einer messbaren Funktion f heißt <u>*NED-Prozess*</u> (near-epoch dependent) bezüglich $\{\varepsilon_k\}_{k\in\mathbb{Z}}$ unter L^p-Norm mit Ratenfunktion $\delta(\cdot)$, wenn zusätzlich gilt:

$$\forall k \in \mathbb{Z}, m \geq 1 : \ \|Y_k - E[Y_k|\mathcal{A}_{k-\lfloor m/2\rfloor}^{k+\lfloor m/2\rfloor}]\|_p \leq \delta(m),$$

wobei $\mathcal{A}_{k-j}^{k+j}$ die von $\varepsilon_{k-j}, \ldots, \varepsilon_{k+j}$ erzeugte σ-Algebra bezeichnet.
Satz 82. *$\{Y_k\}_{k\in\mathbb{Z}}$ sei NED-Prozess mit Ratenfunktion $\delta(\cdot)$ in L^p-Norm bezüglich einer Folge von unabhängigen Zufallsvariablen $\{\varepsilon_k\}_{k\in\mathbb{Z}}$. Dann ist $\{Y_k\}_{k\in\mathbb{Z}}$ schwach $\mathcal{M}$-abhängig mit*

$$Y_k^{(m)} := E[Y_k|\mathcal{A}_{k-\lfloor m/2\rfloor}^{k+\lfloor m/2\rfloor}]$$

bezüglich der gleichen p und $\delta(\cdot)$.

Lineare Prozesse Sei $\{Y_k\}_{k\in\mathbb{Z}}$ ein Prozess mit $Y_k = \sum_{j=-\infty}^{\infty} a_j \varepsilon_{k-j}$, wobei $(a_j)_{j\in\mathbb{Z}}$ reelle Zahlen mit $\sum_{j=-\infty}^{\infty} |a_j| < \infty$ sind und $\{\varepsilon_k\}_{k\in\mathbb{Z}}$ eine Folge von i. i. d. Zufallsvariablen ist. Dann nennt man $\{Y_k\}_{k\in\mathbb{Z}}$ <u>*linear*</u>.

Satz 83. *$\{Y_k\}_{k\in\mathbb{Z}}$ sei ein linearer Prozess. Ist $\|\varepsilon_0\|_p < \infty$, so ist $\{Y_k\}_{k\in\mathbb{Z}}$ schwach $\mathcal{M}$-abhängig in L^p mit*

$$Y_k^{(m)} := \sum_{j=-\lfloor m/2\rfloor}^{\lfloor m/2\rfloor} a_j \varepsilon_{k-j}.$$

Sei $\delta(\cdot)$ die Ratenfunktion von $\{Y_k\}_{k\in\mathbb{Z}}$.
Gilt zusätzlich $|a_j| \ll |j|^{-(A+1)}$ ($|j| \to \infty$) für ein $A > 1$, so gilt $\delta(m) \ll m^{-A}$. Gilt $|a_j| \ll \rho^{|j|}$ ($|j| \to \infty$) für $0 < \rho < 1$, so fällt $\delta(\cdot)$ sogar exponentiell.
Bemerkung 84. *ARMA(p,q)-Prozesse (vgl. Brockwell und Davis (1987), Definition 3.1.2, S. 78) liefern Beispiele für lineare Prozesse: Seien $\{\varepsilon_k\}_{k\in\mathbb{Z}}$ eine Familie von i. i. d. Zufallsvariablen, $a_1, \ldots, a_p, b_1, \ldots, b_q \in \mathbb{R}$ und $a(z) := \sum_{j=1}^{p} a_j z^j$ ein Polynom auf $\mathbb{C}$. Die definierende Gleichung*

$$Y_k - a_1 Y_{k-1} - \cdots - a_p Y_{k-p} = \varepsilon_k + b_1 \varepsilon_{k-1} + \cdots + b_q \varepsilon_{k-q}, \ k \in \mathbb{Z},$$

eines ARMA(p,q)-Prozess $\{Y_k\}_{k\in\mathbb{Z}}$, hat nach Brockwell und Davis (1987), Theorem 3.1.3 (S. 88), eine eindeutige stationäre Lösung

$$Y_k = \sum_{j=-\infty}^{\infty} c_j \varepsilon_{k-j}, \ k \in \mathbb{Z},$$

mit reellen Koeffizienten $(c_j)_{j\in\mathbb{Z}}$, wenn $a(z) \neq 0$ für $|z| = 1$ ist. Für AR(1)-Prozesse $Y_k = aY_{k-1} + \varepsilon_k$ mit $|a| < 1$ gilt sogar (siehe Brockwell und Davis (1987), Example 3.1.2, S. 79) $Y_k = \sum_{j=0}^{\infty} a^j \varepsilon_{k-j}$, $k \in \mathbb{Z}$. In diesem Fall nehmen die Koeffizienten der Darstellung sogar mit einer geometrischen Rate ab. An diesem Beispiel sieht man, dass schwache $\mathcal{M}$-Abhängigkeit auch Prozesse umfasst, die nicht stark mischend sind: AR(1)-Prozesse $Y_k = \frac{1}{2}Y_{k-1} + \varepsilon_k$ mit $\{\varepsilon_k\}_{k\in\mathbb{Z}}$ unabhängig und Bernoulli-verteilt sind nach Andrews (1984) nicht starkt mischend.

Nichtlineare Zeitreihen Sei $\{\varepsilon_k\}_{k\in\mathbb{Z}}$ eine Familie von i. i. d. Zufallsvariablen und G eine messbare Funktion. Wir betrachten den durch die Rekurrenz

$$Y_k = G(Y_{k-1}, \varepsilon_k), \ k \in \mathbb{Z}, \tag{5.24}$$

gegebenen stochastischen Prozess $\{Y_k\}_{k\in\mathbb{Z}}$. Nach Diaconis und Freedman (1999) (vgl. Wu (2005), S. 14152) hat (5.24) eine eindeutige stationäre Lösung, wenn es ein $x_0 \in \mathbb{R}$ und ein $\alpha > 0$ gibt, so dass G folgende Lipschitz-Bedingung erfüllt:

$$E[\log K(\varepsilon_0)] > 0 \ \text{ und } \ K(\varepsilon_0), |G(x_0, \varepsilon_0)| \in L^\alpha,$$

wobei

$$K(\varepsilon) := \sup_{x,x'\in\mathbb{R}, x\neq x'} \frac{|G(x,\varepsilon) - G(x',\varepsilon)|}{|x - x'|}.$$

Iterationen von (5.24) liefern dann (vgl. Wu und Shao (2004), Remark 2) die Darstellung

$$Y_k = f(\ldots, \varepsilon_{k-1}, \varepsilon_k), \ k \in \mathbb{Z}, \tag{5.25}$$

mit einer messbaren Funktion f. Eine solche Darstellung lässt sich interpretieren als physikalisches System, bei dem der Input $\{\varepsilon_k\}_{k\in\mathbb{Z}}$ durch die Funktion f in den Output Y_k transformiert wird (vgl. Wu (2005)). Um approximierende Zufallsvariablen $Y_k^{(m)}$ zu definieren, benutzen wir die von Wu (2005) eingeführte Kopplungsmethode. Seien $\{\varepsilon_k^{(l)}\}_{k\in\mathbb{Z}}$ für jedes $l \in \mathbb{Z}$ von einander und von $\{\varepsilon_k\}_{k\in\mathbb{Z}}$ unabhängige Folgen von i. i. d. Zufallsvariablen, die identisch verteilt sind wie $\{\varepsilon_k\}_{k\in\mathbb{Z}}$. Dann erfüllen die Zufallsvariablen

$$Y_k^{(m)} := f(\ldots, \varepsilon_{k-m-2}^{(k)}, \varepsilon_{k-m-1}^{(k)}, \varepsilon_{k-m}, \ldots, \varepsilon_k),\ \ k \in \mathbb{Z}, m \in \mathbb{N} \tag{5.26}$$

Bedingung (B) aus Definition 68 (vgl. Berkes et al. (2011), S. 2449) und nach Theorem 2 von Wu und Shao (2004) gibt es ein $r \in (0,1)$, so dass aufgrund der Stationarität

$$\|Y_k - Y_k^{(m)}\|_p \ll \exp(-\rho m)$$

für $\rho = -\log r$ gilt (vgl. Berkes et al. (2011), S. 2449; Wu (2005), S 14152).

Bemerkung 85. *Hörmann und Kokoszka (2010) betrachten in ihrem Artikel stochastische Prozesse $Y = \{Y_k\}_{k\in\mathbb{Z}}$ mit Werten in einem Funktionenraum $L^p([0,1], \mathcal{B}_{[0,1]}, \lambda_{[0,1]})$ ($p \in \mathbb{N}$). Als Sonderfall ergeben sich Folgen von konstanten Abbildungen Y_k, also reellen Zufallsvariablen in L^p. Im Artikel von Hörmann und Kokoszka (2010) findet sich ein konsistenter Schätzer für die asymptotische Varianz σ^2 für Prozesse, die eine Darstellung der Form* (5.25) *mit approximierenden Zufallsvariablen* (5.26) *haben. Die dort vorausgesetzte Approximationsgüte ist*

$$\sum_{m=1}^{\infty} \|Y_m - Y_m^{(m)}\|_p < \infty,$$

was im obigen Fall gegeben ist. Hörmann und Kokoszka (2010) bezeichnen Prozesse mit dieser Eigenschaft als $\underline{L^p\text{-}m\text{-}approximierbar}$ und betrachten Kernschätzer für σ^2 der Form

$$\hat{\sigma}^2 = \sum_{|j|\le q_n} \omega_{q_n}(j)\hat{\gamma}_j \tag{5.27}$$

mit

$$\hat{\gamma}_j = \frac{1}{n}\sum_{i=1}^{n-|j|}(Y_i - \bar{Y}_n)(Y_{i+|j|} - \bar{Y}_n).$$

Ist $\{Y_k\}_{k\in\mathbb{Z}} \in L^4$ ein L^4-m-approximierbarer Prozess und gilt

$$\frac{1}{n}\sum_{k,l=0}^{q_n}\sum_{r=1}^{n-1} |\mathrm{Cov}(Y_0(Y_k - Y_k^{(k)}), Y_r^{(r)} Y_{r+l}^{(r+l)})| \overset{n\to\infty}{\longrightarrow} 0, \tag{5.28}$$

so ist $\hat{\sigma}^2$ nach Theorem 4.1 von Hörmann und Kokoszka (2010) für passende Wahlen von q_n und ω_{q_n} konsistent. Ein analoger Schätzer für L^p-m-approximierbare Prozesse, bei denen (5.28) *nicht vorausgesetzt wird, findet sich in Chochola et al. (2013).*

5.2 Assoziierte Prozesse

Nachdem wir uns im letzten Abschnitt auf Prozesse mit eindimensionaler Parametermenge beschränkt haben, wenden wir uns nun Zufallsfeldern zu. Hierbei werden wir Zufallsfelder betrachten, die gemäß Bulinski und Shashkin (2007) einer der Definitionen von Assoziation entsprechen. Dies ist eine in der Literatur viel beschriebene Abhängigkeitsart (vgl. z. B. Esary et al. (1967), Joag-Dev und Proschan (1983), Balan (2005), Bulinski und Shashkin (2006)), für die bereits schwache und starke Invarianzprinzipien existieren. Dabei wird die Abhängigkeitsstruktur der Prozesse durch Forderungen an das Verhalten der Kovarianz charakterisiert. Im Folgenden sei $\mathbf{X} = \{X_\mathbf{t}\}_{\mathbf{t}\in T}$ ($T \subset \mathbb{R}^d$, $d \in \mathbb{N}$) ein stochastischer Prozess. Für eine Teilmenge $I \subset T$ bezeichnen wir mit X_I die Familie $\{X_\mathbf{t}\}_{\mathbf{t}\in I}$. Weiter sei $\mathcal{M}(n)$ für $n \in \mathbb{N}$ die Menge der beschränkten reellwertigen, koordinatenweise monoton wachsenden, Borel-messbaren Funktionen auf $\mathbb{R}^n$. Für Funktionen $f \in \mathcal{M}(|I|)$ meinen wir mit $f(X_I)$ die Anwendung von f auf einen Vektor $\vec{X_I}$ in $\mathbb{R}^{|I|}$, der sich ergibt, wenn man der Familie X_I eine Ordnung gibt. Kommen in einer Formel mehrere Funktionen $f_1(X_I)$, $f_2(X_I)$ usw. vor, so werden diese stets auf den gleichen Vektor $\vec{X_I}$ angewendet.

Definition 86. *(siehe Bulinski und Shashkin (2007), Definition 1.1.1, S. 3)* $\mathbf{X}$ *heißt assoziiert (geschrieben als* $\mathbf{X} \in \mathbf{A}$*), wenn für alle endlichen Teilmengen* $I \subset T$ *und* $f, g \in \mathcal{M}(|I|)$ *gilt:*

$$\mathrm{Cov}(f(X_I), g(X_I)) \geq 0$$

Diese Definition lässt sich abschwächen, indem man disjunkte Teilmengen von T betrachtet:

Definition 87. *(siehe Bulinski und Shashkin (2007), Definition 1.1.2, S. 3)* $\mathbf{X}$ *heißt positiv assoziiert (geschrieben als* $\mathbf{X} \in \mathbf{PA}$*), wenn für alle endlichen disjunkten Teilmengen* $I, J \subset T$ *und* $f \in \mathcal{M}(|I|)$, $g \in \mathcal{M}(|J|)$ *gilt:*

$$\mathrm{Cov}(f(X_I), g(X_J)) \geq 0$$

Die folgende Definition von negativer Assoziation liefert ein Analogon zur positiven Assoziation mit umgekehrtem Vorzeichen. Negative Assoziiertheit wurde von Joag-Dev und Proschan (1983) als weitere Form einer „negativen Abhängigkeit“ eingeführt, die gegenüber den bereits bekannten „negativen“ Abhängigkeitsarten den Vorteil hat, dass monoton wachsende Funktionen von disjunkten Mengen von negativ assoziierten Zufallsvariablen wieder negativ assoziert sind.

Definition 88. *(siehe Bulinski und Shashkin (2007), Definition 1.1.3, S. 3)* $\mathbf{X}$ *heißt negativ assoziiert (geschrieben als* $\mathbf{X} \in \mathbf{NA}$*), wenn für alle endlichen disjunkten Teilmengen* $I, J \subset T$ *und* $f \in \mathcal{M}(|I|)$, $g \in \mathcal{M}(|J|)$ *gilt:*

$$\mathrm{Cov}(f(X_I), g(X_J)) \leq 0$$

Bemerkung 89. 1. *Assoziierte stochastische Prozesse sind insbesondere auch positiv assoziiert: Wie in Esary et al. (1967) angemerkt, lassen sich Funktionen aus $\mathcal{M}(n)$, $n \in \mathbb{N}$, auch als Funktionen aus $\mathcal{M}(m)$ mit $m \geq n$ auffassen, die unabhängig von den zusätzlichen, zu $n+1,\ldots,m$ gehörigen, Parametern sind. Betrachtet man disjunkte Teilmengen $I, J \subset T$, so lassen sich Funktionen $f \in \mathcal{M}(|I|)$ und $g \in \mathcal{M}(|J|)$ folglich auch als Funktionen aus $\mathcal{M}(|I \cup J|) = \mathcal{M}(|I| + |J|)$ auffassen, und die positive Assoziation folgt aus der Assoziation.*

2. *Eine einzelne Zufallsvariable ist assoziiert. Außerdem ist die Vereinigung von unabhängigen Mengen von assoziierten (bzw. positiv oder negativ assoziierten) Zufallsvariablen assoziiert (bzw. positiv oder negativ assoziiert). Folglich ist eine Familie von unabhängigen Zufallsvariablen assoziiert und negativ assoziiert. (siehe Bulinski und Shashkin (2007), Theorem 1.1.8, S. 6)*

3. *Eine Familie $X = \{X_{\mathbf{t}}\}_{\mathbf{t} \in T}$ mit unendlichem Parameterraum ist genau dann assoziiert (bzw. positiv oder negativ assoziiert), wenn alle ihre endlichen Unterfamilien assoziiert (bzw. positiv oder negativ assoziiert) sind.*

Die folgenden Beispiele für assoziierte und positiv oder negativ assoziierte Prozesse wurden aus Bulinski und Shashkin (2007) entnommen.

Beispiel 90. *(siehe Bulinski und Shashkin (2007), Theorem 1.2.1, S. 17, Remark 1.2.4, Theorem 1.2.6, S. 20) Bulinski und Shashkin (2007) beweisen in Anlehnung an Pitt (1982) sowie Joag-Dev und Proschan (1983), dass normalverteilte Zufallsvariablen $X_1, \ldots, X_n$ ($n \in \mathbb{N}$) genau dann assoziiert sind, wenn $\mathrm{Cov}(X_i, X_j) \geq 0$ für alle $i, j \in \{1, \ldots, n\}$ und genau dann negativ assoziiert sind, wenn $\mathrm{Cov}(X_i, X_j) \leq 0$ für alle $i, j \in \{1, \ldots, n\}$. Ein Gauß'sches Feld $X = \{X_{\mathbf{t}}\}_{\mathbf{t} \in \mathbb{Z}^d}$ ist somit genau dann assoziiert, wenn es positiv assoziiert ist. Bekanntlich (vgl. Shiryaev (1996), Chapter II, §, Theorem 1, S. 301) besteht eine Familie von Gauß'schen Zufallsvariablen genau dann aus unabhängigen Zufallsvariablen, wenn die Zufallsvariablen unkorreliert sind. Folglich lässt sich sagen, dass X genau dann*

$$\begin{cases} \textit{assoziiert} \\ \textit{unabhängig} \\ \textit{negativ assoziiert} \end{cases} \quad \textit{ist, wenn} \quad \mathrm{Cov}(X_{\mathbf{s}}, X_{\mathbf{t}}) \begin{cases} \geq \\ = \\ \leq \end{cases} 0 \ \forall \mathbf{s}, \mathbf{t} \in \mathbb{Z}^d, \mathbf{s} \neq \mathbf{t}.$$

Nun betrachten wir eine Erweiterung der Definitionen von positiver und negativer Assoziation auf eine größere Prozessklasse. Dabei werden wir nicht mehr koordinatenweise wachsende Funktionen, sondern Lipschitz-Funktionen betrachten:

Definition 91. *(vgl. Bulinski und Shashkin (2007), Definition 1.5.1, S. 88) Eine Funktion $F : \mathbb{R}^n \to \mathbb{R}$ ($n \in \mathbb{N}$) heißt Lipschitz-Funktion, wenn*

$$Lip(F) := \sup_{\mathbf{x}, \mathbf{y} \in \mathbb{R}^n,\ \mathbf{x} \neq \mathbf{y}} \frac{|F(\mathbf{x}) - F(\mathbf{y})|}{\|\mathbf{x} - \mathbf{y}\|_1} < \infty,$$

wobei $\|\mathbf{x}\|_1 := \sum_{i=1}^n |x_i|$ die l^1-Norm auf $\mathbb{R}^n$ sei.

Die partiellen Lipschitz-Konstanten seien dann für $i = 1, \ldots, n$:

$$Lip_i(F) := \sup_{\substack{x_1,\ldots,x_n,y_i\in\mathbb{R}\\ x_i\neq y_i}} \frac{|F(x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_n) - F(x_1,\ldots,x_{i-1},y_i,x_{i+1},\ldots,x_n)|}{|x_i-y_i|}$$

Mit $BL(n)$ bezeichnen wir die Menge aller beschränkten Lipschitz-Funktionen $F : \mathbb{R}^n \to \mathbb{R}$ und es sei $BL := \bigcup_{n\in\mathbb{N}} BL(n)$.

Definition 92. *(siehe Bulinski und Shashkin (2007), S. 91, oder Cox und Grimmett (1984)) Für eine Familie $X = \{X_{\mathbf{t}}\}_{\mathbf{t}\in T}$ von Zufallsvariablen sind die Cox-Grimmett Koeffizienten von X:*

$$u_r := \sup_{\mathbf{k}\in\mathbb{Z}^d} \sum_{\mathbf{j}\in\mathbb{Z}^d,\|\mathbf{j}-\mathbf{k}\|\geq r} |\mathrm{Cov}(X_{\mathbf{k}}, X_{\mathbf{j}})|,\ r\in\mathbb{N}$$

Für $X \in \mathbf{PA}$ oder $X \in \mathbf{NA}$ mit $E[X_{\mathbf{j}}^2] < \infty$ für alle $\mathbf{j} \in \mathbb{Z}^d$, endliche disjunkte Teilmengen $I, J \subset T$ und Lipschitz-Funktionen $f : \mathbb{R}^{|I|} \to \mathbb{R}$ und $g : \mathbb{R}^{|J|} \to \mathbb{R}$ haben Bulinski und Shabanovich (1998) gezeigt (siehe Bulinski und Shashkin (2007), Theorem 1.5.3 und Theorem 1.5.17), dass X folgende Ungleichung erfüllt:

$$|\mathrm{Cov}(f(X_I), g(X_J))| \leq \sum_{\mathbf{i}\in I}\sum_{\mathbf{j}\in J} Lip_{\mathbf{i}}(f) Lip_{\mathbf{j}}(g) |\mathrm{Cov}(X_{\mathbf{i}}, X_{\mathbf{j}})|. \tag{5.29}$$

Sind die Cox-Grimmett Koeffizienten $(u_r)_{r\in\mathbb{N}}$ endlich und konvergieren für $r \to \infty$ gegen 0, so folgt aus (5.29), dass X für Mengen I und J mit $d(I, J) = r$ auch die Ungleichung (5.30) mit $\theta_r := u_r$ aus der folgenden Definition erfüllt. Dies führt zu folgender Verallgemeinerung von positiver und negativer Assoziation:

Definition 93. *(siehe Bulinski und Shashkin (2007), Definition 1.5.12, S. 94)* **X** *heißt (BL,θ)-abhängig, wenn es eine monoton fallende Nullfolge $\theta = (\theta_r)_{r\in\mathbb{N}}$ gibt, so dass für alle endlichen disjunkten Teilmengen $I, J \subset T$ mit $d(I, J) = r \in \mathbb{N}$ und alle Funktionen $f \in BL(|I|)$ und $g \in BL(|J|)$ gilt:*

$$|\mathrm{Cov}(f(X_I), g(X_J))| \leq Lip(f) Lip(g)(|I| \wedge |J|)\theta_r \tag{5.30}$$

Beispiel 94. *(siehe Bulinski und Shashkin (2007), S. 96) Seien $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ ($m \in \mathbb{N}$) in $\mathbb{Z}^d \setminus \{\underline{\mathbf{0}}\}$ und eine messbare Lipschitz-Funktion $F : \mathbb{R}^m \to \mathbb{R}$ mit $L := \sum_{i=1}^m Lip_i(F) < 1$ gegeben. Weiter sei $\{\varepsilon_{\mathbf{j}}\}_{\mathbf{j}\in\mathbb{Z}^d}$ ein i. i. d. Zufallsfeld mit $E[\varepsilon_{\underline{\mathbf{0}}}] = 0$ und $E[\varepsilon_{\underline{\mathbf{0}}}^2] = 1$. Dann betrachten wir das autoregressiv definierte Zufallsfeld $\{X_{\mathbf{j}}\}_{\mathbf{j}\in\mathbb{Z}^d}$ mit*

$$X_{\mathbf{j}} = F(X_{\mathbf{j}-\mathbf{v}_1}, \ldots, X_{\mathbf{j}-\mathbf{v}_m}) + \varepsilon_{\mathbf{j}},\ \mathbf{j}\in\mathbb{Z}^d.$$

Diese Gleichung hat nach Theorem 1.5.18 von Bulinski und Shashkin (2007) eine stationäre Lösung, die (BL,θ)-abhängig mit

$$\theta_r \leq L^{-1}(1-L)^{-2}\, 3^d \left(r^d L^{r/p} + \sum_{k=r}^{\infty} k^{d-1} L^{k/p} \right),\ r\in\mathbb{N},$$

ist, wobei $p := diam(V) := \max\{\|\mathbf{i} - \mathbf{j}\| : \mathbf{i}, \mathbf{j} \in V\}$.

Beispiel 95. *(Bulinski und Shashkin (2007), Theorem 1.5.17, S. 95) Ein Gauß'sches Zufallsfeld* $\{X_\mathbf{j}\}_{\mathbf{j} \in \mathbb{Z}^d}$ *ist* (BL, θ)*-abhängig mit* $\theta_r = u_r$, $r \in \mathbb{N}$, *wenn die Cox-Grimmett Koeffizienten* $(u_r)_{r \in \mathbb{N}}$ *endlich sind und gegen Null konvergieren.*

Nun zitieren wir das schwache Invarianzprinzip für die von uns betrachteten Assoziationsarten. Der folgende Satz wurde aus Bulinski und Shashkin (2007) entnommen. Die Beweise seiner Teilaussagen gehen zurück auf Newman und Wright (1982) (Theorem 1, Theorem 11), Bulinski und Keane (1996) (Theorem, S. 2906), Zhang und Wen (2001) (Theorem 1) und Shashkin (2003) (Theorem 1):

Satz 96. *(siehe Bulinski und Shashkin (2007), Theorem 1.5.1, S. 255)*
Sei $X = \{X_\mathbf{j}\}_{\mathbf{j} \in \mathbb{Z}^d}$ *ein quadratisch integrierbares, zentriertes Zufallsfeld. Es sei eine der folgenden Bedingungen erfüllt:*

(a) Es ist $d \leq 2$ *und* $X \in \mathbf{PA}$ *ist stationär und erfüllt* (1.3).

(b) $X \in \mathbf{A}$ *und* X *ist schwach stationär. Außerdem gibt es* $\delta > 0$ *und* $\lambda > 0$, *so dass*

$$\sup_{\mathbf{j} \in \mathbb{Z}^d} E[|X_\mathbf{j}|^{2+\delta}] < \infty$$

und für die Cox-Grimmett Koeffizienten $u_r = \mathcal{O}(r^{-\lambda})$ *für* $r \to \infty$ *gilt.*

(c) $X \in \mathbf{NA}$ *ist stationär.*

(d) X *ist stationär und* (BL, θ)*-abhängig mit* $\theta_r = \mathcal{O}(r^{-\lambda})$ *für ein* $\lambda > 3d$.

Dann erfüllt X *das schwache Invarianzprinzip.*

Bemerkung 97. *Bulinski und Shashkin (2007) beziehen sich in ihrem Theorem 1.5.1 auf die in Bemerkung 30 eingeführte Konvergenzdefinition und benutzen als Straffheitskriterium den in Bemerkung 36 definierten Stetigkeitsmodul. Wie in Bemerkung 30 gezeigt wurde, impliziert die Konvergenz „in der* U*-Topologie“ die Konvergenz in Verteilung, so wie wir sie in dieser Arbeit eingeführt haben.*

Bemerkung 98. *Für assoziierte Variablen betrachten Bulinski und Shashkin (2007) zwei konsistente Schätzer (vgl. Bulinski und Shashkin (2007), Theorem 5.1.1, S. 321, Theorem 5.1.4 ,S. 327, und Corollary 5.1.6, S. 331) für die asymptotische Varianz* σ^2 *(siehe* (1.3)*). Für* (BL, θ)*-abhängige Familien wurde in Bulinski (2004) ein konsistenter (vgl. Bulinski (2004), Theorem 2, S. 460) Schätzer eingeführt.*

Anhang

Beweis von Lemma 12. Sei $B \subset Y$ eine abgeschlossene Menge. Für beliebiges $k \in \{1, \dots, n\}$ ist f_k stetig und somit $f_k^{-1}(B)$ abgeschlossen in A_k. Per Definition der Spurtopologie gibt es also eine offene Menge $M_k \in \mathcal{T}_X$, so dass gilt:

$$f_k^{-1}(B) = A_k \setminus (M_k \cap A_k) = A_k \setminus M_k = A_k \cap (X \setminus M_k)$$

Da A_k und $X \setminus M_k$ abgeschlossene Mengen in X sind, ist also auch $f_k^{-1}(B)$ als Schnitt von abgeschlossenen Mengen abgeschlossen in X. Folglich gilt, dass auch

$$f^{-1}(B) = \bigcup_{k=1}^{n} f_k^{-1}(B)$$

als endliche Vereinigung abgeschlossener Mengen abgeschlossen ist. Somit ist f stetig. $\square$

Beweis von Bemerkung 17. Für Funktionen $x, y \in D[0,1]^d$ und einen Quadranten Q gilt

$$\lim_{\mathbf{s}\to\mathbf{t}, \mathbf{s}\in Q}(x(\mathbf{s}) + y(\mathbf{s})) = \lim_{\mathbf{s}\to\mathbf{t}, \mathbf{s}\in Q} x(\mathbf{s}) + \lim_{\mathbf{s}\to\mathbf{t}, \mathbf{s}\in Q} y(\mathbf{s}) = x_Q(\mathbf{t}) + y_Q(\mathbf{t})$$

und

$$\lim_{\mathbf{s}\to\mathbf{t}, \mathbf{s}\in Q} x(\mathbf{s})y(\mathbf{s}) = \lim_{\mathbf{s}\to\mathbf{t}, \mathbf{s}\in Q} x(\mathbf{s}) \lim_{\mathbf{s}\to\mathbf{t}, \mathbf{s}\in Q} y(\mathbf{s}) = x_Q(\mathbf{t})\ y_Q(\mathbf{t}),$$

und somit liefert Lemma 16 die Behauptung. $\square$

Beweis von Bemerkung 19. Per definitionem gilt für $\lambda \in \Lambda$, dass $\lambda(t_1, \dots, t_d) = (\lambda_1(t_1), \dots, \lambda_d(t_d))$, mit bijektiven, streng monoton wachsenden, stetigen Funktionen λ_p. Offensichtlich ist dann

$$\lambda^{-1}(t_1, \dots, t_d) = (\lambda_1^{-1}(t_1), \dots, \lambda_d^{-1}(t_d)).$$

Für alle $p = 1, \dots, d$ gilt $\lambda_p^{-1}(0) = 0$ und $\lambda_p^{-1}(1) = 1$. Außerdem sind die λ_p^{-1} streng monoton wachsend: Angenommen, es gäbe $p \in \{1, \dots, d\}$ und $0 \leq t_1 < t_2 \leq 1$ mit $\lambda_p^{-1}(t_1) \geq \lambda_p^{-1}(t_2)$. Dann würde aus der strengen Monotonie von λ_p folgen:

$$t_1 = \lambda_p(\lambda_p^{-1}(t_1)) \geq \lambda_p(\lambda_p^{-1}(t_2)) = t_2,$$

also ein Widerspruch. Analog lässt sich auch die Stetigkeit von λ_p^{-1} $(p = 1, \dots, d)$ zeigen: Angenommen, es gäbe $t, t_1, \dots \in [0,1]$ mit $t_n \to t$, aber $\lambda_p^{-1}(t_n) \not\to \lambda_p^{-1}(t)$.

Da $[0,1]$ kompakt ist, existiert dann eine Teilfolge $\lambda_p^{-1}(t_{n_k})$ und ein $\tilde{t} \in [0,1]$ mit $\lambda_p^{-1}(t_{n_k}) \to \lambda_p^{-1}(\tilde{t})$. Dann folgt aber aus der Stetigkeit von λ_p

$$t_{n_k} = \lambda_p(\lambda_p^{-1}(t_{n_k})) \overset{k\to\infty}{\longrightarrow} \lambda_p(\lambda_p^{-1}(\tilde{t})) = \tilde{t},$$

also $t_n \not\to t$. $\square$

Beweis von Lemma 21. zu (a): Offensichtlich ist $id : T \to T$, $id(\mathbf{t}) := \mathbf{t}$ in Λ enthalten. Es gilt:

$$\|id - id\|_\infty = 0$$

und

$$d_S(x_n, x) \leq \|x_n - x\, id\|_\infty = \|x_n - x\|_\infty,$$

also folgt die Behauptung.

zu (b): Da x stetig auf dem Kompaktum T ist, ist x nach dem Satz von Heine auch gleichmäßig stetig auf T. Aus $d_S(x_n, x) \overset{n\to\infty}{\longrightarrow} 0$ folgt nach Lemma 18, dass es eine Folge $(\lambda_n)_{n\in\mathbb{N}}$ in Λ gibt mit

$$\|\lambda_n - id\|_\infty \overset{n\to\infty}{\longrightarrow} 0 \text{ und } \|x - x_n\lambda_n\|_\infty \overset{n\to\infty}{\longrightarrow} 0. \qquad (\star)$$

Dann existieren nach Bemerkung 19 auch $\lambda_n^{-1} \in \Lambda$ und es gilt:

$$\|\lambda_n^{-1} - id\|_\infty = \sup\{|\lambda_n^{-1}(\mathbf{t}) - \mathbf{t}| : \mathbf{t} \in T\} = \sup\{|\lambda_n(\mathbf{t}) - \mathbf{t}| : \mathbf{t} \in T\} \overset{n\to\infty}{\longrightarrow} 0 \quad (\star\star)$$

und

$$\begin{aligned}
\|x_n - x\|_\infty &= \sup\{|x_n(\mathbf{t}) - x(\mathbf{t})| : \mathbf{t} \in T\} \\
&\leq \underbrace{\sup\{|\overbrace{x_n(\mathbf{t})}^{=x_n(\lambda_n(\lambda_n^{-1}(\mathbf{t})))} - x(\lambda_n^{-1}(\mathbf{t}))| : \mathbf{t} \in T\}}_{=\sup\{|x_n(\lambda_n(\mathbf{t}))-x(\mathbf{t})|:\mathbf{t}\in T\}\to 0,\text{ wg. }(\star)} \\
&\quad + \underbrace{\sup\{|x(\lambda_n^{-1}(\mathbf{t})) - x(\mathbf{t})| : \mathbf{t} \in T\}}_{\to 0,\text{ wg. }(\star\star)\text{ und }x\text{ glm. stetig}} \\
&\overset{n\to\infty}{\longrightarrow} 0
\end{aligned}$$

$\square$

Beweis von Bemerkung 33. Für eine Stetigkeitsstelle $x \in E$ von f gilt:

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall y \in E \text{ mit } d_E(x,y) < \delta :\ d_{E'}(f(x), f(y)) < \varepsilon$$

Aufgrund der Dreiecksungleichung und der Tatsache, dass $d_E(x,x) = 0 < \delta$, ist dies äquivalent zu

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall y, z \in E \text{ mit } d_E(x,y) < \delta \text{ und } d_E(x,z) < \delta :\ d_{E'}(f(y), f(z)) < \varepsilon$$

Offensichtlich gilt dies genau dann, wenn

$$\forall \varepsilon \in \mathbb{Q}_+\ \exists \delta \in \mathbb{Q}_+\ \forall y, z \in E \text{ mit } d_E(x,y) < \delta,\ d_E(x,z) < \delta :\ d_{E'}(f(y), f(z)) < \varepsilon$$

erfüllt ist. Umgekehrt gilt also für jede Unstetigkeitsstelle x von f:

$$\exists \varepsilon \in \mathbb{Q}_+ \ \forall \delta \in \mathbb{Q}_+ \ \exists y, z \in E \text{ mit } d_E(x,y) < \delta,\ d_E(x,z) < \delta :\ d_{E'}(f(y), f(z)) \geq \varepsilon$$

Folglich gilt für die Menge U der Unstetigkeitsstellen von f:

$$U = \bigcup_{\varepsilon \in \mathbb{Q}_+} \bigcap_{\delta \in \mathbb{Q}_+} U_{\varepsilon,\delta},$$

wobei

$$U_{\varepsilon,\delta} := \{x \in E :\ \exists y, z \in E :\ d_E(x,y) < \delta,\ d_E(x,z) < \delta \text{ und } d_{E'}(f(y), f(z)) \geq \varepsilon\}.$$

Die Mengen $U_{\varepsilon,\delta}$ sind offen, denn für $x \in U_{\varepsilon,\delta}$ gilt: Es gibt $y, z \in E$ mit $d_E(x,y) < \delta$, $d_E(x,z) < \delta$ und $d_{E'}(f(y), f(z)) \geq \varepsilon$. Definiere $\tilde{\delta} := \max\{d_E(x,y), d_E(x,z)\} < \delta$. Dann gilt für alle $\tilde{x} \in E$ mit $d_E(x, \tilde{x}) < \delta - \tilde{\delta}$, dass

$$d_E(\tilde{x}, y) \leq d_E(\tilde{x}, x) + d_E(x, y) < \delta - \tilde{\delta} + \tilde{\delta} = \delta$$

und analog auch $d_E(\tilde{x}, z) < \delta$. Somit ist $\tilde{x} \in U_{\varepsilon,\delta}$. Als offene Mengen sind die $U_{\varepsilon,\delta}$ Borel-messbar und somit ist U als abzählbare Kombination von Vereinigungen und Schnitten messbarer Mengen auch Borel-messbar. Schließlich ist damit die Menge U^c der Stetigkeitsstellen messbar. $\square$

Beweis von Bemerkung 39. Setze:

$$R := \begin{cases} 0, & \text{unter } H_0 \\ \delta I_{\{\mathbf{k}_0 < \mathbf{k} \leq \mathbf{m}_0\}}(\mathbf{k}), & \text{unter } H_A \end{cases}$$

Dann gilt:

$$\begin{aligned} X_{\mathbf{j}} - \bar{X} &= Y_{\mathbf{j}} + \mu + R - (n^{-d} \sum_{\mathbf{1} \leq \mathbf{k} \leq \mathbf{n}} (Y_{\mathbf{k}} + \mu + R)) \\ &= Y_{\mathbf{j}} + R - n^{-d} \sum_{\mathbf{1} \leq \mathbf{k} \leq \mathbf{n}} (Y_{\mathbf{k}} + R) \end{aligned} \qquad \square$$

Beweis von Lemma 41. Wir führen eine vollständige Induktion über d durch.
Induktionsanfang $d = 1$:

$$\sup_{s,t \in [0,1]} \left| \frac{\lfloor nt \rfloor - \lfloor ns \rfloor}{n} - (t-s) \right| \leq 2 \sup_{t \in [0,1]} \left| \frac{\lfloor nt \rfloor - nt}{n} \right| \leq \frac{2}{n} \to 0$$

Induktionsvoraussetzung: Für ein $d > 1$ gelte die Aussage für alle $\tilde{d} < d$.

Induktionsschritt:

$$
\begin{aligned}
&\sup_{\mathbf{s},\mathbf{t}\in[0,1]^d}\left|\prod_{i=1}^{d}\frac{\lfloor nt_i\rfloor-\lfloor ns_i\rfloor}{n}-\prod_{i=1}^{d}(t_i-s_i)\right|\\
&=\sup_{\mathbf{s},\mathbf{t}\in[0,1]^d}\left|\prod_{i=1}^{d-1}\frac{\lfloor nt_i\rfloor-\lfloor ns_i\rfloor}{n}\left\{\frac{\lfloor nt_d\rfloor-\lfloor ns_d\rfloor}{n}-(t_d-s_d)\right\}\right.\\
&\quad\left.+\left\{\prod_{i=1}^{d-1}\frac{\lfloor nt_i\rfloor-\lfloor ns_i\rfloor}{n}-\prod_{i=1}^{d-1}(t_i-s_i)\right\}(t_d-s_d)\right|\\
&\leq\underbrace{\sup_{\mathbf{s},\mathbf{t}\in[0,1]^{d-1}}\left|\prod_{i=1}^{d-1}\frac{\lfloor nt_i\rfloor-\lfloor ns_i\rfloor}{n}\right|}_{\leq 1}\sup_{s_d,t_d\in[0,1]}\left|\frac{\lfloor nt_d\rfloor-\lfloor ns_d\rfloor}{n}-(t_d-s_d)\right|\\
&\quad+\sup_{\mathbf{s},\mathbf{t}\in[0,1]^{d-1}}\left|\prod_{i=1}^{d-1}\frac{\lfloor nt_i\rfloor-\lfloor ns_i\rfloor}{n}-\prod_{i=1}^{d-1}(t_i-s_i)\right|\\
&\leq\underbrace{\sup_{s,t\in[0,1]}\left|\frac{\lfloor nt\rfloor-\lfloor ns\rfloor}{n}-(t-s)\right|}_{=o(1),\text{ nach I.A.}}\\
&\quad+\underbrace{\sup_{\mathbf{s},\mathbf{t}\in[0,1]^{d-1}}\left|\frac{[\lfloor n\mathbf{t}\rfloor-\lfloor n\mathbf{s}\rfloor]}{n^{d-1}}-[\mathbf{t}-\mathbf{s}]\right|}_{=o(1),\text{ nach I.V.}}\\
&=o(1)
\end{aligned}
$$

□

Beweis von Bemerkung 44. zu (a): Es gilt

$$\sum_{\lfloor n\mathbf{s}\rfloor<\mathbf{j}\leq\lfloor n\mathbf{t}\rfloor}(X_{\mathbf{j}}-\bar{X})=\sum_{\lfloor n\mathbf{s}\rfloor<\mathbf{j}\leq\lfloor n\mathbf{t}\rfloor}X_{\mathbf{j}}-\frac{[\lfloor n\mathbf{t}\rfloor-\lfloor n\mathbf{s}\rfloor]}{n^d}\sum_{\underline{\mathbf{1}}\leq\mathbf{j}\leq\underline{\mathbf{n}}}X_{\mathbf{j}};$$

also reicht es zu zeigen, dass

$$\sum_{\lfloor n\mathbf{s}\rfloor<\mathbf{j}\leq\lfloor n\mathbf{t}\rfloor}X_{\mathbf{j}}\overset{!}{=}\sum_{\boldsymbol{\varepsilon}\in\{0,1\}^d}(-1)^{d-\sum_{i=1}^{d}\varepsilon_i}\sum_{\underline{\mathbf{1}}\leq\mathbf{j}\leq\lfloor n(\mathbf{s}+\boldsymbol{\varepsilon}(\mathbf{t}-\mathbf{s}))\rfloor}X_{\mathbf{j}}.$$

Wir führen dazu eine vollständige Induktion über $d\in\mathbb{N}$ durch:

- Induktionsanfang $d=1$:

$$
\begin{aligned}
\sum_{j=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor}X_j&=\sum_{j=1}^{\lfloor nt\rfloor}X_j-\sum_{j=1}^{\lfloor ns\rfloor}X_j\\
&=(-1)^{1-1}\sum_{j=1}^{\lfloor n(s+(t-s))\rfloor}X_j+(-1)^{1-0}\sum_{j=1}^{\lfloor n(s+0\cdot(t-s))\rfloor}X_j
\end{aligned}
$$

- Induktionsvoraussetzung: Für ein $d > 1$ gelte die Aussage für $d-1$.
- Induktionsschritt $d-1 \mapsto d$: Für Vektoren $\mathbf{x} \in \mathbb{R}^d$ sei $\tilde{\mathbf{x}} := (x_1, \ldots, x_{d-1})$. Es gilt:

$$
\begin{aligned}
\sum_{\lfloor n\mathbf{s}\rfloor < \mathbf{j} \leq \lfloor n\mathbf{t}\rfloor} X_{\mathbf{j}} &= \sum_{\lfloor n\tilde{\mathbf{s}}\rfloor < \tilde{\mathbf{j}} \leq \lfloor n\tilde{\mathbf{t}}\rfloor} \sum_{j_d=1}^{\lfloor nt_d\rfloor} X_{\mathbf{j}} - \sum_{\lfloor n\tilde{\mathbf{s}}\rfloor < \tilde{\mathbf{j}} \leq \lfloor n\tilde{\mathbf{t}}\rfloor} \sum_{j_d=1}^{\lfloor ns_d\rfloor} X_{\mathbf{j}} \\
&\overset{I.V.}{=} \sum_{\tilde{\boldsymbol{\varepsilon}} \in \{0,1\}^{d-1}} (-1)^{d-1-\sum_{i=1}^{d-1} \tilde{\varepsilon}_i} \sum_{\mathbf{1} \leq \tilde{\mathbf{j}} \leq \lfloor n(\tilde{\mathbf{s}} + \tilde{\boldsymbol{\varepsilon}}(\tilde{\mathbf{t}} - \tilde{\mathbf{s}}))\rfloor} \sum_{j_d=1}^{\lfloor nt_d\rfloor} X_{\mathbf{j}} \\
&\quad - \sum_{\tilde{\boldsymbol{\varepsilon}} \in \{0,1\}^{d-1}} (-1)^{d-1-\sum_{i=1}^{d-1} \tilde{\varepsilon}_i} \sum_{\mathbf{1} \leq \tilde{\mathbf{j}} \leq \lfloor n(\tilde{\mathbf{s}} + \tilde{\boldsymbol{\varepsilon}}(\tilde{\mathbf{t}} - \tilde{\mathbf{s}}))\rfloor} \sum_{j_d=1}^{\lfloor ns_d\rfloor} X_{\mathbf{j}} \\
&= \sum_{\tilde{\boldsymbol{\varepsilon}} \in \{0,1\}^{d-1}} (-1)^{d-(\sum_{i=1}^{d-1} \tilde{\varepsilon}_i + 1)} \sum_{\mathbf{1} \leq \mathbf{j} \leq \lfloor n(\mathbf{s} + (\tilde{\boldsymbol{\varepsilon}}, \mathbf{1})(\mathbf{t} - \mathbf{s}))\rfloor} X_{\mathbf{j}} \\
&\quad + \sum_{\tilde{\boldsymbol{\varepsilon}} \in \{0,1\}^{d-1}} (-1)^{d-(\sum_{i=1}^{d-1} \tilde{\varepsilon}_i + 0)} \sum_{\mathbf{1} \leq \mathbf{j} \leq \lfloor n(\mathbf{s} + (\tilde{\boldsymbol{\varepsilon}}, \mathbf{0})(\mathbf{t} - \mathbf{s}))\rfloor} X_{\mathbf{j}} \\
&= \sum_{\boldsymbol{\varepsilon} \in \{0,1\}^{d}} (-1)^{d-\sum_{i=1}^{d} \varepsilon_i} \sum_{\mathbf{1} \leq \mathbf{j} \leq \lfloor n(\mathbf{s} + \boldsymbol{\varepsilon}(\mathbf{t} - \mathbf{s}))\rfloor} X_{\mathbf{j}}
\end{aligned}
$$

<u>zu (b):</u> Offensichtlich ist $h_n(\mathbf{s}, \mathbf{t}) = h(\frac{\lfloor n\mathbf{s}\rfloor}{n}, \frac{\lfloor n\mathbf{t}\rfloor}{n})$. Für $\tilde{\alpha} \leq |[\mathbf{t} - \mathbf{s}]| \leq 1 - \tilde{\beta}$ gilt:

$$
|[\mathbf{t} - \mathbf{s}]|(1 - |[\mathbf{t} - \mathbf{s}]|) \geq \tilde{\alpha}\tilde{\beta}
$$

Also erfüllt $\gamma := \min\{\tilde{\alpha}(1-\tilde{\alpha}), \tilde{\beta}(1-\tilde{\beta}), \tilde{\alpha}\tilde{\beta}\} > 0$ die Behauptung. Für den Beweis der gleichmäßigen Stetigkeit stellen wir zunächst fest, dass stetige Funktionen auf Kompakta auch gleichmäßig stetig sind, und es somit ausreicht zu zeigen, dass die Funktion h stetig ist. Die Funktion

$$
(\mathbf{s}, \mathbf{t}) \mapsto |[\mathbf{t} - \mathbf{s}]| = \left| \prod_{i=1}^{d} (t_i - s_i) \right|
$$

ist als Verknüpfung stetiger Funktionen stetig. Somit lässt sich h als auf abgeschlossenen Mengen abschnittsweise stetige Funktion definieren:

$$
h(\mathbf{s}, \mathbf{t}) := \begin{cases} \tilde{\alpha}(1 - \tilde{\alpha}), & 0 \leq |[\mathbf{t} - \mathbf{s}]| \leq \tilde{\alpha} \\ |[\mathbf{t} - \mathbf{s}]|(1 - |[\mathbf{t} - \mathbf{s}]|), & \tilde{\alpha} \leq |[\mathbf{t} - \mathbf{s}]| \leq 1 - \tilde{\beta} \\ \tilde{\beta}(1 - \tilde{\beta}), & 1 - \tilde{\beta} \leq |[\mathbf{t} - \mathbf{s}]| \leq 1 \end{cases}
$$

Dabei ist h im Schnitt der Mengen wohldefiniert und jeweils auf den Mengen $0 \leq |[\mathbf{t} - \mathbf{s}]| \leq \tilde{\alpha}$, $\tilde{\alpha} \leq |[\mathbf{t} - \mathbf{s}]| \leq 1 - \tilde{\beta}$ und $1 - \tilde{\beta} \leq |[\mathbf{t} - \mathbf{s}]| \leq 1$ stetig. Nach Lemma 12 ist h dann auch auf ganz $[0,1]^{2d}$ stetig. □

Beweis von Bemerkung 46. Sei $(\mathbf{s},\mathbf{t}) \in \tilde{A}$ beliebig. Da W ein Wiener Feld ist, ist der Zufallsvektor

$$(W(\underline{\mathbf{1}}), W(\mathbf{s}+\boldsymbol{\varepsilon}_{\mathbf{1}}(\mathbf{t}-\mathbf{s})), \dots, W(\mathbf{s}+\boldsymbol{\varepsilon}_{\mathbf{r}}(\mathbf{t}-\mathbf{s}))),$$

wobei $r = 2^d$ und $\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_r \in \{0,1\}^d$ paarweise verschieden sind, multivariat normalverteilt. Somit sind alle Linearkombinationen seiner Einträge normalverteilt, insbesondere auch $Y(\mathbf{s},\mathbf{t})$. Analog sind auch Vektoren der Form

$$(Y(\mathbf{s}_1,\mathbf{t}_1), \dots, Y(\mathbf{s}_k,\mathbf{t}_k))$$

für $k \in \mathbb{N}$ und $(\mathbf{s}_1,\mathbf{t}_1), \dots, (\mathbf{s}_k,\mathbf{t}_k) \in \tilde{A}$ multivariat normalverteilt. Aus $E[W(\mathbf{t})] = 0$ für alle $\mathbf{t} \in [0,1]^d$ folgt, dass $E[Y(\mathbf{s},\mathbf{t})] = 0$ ist. Nach Bemerkung 7 und wegen $[\mathbf{t}-\mathbf{s}] = \lambda^d((\mathbf{s},\mathbf{t}])$ ist

$$Y(\mathbf{s},\mathbf{t}) = \frac{W((\mathbf{s},\mathbf{t}]) - \lambda^d((\mathbf{s},\mathbf{t}])W((\underline{\mathbf{0}},\underline{\mathbf{1}}])}{\sqrt{\lambda^d((\mathbf{s},\mathbf{t}])(1-\lambda^d((\mathbf{s},\mathbf{t}]))}}.$$

Mit $E[W(A)W(B)] = \lambda^d(A \cap B)$ gilt somit:

$$E[Y(\mathbf{s},\mathbf{t})^2] = \frac{\lambda^d((\mathbf{s},\mathbf{t}]) - 2\lambda^d((\mathbf{s},\mathbf{t}])^2 + \lambda^d((\mathbf{s},\mathbf{t}])^2}{\lambda^d((\mathbf{s},\mathbf{t}])(1-\lambda^d((\mathbf{s},\mathbf{t}]))} = 1$$

Als stetige Transformation eines fast sicher stetigen Prozesses ist Y fast sicher stetig. □

Beweis von Lemma 51. **Vorüberlegung:** Für Mengen $A \subset B$ gilt:

$$0 \leq \sup_{\mathbf{x}\in B} |f(\mathbf{x})| - \sup_{\mathbf{x}\in A} |f(\mathbf{x})| \overset{!}{\leq} \sup_{\mathbf{x}\in B\setminus A} \inf_{\mathbf{y}\in A} |f(\mathbf{x}) - f(\mathbf{y})| \qquad (\star)$$

Es gilt nämlich:

$$\begin{aligned}
&\sup_{\mathbf{x}\in B} |f(\mathbf{x})| - \sup_{\mathbf{x}\in A} |f(\mathbf{x})| \\
=&\max\{\sup_{\mathbf{x}\in A} |f(\mathbf{x})|, \sup_{\mathbf{x}\in B\setminus A} |f(\mathbf{x})|\} - \sup_{\mathbf{y}\in A} |f(\mathbf{y})| \\
=&\max\{0, \sup_{\mathbf{x}\in B\setminus A} |f(\mathbf{x})| - \sup_{\mathbf{y}\in A} |f(\mathbf{y})|\}
\end{aligned}$$

und

$$\begin{aligned}
&\sup_{\mathbf{x}\in B\setminus A} |f(\mathbf{x})| - \sup_{\mathbf{y}\in A} |f(\mathbf{y})| \\
=&\sup_{\mathbf{x}\in B\setminus A} \inf_{\mathbf{y}\in A} \{|f(\mathbf{x})| - |f(\mathbf{y})|\} \\
\leq&\sup_{\mathbf{x}\in B\setminus A} \inf_{\mathbf{y}\in A} \{||f(\mathbf{x})| - |f(\mathbf{y})||\} \\
\leq&\sup_{\mathbf{x}\in B\setminus A} \inf_{\mathbf{y}\in A} \{|f(\mathbf{x}) - f(\mathbf{y})|\}
\end{aligned}$$

Sei $\varepsilon > 0$ beliebig. Da f auf dem Kompaktum $[0,1]^{2d}$ stetig ist, ist f dort auch gleichmäßig stetig und es existiert folglich ein $\delta > 0$, so dass für alle $(\mathbf{s},\mathbf{t}),(\tilde{\mathbf{s}},\tilde{\mathbf{t}}) \in [0,1]^{2d}$ mit

$$\|(\mathbf{s},\mathbf{t}) - (\tilde{\mathbf{s}},\tilde{\mathbf{t}})\| < \delta$$

gilt:

$$|f(\mathbf{s},\mathbf{t}) - f(\tilde{\mathbf{s}},\tilde{\mathbf{t}})| < \varepsilon$$

zu (a): Es ist $A \subset A_n$ für alle $n \in \mathbb{N}$.
Idee: Finde für alle $(\mathbf{s},\mathbf{t}) \in A_n \setminus A$ ein $(\tilde{\mathbf{s}},\tilde{\mathbf{t}}) \in A$ mit

$$\|(\mathbf{s},\mathbf{t}) - (\tilde{\mathbf{s}},\tilde{\mathbf{t}})\| < \delta.$$

Dann gilt

$$|f(\mathbf{s},\mathbf{t}) - f(\tilde{\mathbf{s}},\tilde{\mathbf{t}})| < \varepsilon,$$

also

$$\inf_{(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in A} |f(\mathbf{s},\mathbf{t}) - f(\tilde{\mathbf{s}},\tilde{\mathbf{t}})| \le \varepsilon$$

und folglich

$$\sup_{(\mathbf{s},\mathbf{t})\in A_n\setminus A} \inf_{(\tilde{\mathbf{s}},\tilde{\mathbf{t}})\in A} |f(\mathbf{s},\mathbf{t}) - f(\tilde{\mathbf{s}},\tilde{\mathbf{t}})| \le \varepsilon$$

Somit gilt dann:

$$\left| \sup_{(\mathbf{s},\mathbf{t})\in A_n} |f(\mathbf{s},\mathbf{t})| - \sup_{(\mathbf{s},\mathbf{t})\in A} |f(\mathbf{s},\mathbf{t})| \right| = \sup_{(\mathbf{s},\mathbf{t})\in A_n} |f(\mathbf{s},\mathbf{t})| - \sup_{(\mathbf{s},\mathbf{t})\in A} |f(\mathbf{s},\mathbf{t})| \overset{(\star)}{\le} \varepsilon$$

Aus technischen Gründen führen wir zunächst einige Konstanten ein. Die geforderten Beschränkungen werden im Laufe des Beweises für Abschätzungen gebraucht. Sei $0 < \tilde{\varepsilon}_1 < 1$ mit $\tilde{\varepsilon}_1 < \min\{1 - \sqrt[d]{a}, \delta, (b-a)/2^d\}$ und $0 < \tilde{\varepsilon}_2 < 1$ so, dass $\tilde{\varepsilon}_2 < \min\{b, b-a, \delta\}$. Außerdem sei $\tilde{\delta} > 0$ gegeben mit $\tilde{\delta} < \min\{\tilde{\varepsilon}_1^d(1-\tilde{\varepsilon}_1)^d, \tilde{\varepsilon}_2 b^{d-1}\}$.

Es gibt ein $n_0 \in \mathbb{N}$, so dass für alle $n \ge n_0$ gilt:

$$|a_n - a| < \tilde{\delta} \text{ und } |b_n - b| < \tilde{\delta} \qquad (\star\star)$$

Seien nun $n \ge n_0$ und $(\mathbf{s},\mathbf{t}) \in A_n \setminus A$ gegeben.
1. Fall: $a_n \le [\mathbf{t}-\mathbf{s}] < a < b$
Wir suchen nun $(\tilde{\mathbf{s}},\tilde{\mathbf{t}}) \in A$, der Form, dass

$$\tilde{\mathbf{t}} - \tilde{\mathbf{s}} = \mathbf{t} - \mathbf{s} + \mathbf{x},$$

mit $\mathbf{x} \ge \underline{\mathbf{0}}$. Dann gilt:

$$[\tilde{\mathbf{t}} - \tilde{\mathbf{s}}] = [\mathbf{t} - \mathbf{s} + \mathbf{x}] = [\mathbf{t} - \mathbf{s}] + \sum_{I\subset\{1,\dots,d\}, I\neq\emptyset} \prod_{i\in I} x_i \prod_{j\in I^c} (t_j - s_j)$$

Wir definieren die Menge

$$I_{\tilde{\varepsilon}_1} := \{i \in \{1, \ldots, d\} : 1 - \tilde{\varepsilon}_1 \geq t_i - s_i\}.$$

Es ist $I_{\tilde{\varepsilon}_1} \neq \emptyset$, denn sonst wäre

$$[\mathbf{t} - \mathbf{s}] > (1 - \tilde{\varepsilon}_1)^d \geq a.$$

Nun setzen wir:

$$x_i := \begin{cases} \tilde{\varepsilon}_1, & i \in I_{\tilde{\varepsilon}_1} \\ 0, & i \in I^c_{\tilde{\varepsilon}_1} \end{cases}$$

Für $j \in I_{\tilde{\varepsilon}_1}$ gilt dann $t_j - s_j \leq 1 - \tilde{\varepsilon}_1$, und folglich $s_j + (1 - t_j) \geq \tilde{\varepsilon}_1$. Falls $t_j \leq 1 - \tilde{\varepsilon}_1$ ist, so setzen wir

$$\tilde{t}_j := t_j + \tilde{\varepsilon}_1 \text{ und } \tilde{s}_j := s_j;$$

sonst setzen wir

$$\tilde{t}_j := 1 \text{ und } \tilde{s}_j := s_j - \overbrace{(\tilde{\varepsilon}_1 - (1 - t_j))}^{>0}.$$

Für $i \notin I_{\tilde{\varepsilon}_1}$ setzen wir

$$\tilde{t}_j := t_j \text{ und } \tilde{s}_j := s_j.$$

Dann gilt

$$\underline{\mathbf{0}} \leq \tilde{\mathbf{s}} \leq \tilde{\mathbf{t}} \leq \underline{\mathbf{1}}, \quad \tilde{\mathbf{t}} - \tilde{\mathbf{s}} = \mathbf{t} - \mathbf{s} + \mathbf{x}$$

und

$$\max\{\max\{|\tilde{s}_j - s_j|, |\tilde{t}_j - t_j|\} : j \in \{1, \ldots, d\}\} \leq \tilde{\varepsilon}_1 < \delta.$$

Außerdem gilt wegen $0 < \tilde{\varepsilon}_1 < 1$:

$$\begin{aligned}
[\tilde{\mathbf{t}} - \tilde{\mathbf{s}}] &= [\mathbf{t} - \mathbf{s}] + \sum_{I \subset \{1,\ldots,d\}, I \neq \emptyset} \overbrace{\prod_{i \in I} x_i \prod_{j \in I^c} (t_j - s_j)}^{\geq 0} \\
&\geq [\mathbf{t} - \mathbf{s}] + \underbrace{\prod_{i \in I_{\tilde{\varepsilon}_1}} x_i}_{= \tilde{\varepsilon}_1^{|I_{\tilde{\varepsilon}_1}|}} \underbrace{\prod_{j \in I^c_{\tilde{\varepsilon}_1}} (t_j - s_j)}_{\geq (1 - \tilde{\varepsilon}_1)^{d - |I_{\tilde{\varepsilon}_1}|}} \\
&\geq [\mathbf{t} - \mathbf{s}] + \tilde{\varepsilon}_1^d (1 - \tilde{\varepsilon}_1)^d \\
&\geq a_n + \tilde{\delta} \\
&\overset{(\star\star)}{\geq} a
\end{aligned}$$

Außerdem folgt aus der Wahl von $\mathbf{x}$:

$$\begin{aligned}[\tilde{\mathbf{t}} - \tilde{\mathbf{s}}] &= [\mathbf{t} - \mathbf{s}] + \sum_{I \subset I_{\tilde{\varepsilon}_1}, I \neq \emptyset} \overbrace{\prod_{i \in I} x_i}^{= \tilde{\varepsilon}_1^{|I|} \leq \tilde{\varepsilon}_1} \overbrace{\prod_{j \in I^c} (t_j - s_j)}^{\leq 1} \\ &\leq [\mathbf{t} - \mathbf{s}] + 2^d \tilde{\varepsilon}_1 \\ &\leq a + 2^d \tilde{\varepsilon}_1 \\ &\leq b\end{aligned}$$

2. Fall: $a < b < [\mathbf{t} - \mathbf{s}] \leq b_n$
Für ein $i \in \{1, \ldots, d\}$ setzen wir:

$$x_j = \begin{cases} -\tilde{\varepsilon}_2, & j = i \\ 0, & \text{sonst} \end{cases}$$

Dann seien $\tilde{\mathbf{t}} := \mathbf{t}$ und

$$\tilde{s}_j := \begin{cases} s_i + \tilde{\varepsilon}_2, & j = i \\ s_j, & \text{sonst} \end{cases}.$$

Es gilt $t_i - s_i \geq b \geq \tilde{\varepsilon}_2$, also

$$\underline{\mathbf{0}} \leq \tilde{\mathbf{s}} \leq \tilde{\mathbf{t}} \leq \underline{\mathbf{1}}, \quad \tilde{\mathbf{t}} - \tilde{\mathbf{s}} = \mathbf{t} - \mathbf{s} + \mathbf{x}$$

und

$$\max\{\max\{|\tilde{s}_j - s_j|, |\tilde{t}_j - t_j|\} : j \in \{1, \ldots, d\}\} = \tilde{\varepsilon}_2 < \delta.$$

Außerdem gilt:

$$\begin{aligned}[\tilde{\mathbf{t}} - \tilde{\mathbf{s}}] &= [\mathbf{t} - \mathbf{s}] - \tilde{\varepsilon}_2 \overbrace{\prod_{j \in \{1, \ldots, d\}, j \neq i} (t_j - s_j)}^{\geq b^{d-1}} \\ &\leq b_n - \tilde{\varepsilon}_2 b^{d-1} \\ &\leq b_n - \tilde{\delta} \\ &\leq b\end{aligned}$$

und

$$\begin{aligned}[\tilde{\mathbf{t}} - \tilde{\mathbf{s}}] &= [\mathbf{t} - \mathbf{s}] - \tilde{\varepsilon}_2 \overbrace{\prod_{j \in \{1, \ldots, d\}, j \neq i} (t_j - s_j)}^{\leq 1} \\ &\geq b - \tilde{\varepsilon}_2 \\ &\geq a\end{aligned}$$

zu (b): Der Beweis geht unter Beachtung von $A_n \subset A$ analog zu (a). $\square$

Beweis von Lemma 56. Seien $(\mathbf{s},\mathbf{t}),(\mathbf{s}+\mathbf{h},\mathbf{t}+\mathbf{f}) \in D$. Wir zeigen zunächst für $d=1$:

$$E[X(s,t)X(s+h,t+f)] \overset{!}{=} 1 - \frac{|h|+|f|}{2(t-s)(1-(t-s))} + \mathrm{o}(|h|+|f|),\ \ h,f \to 0$$

Es gilt:

$$\begin{aligned}
&E[X(s,t)X(s+h,t+f)]\\
=&\frac{1}{\sqrt{(t-s)(1-(t-s))}}\frac{1}{\sqrt{(t-s+f-h)(1-(t-s+f-h))}}\\
&\cdot E[B(t)B(t+f)-B(t)B(s+h)-B(s)B(t+f)+B(s)B(s+h)]\\
=:&\frac{1}{\sqrt{(t-s)(1-(t-s))}}\frac{1}{\sqrt{(t-s+f-h)(1-(t-s+f-h))}}\cdot g(s,t,h,f)
\end{aligned}$$

Dabei ist

$$\begin{aligned}
g(s,t,h,f) &= \min(t,t+f)-\min(t,s+h)-\min(s,t+f)+\min(s,s+h)\\
&\quad -t^2-tf+st+th+st+sf-s^2-sh\\
&=\min(t,t+f)-\min(t,s+h)-\min(s,t+f)+\min(s,s+h)\\
&\quad -(t-s)f+(t-s)h-(t-s)^2.
\end{aligned}$$

Wir untersuchen das Verhalten dieser Funktion für $h,f\to 0$. Da $s+\alpha\le t$ ist, können wir uns für $h,f<\frac{\alpha}{2}$ auf die Betrachtung folgender Fälle beschränken:

- $h,f\ge 0:\ \ s\le s+h\le t\le t+f$
- $h<0,f\ge 0:\ \ s+h\le s<t\le t+f$
- $h\ge 0,f<0:\ \ s\le s+h\le t+f<t$
- $h<0,f<0:\ \ s+h<s\le t+f<t$

Einsetzen und Umformen liefert

$$g(s,t,h,f)=(t-s)(1-(t-s))+R,$$

mit

$$R := \begin{cases} -(1-(t-s))h-(t-s)f, & h,f\ge 0\\ (t-s)(h-f), & h<0,f\ge 0\\ (1-(t-s))(f-h), & h\ge 0, f<0\\ (1-(t-s))f+(t-s)h, & h,f<0.\end{cases}$$

Beispielsweise im Fall $h,f\ge 0$ gilt:

$$\begin{aligned}
g(s,t,h,f) &= t-(s+h)-s+s-(t-s)f+(t-s)h-(t-s)^2\\
&=t-s-(t-s)^2-(1-(t-s))h-(t-s)f\\
&=(t-s)(1-(t-s))-(1-(t-s))h-(t-s)f
\end{aligned}$$

Zur Vereinfachung schreiben wir $a := t - s$. Nun führen wir für die Funktion

$$F(h, f) := \frac{1}{\sqrt{(a + f - h)(1 - a - f + h)}}$$

eine Taylor-Entwicklung an der Stelle $(0, 0)$ durch:

$$\begin{aligned} F(h, f) &= F(0, 0) + \frac{\partial F(0, 0)}{\partial h} h + \frac{\partial F(0, 0)}{\partial f} f + \mathrm{o}(|h| + |f|) \\ &= \frac{1}{\sqrt{a(1 - a)}} + \frac{1 - 2a}{2\sqrt{(a(1 - a))^3}} h - \frac{1 - 2a}{2\sqrt{(a(1 - a))^3}} f + \mathrm{o}(|h| + |f|) \end{aligned}$$

Setzt man dies in die Kovarianz ein und nutzt aus, dass $R = \mathrm{o}(1)$ für $h, f \to 0$, so erhält man:

$$\begin{aligned} &E[X(s, t)X(s + h, t + f)] \\ =& \left(\frac{1}{a(1 - a)} + \frac{1 - 2a}{2(a(1 - a))^2} h - \frac{1 - 2a}{2(a(1 - a))^2} f + \mathrm{o}(|h| + |f|) \right) \Big(a(1 - a) + R \Big) \\ =& 1 + \frac{1 - 2a}{2a(1 - a)} (h - f) + \frac{R}{a(1 - a)} + \frac{1 - 2a}{2(a(1 - a))^2} \underbrace{R(h - f)}_{=o(|h|+|f|)} + \mathrm{o}(|h| + |f|) \\ =& 1 + \underbrace{\frac{1 - 2a}{2a(1 - a)} (h - f) + \frac{R}{a(1 - a)}}_{:=(\star)} + \mathrm{o}(|h| + |f|) \end{aligned}$$

Einsetzen der möglichen Werte von R und Zusammenfassen liefert nun

$$(\star) = -\frac{|h| + |f|}{2a(1 - a)},$$

und somit die behauptete Gleichung. Beispielsweise im Fall $h \geq 0, f \geq 0$ ist $R = -(1 - a)|h| - a|f|$, also

$$\begin{aligned} (\star) &= \frac{1 - 2a}{2a(1 - a)} (|h| - |f|) - \frac{(1 - a)|h| + a|f|}{a(1 - a)} \\ &= \frac{(1 - 2a - 2 + 2a)|h| - (1 - 2a + 2a)|f|}{2a(1 - a)} \\ &= -\frac{|h| + |f|}{2a(1 - a)}. \end{aligned}$$

Analog zum Fall $d = 1$ zeigen wir nun für $d = 2$ und $(\mathbf{h}, \mathbf{f}) \to \underline{\mathbf{0}}$:

$$\begin{aligned} &E[X(\mathbf{s}, \mathbf{t})X(\mathbf{s} + \mathbf{h}, \mathbf{t} + \mathbf{f})] \\ \overset{!}{=}& 1 - \frac{(t_1 - s_1)(|h_2| + |f_2|) + (t_2 - s_2)(|h_1| + |f_1|)}{2[\mathbf{t} - \mathbf{s}](1 - [\mathbf{t} - \mathbf{s}])} + \mathrm{o}(|h_1| + |h_2| + |f_1| + |f_2|) \end{aligned}$$

Definiere

$$\bar{Z}(\mathbf{s},\mathbf{t}) := W(\mathbf{t}) - W(t_1,s_2) - W(s_1,t_2) + W(\mathbf{s})$$

und

$$\begin{aligned}
R :=& E\left[\bar{Z}(\mathbf{s},\mathbf{t})\bar{Z}(\mathbf{s}+\mathbf{h},\mathbf{t}+\mathbf{f})\right]\\
=& E\left[W(\mathbf{t})W(\mathbf{t}+\mathbf{f}) - W(\mathbf{t})W(t_1+f_1,s_2+h_2) - W(\mathbf{t})W(s_1+h_1,t_2+f_2)\right.\\
&+W(\mathbf{t})W(\mathbf{s}+\mathbf{h}) - W(t_1,s_2)W(\mathbf{t}+\mathbf{f}) + W(t_1,s_2)W(t_1+f_1,s_2+h_2)\\
&+W(t_1,s_2)W(s_1+h_1,t_2+f_2) - W(t_1,s_2)W(\mathbf{s}+\mathbf{h})\\
&-W(s_1,t_2)W(\mathbf{t}+\mathbf{f}) + W(s_1,t_2)W(t_1+f_1,s_2+h_2)\\
&+W(s_1,t_2)W(s_1+h_1,t_2+f_2) - W(s_1,t_2)W(\mathbf{s}+\mathbf{h}) + W(\mathbf{s})W(\mathbf{t}+\mathbf{f})\\
&\left.-W(\mathbf{s})W(t_1+f_1,s_2+h_2) - W(\mathbf{s})W(s_1+h_1,t_2+f_2) + W(\mathbf{s})W(\mathbf{s}+\mathbf{h})\right].
\end{aligned}$$

Dann gilt für $(\mathbf{s},\mathbf{t}),(\mathbf{s}+\mathbf{h},\mathbf{t}+\mathbf{f}) \in D$:

$$\begin{aligned}
&E[X(\mathbf{s},\mathbf{t})X(\mathbf{s}+\mathbf{h},\mathbf{t}+\mathbf{f})]\\
=&\frac{1}{\sqrt{[\mathbf{t}-\mathbf{s}](1-[\mathbf{t}-\mathbf{s}])}}\frac{1}{\sqrt{[\mathbf{t}+\mathbf{f}-(\mathbf{s}+\mathbf{h})](1-[\mathbf{t}+\mathbf{f}-(\mathbf{s}+\mathbf{h})])}}\cdot g(\mathbf{s},\mathbf{t},\mathbf{h},\mathbf{f}),
\end{aligned}$$

mit

$$\begin{aligned}
&g(\mathbf{s},\mathbf{t},\mathbf{h},\mathbf{f})\\
:=&E\left[\{\bar{Z}(\mathbf{s},\mathbf{t})-[\mathbf{t}-\mathbf{s}]W(\underline{\mathbf{1}})\}\{\bar{Z}(\mathbf{s}+\mathbf{h},\mathbf{t}+\mathbf{f})-[\mathbf{t}+\mathbf{f}-(\mathbf{s}+\mathbf{h})]W(\underline{\mathbf{1}})\}\right]\\
=&-[\mathbf{t}-\mathbf{s}]\underbrace{E[W(\underline{\mathbf{1}})\bar{Z}(\mathbf{s}+\mathbf{h},\mathbf{t}+\mathbf{f})]}_{=[\mathbf{t}+\mathbf{f}-(\mathbf{s}+\mathbf{h})]}-[\mathbf{t}+\mathbf{f}-(\mathbf{s}+\mathbf{h})]\underbrace{E[W(\underline{\mathbf{1}})\tilde{Z}(\mathbf{s},\mathbf{t})]}_{=[\mathbf{t}-\mathbf{s}]}\\
&+[\mathbf{t}-\mathbf{s}][\mathbf{t}+\mathbf{f}-(\mathbf{s}+\mathbf{h})]E[W(\underline{\mathbf{1}})W(\underline{\mathbf{1}})]+E\left[\bar{Z}(\mathbf{s},\mathbf{t})\bar{Z}(\mathbf{s}+\mathbf{h},\mathbf{t}+\mathbf{f})\right]\\
=&-[\mathbf{t}-\mathbf{s}][\mathbf{t}+\mathbf{f}-(\mathbf{s}+\mathbf{h})]+E\left[\bar{Z}(\mathbf{s},\mathbf{t})\bar{Z}(\mathbf{s}+\mathbf{h},\mathbf{t}+\mathbf{f})\right]\\
=&-[\mathbf{t}-\mathbf{s}][\mathbf{t}+\mathbf{f}-(\mathbf{s}+\mathbf{h})]+R\\
=&-[\mathbf{t}-\mathbf{s}]^2+[\mathbf{t}-\mathbf{s}]\{(t_1-s_1)(h_2-f_2)+(t_2-s_2)(h_1-f_1)\}\\
&-[\mathbf{t}-\mathbf{s}][\mathbf{h}-\mathbf{f}]+R\\
=&[\mathbf{t}-\mathbf{s}](1-[\mathbf{t}-\mathbf{s}])+[\mathbf{t}-\mathbf{s}]\{(t_1-s_1)(h_2-f_2)+(t_2-s_2)(h_1-f_1)\}\\
&-[\mathbf{t}-\mathbf{s}]+R+\mathrm{o}(|h_1|+|h_2|+|f_1|+|f_2|),
\end{aligned}$$

für $(\mathbf{h},\mathbf{f}) \to \underline{\mathbf{0}}$, wobei in den letzten zwei Umformungen die Identität

$$[\mathbf{x}+\mathbf{y}] = [\mathbf{x}] + x_1y_2 + x_2y_1 + [\mathbf{y}]$$

und die Tatsache, dass $[\mathbf{h}-\mathbf{f}] = o(|h_1|+|h_2|+|f_1|+|f_2|)$ ist, benutzt wurde. Zur Vereinfachung schreiben wir von nun an $\mathbf{a} := \mathbf{t}-\mathbf{s}$. Wir machen wieder eine Fallunterscheidung und nutzen aus, dass wegen $\underline{\alpha} \leq \mathbf{t}-\mathbf{s}$ $(\mathbf{h},\mathbf{f})$ klein genug gewählt werden können, dass auch $\mathbf{s}+\mathbf{h} \leq \mathbf{t}$ und $\mathbf{s} \leq \mathbf{t}+\mathbf{f}$ gilt. Betrachtet werden folgende Fälle:

$$\mathbf{h},\mathbf{f} \geq \underline{\mathbf{0}}: \ \mathbf{s} \leq \mathbf{s}+\mathbf{h} \leq \mathbf{t} \leq \mathbf{t}+\mathbf{f}$$

$$\mathbf{h} \geq \underline{\mathbf{0}}, \mathbf{f} < \underline{\mathbf{0}}: \ \mathbf{s} \leq \mathbf{s}+\mathbf{h} \leq \mathbf{t}+\mathbf{f} < \mathbf{t}$$
$$\mathbf{h} < \underline{\mathbf{0}}, \mathbf{f} \geq \underline{\mathbf{0}}: \ \mathbf{s}+\mathbf{h} < \mathbf{s} < \mathbf{t} \leq \mathbf{t}+\mathbf{f}$$
$$\mathbf{h}, \mathbf{f} < \underline{\mathbf{0}}: \ \mathbf{s}+\mathbf{h} \leq \mathbf{s} \leq \mathbf{t}+\mathbf{f} \leq \mathbf{t}$$
$$\mathbf{f} \geq \underline{\mathbf{0}}, h_1 \geq 0, h_2 < 0: \begin{cases} s_1 \leq s_1+h_1 \leq t_1 \leq t_1+f_1 \\ s_2+h_2 < s_2 < t_2 \leq t_2+f_2 \end{cases}$$
$$\mathbf{f} \geq \underline{\mathbf{0}}, h_1 < 0, h_2 \geq 0: \begin{cases} s_1+h_1 < s_1 < t_1 \leq t_1+f_1 \\ s_2 \leq s_2+h_2 \leq t_2 \leq t_2+f_2 \end{cases}$$
$$\mathbf{f} < \underline{\mathbf{0}}, h_1 \geq 0, h_2 < 0: \begin{cases} s_1 \leq s_1+h_1 \leq t_1+f_1 < t_1 \\ s_2+h_2 < s_2 \leq t_2+f_2 < t_2 \end{cases}$$
$$\mathbf{f} < \underline{\mathbf{0}}, h_1 < 0, h_2 \geq 0: \begin{cases} s_1+h_1 < s_1 \leq t_1+f_1 < t_1 \\ s_2 \leq s_2+h_2 \leq t_2+f_2 < t_2 \end{cases}$$
$$\mathbf{h} \geq \underline{\mathbf{0}}, f_1 \geq 0, f_2 < 0: \begin{cases} s_1 \leq s_1+h_1 \leq t_1 \leq t_1+f_1 \\ s_2 \leq s_2+h_2 \leq t_2+f_2 < t_2 \end{cases}$$
$$\mathbf{h} \geq \underline{\mathbf{0}}, f_1 < 0, f_2 \geq 0: \begin{cases} s_1 \leq s_1+h_1 \leq t_1+f_1 < t_1 \\ s_2 \leq s_2+h_2 \leq t_2 \leq t_2+f_2 \end{cases}$$
$$\mathbf{h} < \underline{\mathbf{0}}, f_1 \geq 0, f_2 < 0: \begin{cases} s_1+h_1 < s_1 < t_1 \leq t_1+f_1 \\ s_2+h_2 < s_2 \leq t_2+f_2 < t_2 \end{cases}$$
$$\mathbf{h} < \underline{\mathbf{0}}, f_1 < 0, f_2 \geq 0: \begin{cases} s_1+h_1 < s_1 \leq t_1+f_1 < t_1 \\ s_2+h_2 < s_2 < t_2 \leq t_2+f_2 \end{cases}$$
$$h_1 \geq 0, h_2 < 0, f_1 \geq 0, f_2 < 0: \begin{cases} s_1 \leq s_1+h_1 \leq t_1 \leq t_1+f_1 \\ s_2+h_2 < s_2 \leq t_2+f_2 < t_2 \end{cases}$$
$$h_1 < 0, h_2 \geq 0, f_1 < 0, f_2 \geq 0: \begin{cases} s_1+h_1 < s_1 \leq t_1+f_1 < t_1 \\ s_2 \leq s_2+h_2 \leq t_2 \leq t_2+f_2 \end{cases}$$
$$h_1 \geq 0, h_2 < 0, f_1 < 0, f_2 \geq 0: \begin{cases} s_1 \leq s_1+h_1 \leq t_1+f_1 < t_1 \\ s_2+h_2 < s_2 < t_2 \leq t_2+f_2 \end{cases}$$
$$h_1 < 0, h_2 \geq 0, f_1 \geq 0, f_2 < 0: \begin{cases} s_1+h_1 < s_1 < t_1 \leq t_1+f_1 \\ s_2 \leq s_2+h_2 \leq t_2+f_2 < t_2 \end{cases}$$

Einsetzen und Umformen liefert für R:

$$R = \begin{cases} [\mathbf{a}-\mathbf{h}], & \mathbf{h},\mathbf{f} \geq \underline{\mathbf{0}} \\ [\mathbf{a}+\mathbf{f}-\mathbf{h}], & \mathbf{h} \geq \underline{\mathbf{0}}, \mathbf{f} < \underline{\mathbf{0}} \\ [\mathbf{a}], & \mathbf{h} < \underline{\mathbf{0}}, \mathbf{f} \geq \underline{\mathbf{0}} \\ [\mathbf{a}+\mathbf{f}], & \mathbf{h},\mathbf{f} < \underline{\mathbf{0}} \\ (a_1-h_1)a_2, & \mathbf{f} \geq \underline{\mathbf{0}}, h_1 \geq 0, h_2 < 0 \\ a_1(a_2-h_2), & \mathbf{f} \geq \underline{\mathbf{0}}, h_1 < 0, h_2 \geq 0 \\ (a_1+f_1-h_1)(a_2+f_2), & \mathbf{f} < \underline{\mathbf{0}}, h_1 \geq 0, h_2 < 0 \\ (a_1+f_1)(a_2+f_2-h_2), & \mathbf{f} < \underline{\mathbf{0}}, h_1 < 0, h_2 \geq 0 \\ (a_1-h_1)(a_2+f_2-h_2), & \mathbf{h} \geq \underline{\mathbf{0}}, f_1 \geq 0, f_2 < 0 \\ (a_1+f_1-h_1)(a_2-h_2), & \mathbf{h} \geq \underline{\mathbf{0}}, f_1 < 0, f_2 \geq 0 \\ a_1(a_2+f_2), & \mathbf{h} < \underline{\mathbf{0}}, f_1 \geq 0, f_2 < 0 \\ (a_1+f_1)a_2, & \mathbf{h} < \underline{\mathbf{0}}, f_1 < 0, f_2 \geq 0 \\ (a_1-h_1)(a_2+f_2), & h_1 \geq 0, h_2 < 0, f_1 \geq 0, f_2 < 0 \\ (a_1+f_1)(a_2-h_2), & h_1 < 0, h_2 \geq 0, f_1 < 0, f_2 \geq 0 \\ (a_1+f_1-h_1)a_2, & h_1 \geq 0, h_2 < 0, f_1 < 0, f_2 \geq 0 \\ a_1(a_2+f_2-h_2), & h_1 < 0, h_2 \geq 0, f_1 \geq 0, f_2 < 0 \end{cases}$$

Beispielsweise im Fall $\mathbf{h},\mathbf{f} \geq \underline{\mathbf{0}}$ gilt:

$$\begin{aligned} R = {} & \mathbf{t} \wedge (\mathbf{t}+\mathbf{f}) - \mathbf{t} \wedge (t_1+f_1, s_2+h_2) - \mathbf{t} \wedge (s_1+h_1, t_2+f_2) + \mathbf{t} \wedge (\mathbf{s}+\mathbf{h}) \\ & - (t_1,s_2) \wedge (\mathbf{t}+\mathbf{f}) + (t_1,s_2) \wedge (t_1+f_1, s_2+h_2) + (t_1,s_2) \wedge (s_1+h_1, t_2+f_2) \\ & - (t_1,s_2) \wedge (\mathbf{s}+\mathbf{h}) - (s_1,t_2) \wedge (\mathbf{t}+\mathbf{f}) + (s_1,t_2) \wedge (t_1+f_1, s_2+h_2) \\ & + (s_1,t_2) \wedge (s_1+h_1, t_2+f_2) - (s_1,t_2) \wedge (\mathbf{s}+\mathbf{h}) \\ & + \mathbf{s} \wedge (\mathbf{t}+\mathbf{f}) - \mathbf{s} \wedge (t_1+f_1, s_2+h_2) - \mathbf{s} \wedge (s_1+h_1, t_2+f_2) + \mathbf{s} \wedge (\mathbf{s}+\mathbf{h}) \\ & = \underbrace{[\mathbf{t}] - t_1(s_2+h_2) - t_2(s_1+h_1) + [\mathbf{s}+\mathbf{h}]}_{=[\mathbf{t}-\mathbf{s}-\mathbf{h}]} \\ & - t_1 s_2 + t_1 s_2 + s_2(s_1+h_1) - s_2(s_1+h_1) - s_1 t_2 + s_1(s_2+h_2) + s_1 t_2 \\ & - s_1(s_2+h_2) + [\mathbf{s}] - [\mathbf{s}] - [\mathbf{s}] + [\mathbf{s}] \\ & = [\mathbf{a}-\mathbf{h}] \end{aligned}$$

R lässt sich schreiben als $[\mathbf{a}] + \tilde{R}$, mit

$$\tilde{R} = \begin{cases} -a_1h_2 - a_2h_1 + [\mathbf{h}], & \mathbf{h}, \mathbf{f} \geq \underline{\mathbf{0}} \\ a_1(f_2 - h_2) + a_2(f_1 - h_1) + [\mathbf{f} - \mathbf{h}], & \mathbf{h} \geq \underline{\mathbf{0}}, \mathbf{f} < \underline{\mathbf{0}} \\ 0, & \mathbf{h} < \underline{\mathbf{0}}, \mathbf{f} \geq \underline{\mathbf{0}} \\ a_1f_2 + a_2f_1 + [\mathbf{f}], & \mathbf{h}, \mathbf{f} < \underline{\mathbf{0}} \\ -a_2h_1, & \mathbf{f} \geq \underline{\mathbf{0}}, h_1 \geq 0, h_2 < 0 \\ -a_1h_2, & \mathbf{f} \geq \underline{\mathbf{0}}, h_1 < 0, h_2 \geq 0 \\ a_1f_2 + a_2(f_1 - h_1) + f_2(f_1 - h_1), & \mathbf{f} < \underline{\mathbf{0}}, h_1 \geq 0, h_2 < 0 \\ a_2f_1 + a_1(f_2 - h_2) + f_1(f_2 - h_2), & \mathbf{f} < \underline{\mathbf{0}}, h_1 < 0, h_2 \geq 0 \\ a_1(f_2 - h_2) - a_2h_1 - h_1(f_2 - h_2), & \mathbf{h} \geq \underline{\mathbf{0}}, f_1 \geq 0, f_2 < 0 \\ a_2(f_1 - h_1) - a_1h_2 - h_2(f_1 - h_1), & \mathbf{h} \geq \underline{\mathbf{0}}, f_1 < 0, f_2 \geq 0 \\ a_1f_2, & \mathbf{h} < \underline{\mathbf{0}}, f_1 \geq 0, f_2 < 0 \\ a_2f_1, & \mathbf{h} < \underline{\mathbf{0}}, f_1 < 0, f_2 \geq 0 \\ a_1f_2 - a_2h_1 - h_1f_2, & h_1 \geq 0, h_2 < 0, f_1 \geq 0, f_2 < 0 \\ a_2f_1 - a_1h_2 - h_2f_1, & h_1 < 0, h_2 \geq 0, f_1 < 0, f_2 \geq 0 \\ a_2(f_1 - h_1), & h_1 \geq 0, h_2 < 0, f_1 < 0, f_2 \geq 0 \\ a_1(f_2 - h_2), & h_1 < 0, h_2 \geq 0, f_1 \geq 0, f_2 < 0. \end{cases}$$

Nun entwickeln wir die Funktion

$$F(\mathbf{h}, \mathbf{f}) := \frac{1}{\sqrt{[\mathbf{a} + \mathbf{f} - \mathbf{h}](1 - [\mathbf{a} + \mathbf{f} - \mathbf{h}])}}$$

an der Stelle $(\mathbf{0}, \mathbf{0})$:

$$\begin{aligned} F(\mathbf{h}, \mathbf{f}) = {} & F(\underline{\mathbf{0}}, \underline{\mathbf{0}}) + \frac{\partial F(\underline{\mathbf{0}}, \underline{\mathbf{0}})}{\partial h_1} h_1 + \frac{\partial F(\underline{\mathbf{0}}, \underline{\mathbf{0}})}{\partial h_2} h_2 + \frac{\partial F(\underline{\mathbf{0}}, \underline{\mathbf{0}})}{\partial f_1} f_1 + \frac{\partial F(\mathbf{0}, \mathbf{0})}{\partial f_2} f_2 \\ & + \mathrm{o}(|h_1| + |h_2| + |f_1| + |f_2|) \\ = {} & \frac{1}{\sqrt{[\mathbf{a}](1 - [\mathbf{a}])}} + \mathrm{o}(|h_1| + |h_2| + |f_1| + |f_2|) \\ & + \frac{(1 - 2[\mathbf{a}])a_2}{2\sqrt{([\mathbf{a}](1 - [\mathbf{a}]))^3}} h_1 + \frac{(1 - 2[\mathbf{a}])a_1}{2\sqrt{([\mathbf{a}](1 - [\mathbf{a}]))^3}} h_2 \\ & - \frac{(1 - 2[\mathbf{a}])a_2}{2\sqrt{([\mathbf{a}](1 - [\mathbf{a}]))^3}} f_1 - \frac{(1 - 2[\mathbf{a}])a_1}{2\sqrt{([\mathbf{a}](1 - [\mathbf{a}]))^3}} f_2 \\ = {} & \frac{1}{\sqrt{[\mathbf{a}](1 - [\mathbf{a}])}} + \frac{(1 - 2[\mathbf{a}])(a_1(h_2 - f_2) + a_2(h_1 - f_1))}{2\sqrt{([\mathbf{a}](1 - [\mathbf{a}]))^3}} \\ & + \mathrm{o}(|h_1| + |h_2| + |f_1| + |f_2|) \end{aligned}$$

Um die nachfolgende Rechnung übersichtlicher zu machen, definieren wir

$$L := a_1(h_2 - f_2) + a_2(h_1 - f_1)$$

und

$$T := |h_1| + |h_2| + |f_1| + |f_2|.$$

Es gilt:

$$\begin{aligned}
&E[X(\mathbf{s},\mathbf{t})X(\mathbf{s}+\mathbf{h},\mathbf{t}+\mathbf{f})] \\
=&\left\{\frac{1}{[\mathbf{a}](1-[\mathbf{a}])} + \frac{(1-2[\mathbf{a}])L}{2([\mathbf{a}](1-[\mathbf{a}]))^2} + \mathrm{o}(T)\right\} \cdot g(\mathbf{s},\mathbf{t},\mathbf{h},\mathbf{f}) \\
=&\left\{\frac{1}{[\mathbf{a}](1-[\mathbf{a}])} + \frac{(1-2[\mathbf{a}])L}{2([\mathbf{a}](1-[\mathbf{a}]))^2} + \mathrm{o}(T)\right\} \cdot \{[\mathbf{a}](1-[\mathbf{a}]) + [\mathbf{a}]L + \tilde{R} + \mathrm{o}(T)\} \\
=&1 + \frac{(1-2[\mathbf{a}])L}{2[\mathbf{a}](1-[\mathbf{a}])} + \frac{L}{1-[\mathbf{a}]} + \frac{(1-2[\mathbf{a}])L^2}{2[\mathbf{a}](1-[\mathbf{a}])^2} \\
&+ \frac{\tilde{R}}{[\mathbf{a}](1-[\mathbf{a}])} + \frac{(1-2[\mathbf{a}])L\tilde{R}}{2([\mathbf{a}](1-[\mathbf{a}]))^2} + \mathrm{o}(T)
\end{aligned}$$

Unter Ausnutzung von $L^2 = \mathcal{O}(T) \cdot \mathrm{o}(1) = \mathrm{o}(T)$ und $L\tilde{R} = \mathcal{O}(T) \cdot \mathrm{o}(1) = \mathrm{o}(T)$, jeweils für $\mathbf{h},\mathbf{f} \to \underline{\mathbf{0}}$, gilt somit:

$$E[X(\mathbf{s},\mathbf{t})X(\mathbf{s}+\mathbf{h},\mathbf{t}+\mathbf{f})] = 1 + \frac{L + 2\tilde{R}}{2[\mathbf{a}](1-[\mathbf{a}])} + \mathrm{o}(T)$$

Einsetzen der Werte für $\tilde{R}$ liefert schließlich, dass

$$L + 2\tilde{R} = -a_1(|h_2| + |f_2|) - a_2(|h_1| + |f_1|) + \mathrm{o}(T),$$

und folglich die behauptete Entwicklung der Kovarianzfunktion. □

Beweis von Lemma 60. Zu (a): Seien $C > 0$ und $\varepsilon > 0$ beliebig. Zu zeigen ist:

$$\exists n_0 \in \mathbb{N}\ \forall n \geq n_0 :\ P(|X_n| \leq C) \leq \varepsilon$$

Da $Y_n = \mathcal{O}_P(1)$ gibt es ein $\tilde{C} > 0$, so dass

$$\exists n_0 \in \mathbb{N}\ \forall n \geq n_0 :\ P(|Y_n| \geq \tilde{C}) \leq \varepsilon.$$

Außerdem lässt sich ein solches n_0 so wählen, dass für alle $n \geq n_0$ gilt $a_n > C + \tilde{C}$. Somit gilt für $n \geq n_0$:

$$\begin{aligned}
P(|X_n| \leq C) &\leq P(X_n \leq C) \\
&= P(a_n \leq C + Y_n) \\
&\leq P(a_n \leq C + |Y_n|) \\
&= P(a_n \leq C + |Y_n|,\ |Y_n| \geq \tilde{C}) + P(a_n \leq C + |Y_n|,\ |Y_n| < \tilde{C}) \\
&\leq P(|Y_n| \geq \tilde{C}) + \underbrace{P(a_n \leq C + \tilde{C})}_{=0} \leq \varepsilon
\end{aligned}$$

Zu (b): Seien $C > 0$ und $\varepsilon > 0$ beliebig. Wegen $B_n \xrightarrow{P} \infty$ existiert ein $\tilde{n}_0 \in \mathbb{N}$, so dass für alle $n \geq \tilde{n}_0$ gilt

$$P(|B_n| \leq 2C) \leq \varepsilon/2.$$

$A_n \xrightarrow{P} 1$ liefert außerdem, dass es ein $\bar{n}_0 \in \mathbb{N}$ gibt, so dass für alle $n \geq \bar{n}_0$ gilt:

$$P\left(A_n \leq \frac{1}{2}\right) \leq P\left(|A_n - 1| \geq \frac{1}{2}\right) \leq \varepsilon/2$$

Für $n \geq \max\{\tilde{n}_0, \bar{n}_0\}$ gilt also:

$$\begin{aligned}
P(|X_n| \leq C) &= P(|A_n||B_n| \leq C) \\
&= P\left(A_n|B_n| \leq C, A_n > \frac{1}{2}\right) + P\left(|A_n||B_n| \leq C, A_n \leq \frac{1}{2}\right) \\
&\leq P\left(|B_n| \leq \frac{C}{A_n}, A_n > \frac{1}{2}\right) + P\left(A_n \leq \frac{1}{2}\right) \\
&\leq P(|B_n| \leq 2C) + P\left(A_n \leq \frac{1}{2}\right) \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

□

Beweis von Lemma 74. zu (i): Definiere $\varepsilon_n = \max_{1 \leq i \leq n} |\frac{i}{n} - \frac{s_i^2}{\sigma^2 n}|$ und sei $k(n) \in \{1, \ldots, n\}$ die Stelle, an der das Maximum angenommen wird. Also gilt:

$$\varepsilon_n = \left| \frac{k(n)}{n} - \frac{s_{k(n)}^2}{\sigma^2 n} \right| = \underbrace{\frac{k(n)}{n}}_{\leq 1} \left| 1 - \frac{s_{k(n)}^2}{\sigma^2 k(n)} \right|$$

Dann gilt $\varepsilon_n \to 0$ $(n \to \infty)$:

- 1. Fall: $\exists C \in \mathbb{N}\, \forall n \in \mathbb{N}: \; k(n) \leq C$
  Da die s_n^2 nichtfallend sind, gilt:

$$\varepsilon_n = \left| \frac{k(n)}{n} - \frac{s_{k(n)}^2}{\sigma^2 n} \right| \leq \underbrace{\frac{k(n)}{n}}_{\leq \frac{C}{n}} + \underbrace{\frac{s_{k(n)}^2}{\sigma^2 n}}_{\leq \frac{s_C^2}{\sigma^2 n}} \to 0 \; (n \to \infty)$$

- 2. Fall: $\forall C \in \mathbb{N}\, \exists n_0 \in \mathbb{N}: \; k(n_0) \geq C$
  Offensichtlich gilt $k(n) \leq k(n+1)$ für alle $n \in \mathbb{N}$, d.h. $k(n)$ ist monoton wachsend. Sei $\varepsilon > 0$ beliebig. Wegen (5.13) gilt $\frac{s_n^2}{\sigma^2 n} \to 1$, also gibt es ein $C \in \mathbb{N}$, so dass $|1 - \frac{s_C^2}{\sigma^2 C}| < \varepsilon$. Nach Annahme und aufgrund der Monotonie von $k(n)$ existiert ein $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$ gilt $k(n) \geq C$. Also:

$$\left| 1 - \frac{s_{k(n)}^2}{\sigma^2 k(n)} \right| < \varepsilon \quad \forall n \geq n_0$$

  Wegen $k(n)/n \leq 1$ impliziert dies die Behauptung.

zu (ii): Nach Definition von ε_n gilt für $i = 1, \dots, n$:

$$0 \leq s_i^2/\sigma^2 = i + (s_i^2/\sigma^2 - i) \leq i + \varepsilon_n n \leq n + \varepsilon_n n$$

Da $0 \leq i \leq n$ gilt somit

$$\max_{1 \leq i \leq n} |W_1(s_i^2/\sigma^2) - W_1(i)| \leq \sup_{0 \leq s \leq n} \sup_{0 \leq t \leq \varepsilon_n n} |W_1(s+t) - W_1(s)|.$$

Für $\delta > 0$ beliebig liefert eine Anwendung von Lemma 1.2.1 aus Csörgő und Révész (1981) (mit $h = \varepsilon_n n$, $T = n + \varepsilon_n n$, $v = \delta\, \varepsilon_n^{\alpha - 1/2}$ und $\varepsilon = 1$):

$$\begin{aligned} &P\big(\sup_{0 \leq s \leq n} \sup_{0 \leq t \leq \varepsilon_n n} |W_1(s+t) - W_1(s)| > \delta \sqrt{n} \varepsilon_n^\alpha\big) \\ &\leq \underbrace{\frac{C(n + \varepsilon_n n)}{\varepsilon_n n}}_{=C/\varepsilon_n + C} e^{-\delta^2 \varepsilon_n^{2\alpha - 1}/3} = \mathcal{O}\left(\frac{1}{\varepsilon_n} e^{-\delta^2 \varepsilon_n^{2\alpha-1}/3}\right) \end{aligned}$$

für ein $C > 0$. Wegen $0 < \alpha < 1/2$ gilt $2\alpha - 1 < 0$ und somit $\varepsilon_n^{2\alpha-1} \to \infty$. Außerdem gilt

$$\lim_{y \to \infty} \frac{1}{y^{1/(2\alpha-1)}} e^{-y\, \delta^2/3} = 0,$$

und somit

$$\mathcal{O}\left(\frac{1}{\varepsilon_n} e^{-\delta^2 \varepsilon_n^{2\alpha-1}/3}\right) = \mathrm{o}(1).$$

zu (iii): Nach Annahme 2 ist $\gamma_n = \mathrm{o}(1)$. Da $(t_n^2)_{n \in \mathbb{N}}$ nichtnegativ und monoton wachsend ist, gilt $0 \leq t_i^2 \leq t_n^2 \leq \gamma_n n$ für $i = 1, \dots, n$. Folglich gilt:

$$\max_{1 \leq i \leq n} |W_2(t_i^2/\sigma^2)| \leq \sup_{0 \leq t \leq \gamma_n n/\sigma^2} |W_2(t)|$$

Da $\{W_2(t)\}_{t \geq 0}$ o. E. ein rechtsstetiges Martingal ist, folgt aus der Doob-Ungleichung (vgl. Bauer (2001), Satz 46.4) für $\varepsilon > 0$:

$$P\left(\sup_{0 \leq t \leq \gamma_n n/\sigma^2} |W_2(t)| > \varepsilon \sqrt{n} \gamma_n^\alpha\right) \leq \frac{1}{\varepsilon \gamma_n^\alpha \sqrt{n}} E\left[|W_2(\gamma_n n/\sigma^2)|\right]$$

Mit der Cauchy-Schwarz Ungleichung und der Eigenschaft $EW(t)^2 = t$ folgt schließlich:

$$P\left(\sup_{0 \leq t \leq \gamma_n n/\sigma^2} |W_2(t)| > \varepsilon \sqrt{n} \gamma_n^\alpha\right) \leq \frac{1}{\sigma \varepsilon \gamma_n^\alpha \sqrt{n}} \sqrt{\gamma_n n} = \frac{\gamma_n^{1/2-\alpha}}{\sigma \varepsilon} \overset{n \to \infty}{\longrightarrow} 0$$

□

Literaturverzeichnis

R. J. Adler. *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes*, volume 12 of *Lecture Notes-Monograph Series*. Institute of Mathematical Statistics, Hayward, CA, 1990.

D. W. K. Andrews. Non-strong mixing autoregressive processes. *Journal of Applied Probability*, 21(4):930–934, 1984.

J. Antoch und M. Hušková. Tests and estimators for epidemic alternatives. *Tatra Mountains Mathematical Publications*, 7:311–329, 1996.

J. A. D. Aston und C. Kirch. Detecting and estimating changes in dependent functional data. *Journal of Multivariate Analysis*, 109:204–220, 2012.

A. Aue, S. Hörmann, L. Horváth, und M. Reimherr. Break detection in the covariance structure of multivariate time series models. *The Annals of Statistics*, 37 (6B):4046–4087, 2009.

R. M. Balan. A strong invariance principle for associated random fields. *The Annals of Probability*, 33(2):823–840, 2005.

H. Bauer. *Wahrscheinlichkeitstheorie*. de Gruyter, Berlin, 5. Auflage, 2001.

I. Berkes, S. Hörmann, und L. Horváth. The functional central limit theorem for a family of GARCH observations with applications. *Statistics and Probability Letters*, 78(16):2725–2730, 2008.

I. Berkes, S. Hörmann, und J. Schauer. Asymptotic results for the empirical process of stationary sequences. *Stochastic Processes and their Applications*, 119(4): 1298–1324, 2009.

I. Berkes, S. Hörmann, und J. Schauer. Split invariance principles for stationary processes. *The Annals of Probability*, 39(6):2441–2473, 2011.

P. J. Bickel und M. J. Wichura. Convergence criteria for multiparameter stochastic processes and some applications. *The Annals of Mathematical Statistics*, 42(5): 1656–1670, 1971.

P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Mathematical Statistics. Wiley, New York, 1968.

P. Billingsley. *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics. Wiley, New York, 3rd edition, 1995.

P. Billingsley. *Convergence of Probability Measures.* Wiley Series in Probability and Mathematical Statistics. Wiley, New York, 2nd edition, 2008.

P. J. Brockwell und R. A. Davis. *Time Series: Theory and Methods.* Springer Series in Statistics. Springer, New York, 1987.

B. E. Brodsky und B. S. Darkhovsky. *Nonparametric Methods in Change-Point Problems*, volume 243 of *Mathematics and Its Applications.* Kluwer Academic Publishers, Dordrecht, 1993.

A. V. Bulinski. Statistical version of the central limit theorem for vector-valued random fields. *Mathematical Notes*, 76(3):455–464, 2004.

A. V. Bulinski und M. S. Keane. Invariance principle for associated random fields. *Journal of Mathematical Sciences*, 81(5):2905–2911, 1996.

A. V. Bulinski und E. Shabanovich. Asymptotical behaviour for some functionals of positively and negatively dependent random fields. *Fundamentalnaya i Prikladnaya Matematika*, 4(2):479–492, 1998.

A. V. Bulinski und A. Shashkin. Strong invariance principle for dependent random fields. *Lecture Notes-Monograph Series*, 48:128–143, 2006.

A. V. Bulinski und A. Shashkin. *Limit Theorems for Associated Random Fields and Related Systems*, volume 10 of *Advanced Series on Statistical Science and Applied Probability.* World Scientific, Singapore, 2007.

O. Chochola, M. Hušková, Z. Prášková, und J. G. Steinebach. Robust monitoring of CAPM portfolio betas. *Journal of Multivariate Analysis*, 115:374–395, 2013.

J. T. Cox und G. Grimmett. Central limit theorems for associated random variables and the percolation model. *The Annals of Probability*, 12(2):514–528, 1984.

M. Csörgő und P. Révész. *Strong Approximations in Probability and Statistics.* Academic Press, New York, 1981.

M. Csörgő und L. Horváth. *Limit Theorems in Change-Point Analysis.* Wiley, Chichester, 1997.

P. Diaconis und D. Freedman. Iterated random functions. *SIAM Review*, 41(1): 45–76, 1999.

M. Döring. *Mehrdimensionale Change-Point-Schätzung mit U-Statistiken.* Dissertation, Technische Universität Dresden, 2007.

J. D. Esary, F. Proschan, und D. W. Walkup. Association of random variables, with applications. *The Annals of Mathematical Statistics*, 38(5):1466–1474, 1967.

S. Hörmann und P. Kokoszka. Weakly dependent functional data. *The Annals of Statistics*, 38(3):1845–1884, 2010.

M. Hušková. Estimators for epidemic alternatives. *Commentationes Mathematicae Universitatis Carolinae*, 36(2):279–291, 1995.

J. Jacod und A. N. Shiryaev. *Limit Theorems for Stochastic Processes*, volume 288 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 2nd edition, 2003.

D. Jarušková. Detection of transient change in mean – a linear behavior inside epidemic interval. *Kybernetika*, 47(6):866–879, 2011.

D. Jarušková und V. Piterbarg. Log-likelihood ratio test for detecting transient change. *Statistics and Probability Letters*, 81:552–559, 2011.

K. Joag-Dev und F. Proschan. Negative association of random variables, with applications. *The Annals of Statistics*, 11(1):286–295, 1983.

D. Khoshnevisan. *Multiparameter Processes: An Introduction to Random Fields*. Springer Monographs in Mathematics. Springer, New York, 2002.

C. Kirch. *Introduction to Change-Point Analysis*. Vorlesungsskript, Universität Kaiserslautern, 2008.

A. Klenke. *Wahrscheinlichkeitstheorie*. Springer, Berlin, 2006.

J. Komlós, P. Major, und G. Tusnády. An approximation of partial sums of independent rv'-s, and the sample df. I. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 32(1):111–131, 1975.

J. Komlós, P. Major, und G. Tusnády. An approximation of partial sums of independent rv's, and the sample df. II. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 34(1):33–58, 1976.

F. Lavancier. Processus empirique de fonctionnelles de champs gaussiens à longue mémoire. *Publications IRMA, Lille*, 63(IX):1–26, 2005. Preprint.

B. Levin und J. Kline. The cusum test of homogeneity with an application to spontaneous abortion epidemiology. *Statistics in Medicine*, 4(4):469–488, 1985.

M. A. Lifshits. Absolute continuity of functionals of "supremum" type for Gaussian processes. *Journal of Mathematical Sciences*, 27(5):3103–3112, 1984.

W. Liu und Z. Lin. Strong approximation for a class of stationary processes. *Stochastic Processes and their Applications*, 119(1):249–280, 2009.

P. Major. The approximation of partial sums of independent rv's. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 35(3):213–220, 1976.

P. Major. An improvement of Strassen's invariance principle. *The Annals of Probability*, 7(1):55–61, 1979.

G. Neuhaus. *Zur Theorie der Konvergenz stochastischer Prozesse mit mehrdi-*

mensionalem Zeitparameter. Dissertation, Westfälische Wilhelms-Universität zu Münster, 1969.

G. Neuhaus. On weak convergence of stochastic processes with multidimensional time parameter. *The Annals of Mathematical Statistics*, 42(4):1285–1295, 1971.

C. M. Newman. Normal fluctuations and the FKG inequalities. *Communications in Mathematical Physics*, 74(2):119–128, 1980.

C. M. Newman und A. L. Wright. Associated random variables and martingale inequalities. *Probability Theory and Related Fields*, 59(3):361–371, 1982.

V. I. Piterbarg. *Asymptotic Methods in the Theory of Gaussian Processes and Fields*, volume 148 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, 1996.

L. D. Pitt. Positively correlated normal variables are associated. *The Annals of Probability*, 10(2):496–499, 1982.

Yu. V. Prokhorov. Convergence of random processes and limit theorems in probability theory. *Theory of Probability and Its Applications*, 1:157–214, 1956.

A. P. Shashkin. Invariance principle for a (BL,θ)-dependent random field. *Russian Mathematical Surveys*, 58(3):617–618, 2003.

A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer, New York, 2nd edition, 1996.

M. L. Straf. Weak convergence of stochastic processes with several parameters. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, volume 2, pages 187–221, 1972.

V. Strassen. An invariance principle for the law of the iterated logarithm. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 3(3):211–226, 1964.

S. Timmann. *Repetitorium der Analysis: Teil 1*. Binomi, Springe, 3. Auflage, 2006.

B. S. Tsirel'son. The density of the distribution of the maximum of a Gaussian process. *Teoriya Veroyatnostei i ee Primeneniya*, 20(4):865–873, 1975.

M. J. Wichura. Inequalities with applications to the weak convergence of random processes with multi-dimensional time parameters. *The Annals of Mathematical Statistics*, 40(2):681–687, 1969.

W. B. Wu. Nonlinear system theory: Another look at dependence. *Proceedings of the National Academy of Sciences of the United States of America*, 102(40): 14150–14154, 2005.

W. B. Wu. Strong invariance principles for dependent random variables. *The Annals of Probability*, 35(6):2294–2320, 2007.

W. B. Wu und X. Shao. Limit theorems for iterated random functions. *Journal of Applied Probability*, 41(2):425–436, 2004.

Q. Yao. Tests for change-points with epidemic alternatives. *Biometrika*, 80:171–191, 1993.

L.-X. Zhang und J. Wen. A weak convergence for negatively associated fields. *Statistics and Probability Letters*, 53(3):259–267, 2001.

Erklärung

Hiermit versichere ich, die vorliegende Arbeit selbstständig angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht zu haben.
Köln, den 20. Dezember 2012